Sophia Fiedler

The Grammar-in-use of Direct Reported Thought in French and German

An Interactional and Multimodal Analysis

Dissertation submitted for the joint degree
Docteure ès Lettres at the Faculté des Lettres et Sciences Humaines,
Institut des Sciences du langage, Centre de Linguistique appliquée
Université de Neuchâtel
and
Doktorin der Philosophie at the Fakultät für Geisteswissenschaften,
Institut für Germanistik, Universität Hamburg

by Sophia Fiedler

Supervisors: Prof. Dr. Simona Pekarek Doehler and Prof. Dr. Wolfgang Imo
Jury Members: Prof. Dr. Elwys De Stefani, Prof. Dr. Natalia Filatkina
and Prof. Dr. Jan Lindström

Defended February 16, 2023

Acknowledgements

kinds of (linguistic and non-linguistic) matters; Yue, for being my faithful friend and office mate in Hamburg, despite all COVID-related obstacles; and Adam and Yotam, who often were my virtual officemates when the pandemic forced us all to stay home; Virginia, my writing buddy, who has become a dear friend over the past few years; and Juju, for sending me care packages, recommending non-linguistic books, cheering me up, making me laugh and forget about my thesis, and for being such a faithful friend.

I also want to thank my parents, who always believed I could write a thesis. Your interest and curiosity in what I do were a great motivation for me. A special thanks to my sister Robin, for the most motivating present in the world.

Finally, I want to thank Guillaume—for all the walks during which we discussed linguistic matters, for your presence at the neighboring desk in our study at home, where we have worked side by side for so many hours during lockdown, for not-working with me when it was necessary, for your shared enthusiasm for 'very good data,' for reading my whole manuscript from A to Z, for your constant encouraging words, *et pour croire profondément que les gens normaux n'ont rien d'exceptionnel*.

Inhaltsverzeichnis

1.   Introduction

Joey: "Have you told him how you feel?"
Phoebe: Yes. °hh (brief pause) Not out loud.
(Friends, Season 3, Episode 18, 08:27)

Ein solches Reden ist wahrhaft lautes Denken.
- Heinrich v. Kleist

In his famous *Tractatus Logico-Philosophicus*, Ludwig Wittgenstein writes:

> Die Sprache verkleidet den Gedanken. Und zwar so, dass man nach der äußeren Form des Kleides, nicht auf die Form des bekleideten Gedankens schließen kann; weil die äußere Form des Kleides nach ganz anderen Zwecken gebildet ist, als danach, die Form des Körpers erkennen zu lassen.[1] (Wittgenstein 2014 [1921]: 30)

This *Kleid* ('clothes'), as Wittgenstein calls it—namely, the language that disguises thought—could be, in other words, the utterance that speakers use to make a thought publicly available to their interlocutors. The original form of the thought becomes irrelevant because such an utterance is "constructed with quite another object" (see *infra*). Wittgenstein's description is exactly what the present work is concerned with: analyzing how speakers share their thoughts with others and how they "disguise" them, grammatically and bodily.

In this book, I investigate the grammatical format of direct reported thought[2] in talk-in-interaction. Analyzing a corpus of French and German video-recordings of everyday talk, I show that reported thought is a recurrent multimodal format across languages to implement social actions, and more specifically for taking a stance. I demonstrate that when speakers use direct reported thought in interaction, they exploit grammatical and bodily resources.

Drawing on Interactional Linguistics, I focus on the grammatical and bodily patterns that speakers use to frame talk as purported thoughts from the past. The main criteria determining my phenomenon of direct reported thought were (1) the introduction of the direct reported thought by a quotative that frames upcoming talk as thought,[3] (2) the embeddedness of the quotation into a larger narrative sequence, and (3) the depictive character of the quote. I thereby define "thought" as inner speech that remains exclusively available to the "thinker" unless it is made publicly available to others and framed as such through a quotative. Consequently, the present work exclusively investigates self-quotation (Golato 2002b), thus recognizing that "the

---

[1]   The English translation is as follows: "Language disguises the thought; so that from the external form of the clothes one cannot infer the form of the thought they clothe, because the external form of the clothes is constructed with quite another object than to let the form of the body be recognized." (Wittgenstein 1922: 63, translated by Charles Kay Ogden)

[2]   The term *direct reported thought* can describe (1) the entire grammatical pattern [quotative + quote] and (2) the actual content that speakers report as thought. In order to distinguish the two meanings, I will refer to the content as *purported thought*.

[3]   An exception from this clear framing are the French quotatives *j'étais là* ('I was there') and *je suis là* ('I am there').

thoughts, feelings, experiences, hopes and expectations of persons are generally treated as theirs to know and describe" (Heritage 2013: 559).

I chose the label "direct reported thought" instead of "enactment" for reasons of terminological exactitude. First, because I restrict myself to what is displayed as thought (not speech or hypothetical speech), the term of enactment seemed too large. Second, because I focus on a specific grammatical format—namely, *direct* reported thought, in contrast to *indirect* reported thought—only an equally grammatical term could mirror this exact format of a bipartite structure without a complementizer. And third, in contrast to Holt (2007), who chose the term "enactment" to convey the non-reporting character of the quote, I wanted to explicitly stress that speakers display their quotes *as if they might have been thought* throughout the narration in which the quotes are embedded. This choice is independent of any supposition of whether what is quoted has been thought, said, or neither of these.

In contrast to direct reported speech, direct reported thought has received little attention in (interactional) linguistic research. While clause-combining has been amply discussed for reporting formats more generally, their multimodal analysis (except for prosody) is still in its infancy. The present work addresses these two research gaps. I show that speakers use multiple resources while quoting their thoughts, as Excerpts 1.1 and 1.2 below illustrate: Linguistic resources like the quotatives *j'étais là* ('I was there') and *dacht so* ('thought like') introduce the quote and the response cries[4], bodily resources like shrugs or a gaze up, and prosodic resources including sudden changes in pitch and volume.

Ex. 1.1: *de la pub* 'advertisement' (Pauscaf_02_17, 08min20)

```
04 MAT: %ch'tais °là°   ↑o↓:h °putAIn;°
           I was there    oh     shit
    mat  %leans back-->l.05

05      @+en:COre-+@%
            again
    mat  @hand shrug@
         +gz up-->+
                -->%
```

Ex. 1.2: *des wird nix* 'that's not going to happen' (KAFKU_04, 26min10)

```
02 HAN:  =dacht so ja T+O+LL, @(0.5)@
           thought like yeah great
    han                  +gz up+gz left-->l.03
                                @shrug@

03      des wird +NIX,
         that's not going to happen
    han          -->+
```

In both data excerpts, the quotative is followed by a response cry: a curse in French, *oh putain* ('oh shit'), and an ironic positive evaluation in German, *ja toll* ('yeah great'). The quotes themselves are both short and accompanied by a gaze away and a shrug. The quote does not

---

[4]   Response cries have been defined by Goffman (1981) as "signs meant to be taken to index directly the state of the transmitter" (*ibid*.: 116).

have clausal status in either of the two excerpts, which complicates the syntactic categorization of quotative and quote. With the direct reported, the speakers take an affective stance toward preceding events that they reported before (not in transcript). These two short examples illustrate that the quotations are built similarly across both languages: A quotative introduces the quote (i.e., a response cry), and the speaker's bodily conduct contributes to the display of an affective stance. Together, these features form a "multimodal assembly" (Pekarek Doehler *et al.* 2021) for the same action formation—namely, for stance-taking.

The interactional and multimodal analysis of direct reported thought raises new analytic questions—among others, how the speaker's bodily conduct affects the syntactic analysis of quotatives and quotes. In a nutshell, this study addresses the following questions:

1) What are the typical formats speakers use for introducing direct reported thought?
2) What are the interactional functions that these formats and, more generally, the self-quotation of thought accomplish in everyday talk?
3) How do multimodal accomplishments allow us to differentiate between different actions that can be carried out with direct reported thought?
4) How does the multimodal nature of direct reported thought inform us about the grammatical description of the bipartite structure [quotative + quote] and how are the different resources laminated (Goodwin 2013, 2000)?

The research focus of the present book started with the action of reporting thought. In my German data, I found *(sich) denken* ('to think (to oneself)') in first-person singular to be the most frequent verb to introduce direct reported thought in telling sequences (present and past tense). In French, when searching for direct reported thought in a comparable sequential context, I found two quotatives to be the most frequent: *se dire* ('to say to oneself') in first-person singular (mainly in analytic past tense) and *être là* ('to be there') in first-person singular, rarely in present tense and exclusively in its synthetic past form *j'étais là*. A closer analysis reveals that *j'étais là* and *je suis là* are, in fact, hybrid formats that hover between projecting direct reported thought and speech. Despite this difference when compared to all other investigated quotatives, I included *j'étais là* and *je suis là* for three main reasons: (1) Because the present work starts from the action of reported thought, excluding *j'étais là* and *je suis là* would have meant excluding a format that can be used for similar actions as the German *ich denke (mir)* in present and past tense—namely, affective stance-taking. Because *j'étais là* and *je suis là* can still *also* introduce reported thought (even though out of an array of other possibilities), ensuring an action-based comparability of the German and French forms was the predominant argument for keeping *j'étais là* and *je suis là* in my collection. My analysis will show that context often allows a best guess of whether something has been said or thought. (2) The ambiguous nature of the introduced quote may bear a certain interactional potential as to let the recipient decide whether something has been said or thought, which, especially in narrations of complaints or strong negative assessments, may be a resource that speakers can exploit to diminish responsibility. (3) Finally, even though, grammatically speaking, *j'étais là* and *je suis là* are not, like the other investigated quotatives, complement-taking predicates, the quoting device is used in a similar bipartite structure of [quotative + quote]. The present work

aims to investigate whether these formal differences may be of minor importance given the parallels concerning action formation.

After the action-based approach of the data, the main criterion for the selection of quotatives of thought in both languages was frequency.[5] In German, the quotative *(und) ich so* ('and I like') is by far not as frequent as *ich denke (mir)*, *ich dachte (mir)*, or *ich habe (mir) gedacht* in my data (*n* = 16 compared to *n* = 188). Additionally, *(und) ich so* does not contain a verb and therefore cannot be conjugated in present and past tense, which is an important feature of all the quotatives I investigated. As I am interested in clause-combining, a verbless quotative would not allow me to raise the same syntactic questions, such as those regarding aspects of governance and dependency, syntactic marking, or the syntactic status of quotative and quote.

In my French data, *penser* ('to think') is the verb that I first assumed to report thought, in parallel to German *denken* ('to think'). When analyzing my data, however, I realized that French speakers use *penser* in first person singular, present, and past forms to introduce various actions *except for* direct reported thought; in interaction, *se dire* and *être là* in first-person singular present and past tense are used instead. Verbless quotatives such as *genre* or *comme* ('like') were excluded from my analyses for the same reasons as for the German *(und) ich so*: both quotatives cannot be conjugated according to tense or aspect and are rare in my data (*n* = 1 for standalone *genre*[6] and no occurrence of *comme*).[7]

The action of quoting thought forms the central research interest in the present work. As such, my aim is to find out why speakers publicly display themselves reporting *thought* in specific situations. It is not my goal to show that reported thought differs from reported speech; however, I will show that there are certain differences between these in interaction. I thereby do not try to find out whether what speakers *display* as thought has been thought "verbatim" nor whether what thinkers present as thought has not been said. The determination of these questions lies outside the scope of the present work and, more importantly, outside the methodological capacities that Conversation Analysis and Interactional Linguistics offer. Therefore, I conceptualize direct reported thought as a mere *display* that is understood as such by participants—and therefore also by analysts—because

> [w]hen an individual plays a part he implicitly requests his observers to take seriously the impression that is fostered before them. They are asked to believe that […], in general, matters are what they appear to be. (Goffman 1959: 28)

To investigate the phenomenon of interest, this book approaches direct reported thought, first, from a theoretical perspective. In Chapter 2, I outline the main ideas of a *grammar-in-use* that I apply in the present work. The notions of emergent grammar and projection thereby constitute the core of my explanations because they both move temporality in the foreground of the analysis of grammar in spoken language. Reflections on complementation, clause-combining

---

[5]   The fact that I only treat direct reported thought in this work lead to the exclusion of quotatives like *ich meine* ('I mean') in present and past tense or *ich sage* ('I say') in present and past tense. A closer investigation of the direct reporting format with these verbs during my collection-building process showed that with these quotatives, speakers seem to quote speech, not thought.

[6]   I do have, however, 4 occurrences of *genre* in combination with *j'étais là*.

[7]   In the case of *comme*, the low frequency may be due to it being more common in the French-speaking part of Canada (Blondeau & Moreno 2018; Dostie 2020; Levey *et al.* 2013).

and on the grammar-body-interface complete my description of the grammar in use of the reported thought format.

Chapter 3 more generally treats the form and function of direct reported thought. The chapter describes the bipartite structure [quotative + quote], the delimitation of the quote, multimodal features, and the interactional functions of direct reported thought. To do so, I primarily resort to prior research on direct reported speech, which has been investigated much more thoroughly than direct reported thought. The formal convergences—concerning syntactic, multimodal, and activity contextual features—between the two reporting formats allow such a parallelization. On an interactional level, however, certain divergences occur. When this is the case, such divergences are specifically addressed.

In Chapter 4, I present several arguments for a cross-linguistic analysis. I therefore summarize qualitative and quantitative evidence in favor of potentially universal features of everyday talk that arise from the inherent sociality of everyday talk-in-interaction.

Chapter 5 then gives an overview of the data that I use in the present work and my collection-building process, followed by a short introduction to Conversation Analysis and Interactional Linguistics.

The two analytic chapters, Chapters 6 and 7, are structured in a parallel way: After a brief introduction, I analyze the German quotatives and then the French quotatives. All analytic subsections contain quantifying results,[8] followed by a syntactic analysis and an interactional (i.e., sequential and multimodal) analysis of concrete data excerpts. Whereas Chapter 6 treats the past-tense quotatives in French and German, Chapter 7 analyzes the (narrative) present tense forms. Both chapters conclude with a comparative chapter, in which two French and two German data excerpts are juxtaposed to demonstrate the parallels in action formation and bodily conduct between the languages.

My thesis concludes with Chapter 8, in which I summarize and discuss my main findings and present new research perspectives that arose from the results of the work. I close with a description of the general impact that an interactional and multimodal analysis may have on the grammatical—and especially syntactic—description of the direct reported thought format in French and German.

---

[8]   To avoid confusion between a simple quantitative and a statistic analysis, I used the attribute *quantifying* instead of *quantitative* in the respective sections.

2.   Grammar in Interaction

Since the 1980s, linguists have walked new paths to explain grammar. The systematic study of spoken language suggests that grammar-in-use is an emergent, incrementally built system. This system is dynamic, constantly evolving, and it is negotiated and adapted among interlocutors in everyday talk (Hopper 1987, 2011; Ochs *et al.* 1996; Pekarek Doehler 2021b).[9] From this perspective, grammar is not a system of rules that, once established, remains stable, as suggested by traditional written grammars that stipulate normative rules for language use[10]; instead, grammar-in-use evolves in constant adaptation to the interactional necessities that emerge from humans being essentially social.

> [G]rammar as an outcome of lived sociality […] stands in a relatively intimate relation to social interaction. It is designed for interactional ends and as such must recon with the architecture and dynamics of turns, sequences, activities, participant frameworks, stances, trouble, expectations, contingencies, and other relevant interactional actualities. (Schegloff *et al.* 1996: 36)

In line with this description, I understand grammar as a temporal, constantly changing system growing out of everyday human sociality through routinized patterns that speakers recurrently use to interact.

Out of this conceptualization of grammar, the notion of 'grammar in interaction' evolved, investigating grammatical principles in everyday talk. Since the seminal book *Interaction and Grammar*, edited by Elinor Ochs, Emanuel A. Schegloff, and Sandra A. Thompson (1996), grammar in interaction has grown into a major research field within the CA and IL frameworks. Two notions that developed in reaction to structuralist and generative grammar approaches are essential to describe grammar in interaction: emergent grammar (Hopper 1987, 2011) and (grammatical and interactional) projection (Auer 2002), which foreground the temporal character of grammar-in-use (Mushin & Pekarek Doehler 2021).

Attributing a key role to temporality in language use, Mushin and Pekarek Doehler (2021) rightly point out that "[p]rojection is the counterpart of emergence in the temporal unfolding of talk and actions" (*ibid*.: 13). Their description clarifies that projection (i.e., the foreshadowing of what comes next) is based on underlying grammatical and interactional structures. Projection makes next items expectable for interlocutors, making emergence not an arbitrary assemblage of words, but a forward-moving, temporal process within a network of projectable items.

In the remainder of this chapter, I will first explain the concepts of emergent grammar and grammatical and interactional projection. Then, I will briefly outline the notion of clause combining, applying emergence and projection to concrete grammatical units of complex syntactic patterns. Finally, I will discuss how grammar and physical bodies are interrelated in everyday talk and how body movements may (not) fit a grammar description. I will conclude with a brief summary.

---

[9]   There are also branches of Generative Grammar that include spoken language in their theory, especially phonology. Because these approaches do not work with naturally occurring interaction, they necessarily remain abstract. For an overview of the syntax-prosody interface in Generative Grammar, see Selkirk (2011).

[10]   Such a normative description is essential for language learning and teaching.

2.1.   Emergent Grammar

The notion of emergent grammar emphasizes the temporality and interactionality of grammar in everyday talk.[11] Hopper (1987), who coined the approach of emergent grammar in linguistics, describes the notion of emergence as follows:

> [G]rammar […] like speech itself must be viewed as a real-time, social phenomenon, and therefore is temporal; its structure is always deferred, always in a process but never arriving, and therefore emergent; […] structure, or regularity, comes out of discourse and is shaped by discourse as much as it shapes discourse in an on-going process. Grammar is hence not to be understood as a pre-requisite for discourse […]. Its forms are not fixed templates, but are negotiable in face-to-face interaction in ways that reflect the individual speaker's past experience of these forms […]. (*ibid*.: 141f)

Emergence thus refers to the spontaneous assembly of words in interaction, which are patched together in permanent response to and negotiation with the interlocutor (see also Bybee & Hopper 2001: 3). The notion can therefore apply to only spoken, not written, language.

In taking everyday spoken language as a basis, linguists are confronted with irregularities, incompleteness, restarts, overlaps, etc. Emergent grammar does not aim at accounting for all these features of spoken language; instead, it focuses on the patterns that speakers repeatedly use in interaction until certain patterns become routinized. The goal is to find recurrent structures that "shape" and are "shaped by discourse" (see *supra*). Hopper (2011) argues against the idea of a grammatical system fully established at a point in time:

> Grammar at the conversational level is […] the result of spontaneously creating new combinations of forms. These forms are fragments of previously heard expressions assembled on the fly […]. They are improvised in much the same way that a jazz performance is improvised: publicly and jointly. (*ibid*.: 32)

Routinization is thus primarily governed by two principles: spontaneity and the interactive and publicly displayed formation of new structures so they can be used and reused by speakers.[12]

The spontaneous formation of certain linguistic structures "on the fly"—and in response to contingencies of ongoing interaction—is what Hopper (2011) calls *emergent*.[13] A classic example of such emergence is the format of increments (Auer 2007; Horlacher 2015; Imo

---

[11]   The *in situ* temporality of grammatical trajectories has also been conceptualized under the notion of *online syntax*. In his paper *Syntax als Prozess* ('syntax as process'), Auer (2005) proposes to establish a general theory for spoken language (in contrast to written language). He argues that a theory of spoken language is necessary, not as a normative grammar prescribing learners how to speak a language, but for describing its "specific forms of production and reception" (*ibid*.: 2, my translation). The author establishes three features that such a description of spoken language would need to contain (*cf. ibid.*): it must describe syntax as an incremental phenomenon produced in real time, must consider the dialogic character of spoken language, and must account for the fact that some (even complex) syntactic formats may be "more or less strongly patterned" (*ibid*., my translation).

[12]   Routinizations over time are often evidenced by diachronic research, but synchronic studies have described similar processes, demonstrating the formation of "new combinations of forms" (Ford 1993; Golato 2002a; Günthner & Imo 2003; Heinemann 2005; Imo 2011a, 2012b; Ochs *et al.* 1996; Pekarek Doehler *et al.* 2010; Pekarek Doehler 2016, 2021b).

[13]   In contrast to this, he refers to processes in which rules are formed over a long stretch of time as *emerging* (Hopper 2011: 27f; see also Linell 2013).

2011b, 2015; Luke *et al.* 2012; Stoenica 2020; Stoenica & Pekarek Doehler 2020).[14] Consider the following excerpt, taken from Stoenica and Pekarek Doehler (2020) on French relative clauses as increments (*ibid.*: 310):

Ex. 2.1: 'Que Romain il nous a parlé' [Corpus Pauscaf (Pause 18): 02min10–02min18]

```
01    DAV:    j'ai regardé le film Safe.
              'I've watched the movie Safe'

02            ×*(0.6)*×
      geb     ×raises eyebrows and wrinkles brow×
              *quick middle-distance look to left*
```


fig.1

```
03    DAV:    que       Romain il nous a    parlé    *une  fois.=
              PRO.REL-DO Romain  he us AUX tell.PP   DET  time
              'that Romain once told us about'
      geb                                            *quick middle-
                                                     distance look
                                                     to left->1.04

04    GEB:    =euh*avec [euh
                   'with'
              --->*looks at dav------------>>
```

In this excerpt, the increment—here, a relative clause—is added to David's informing (line 01) incrementally after a pause (line 02). The use of this format may be a response to the interlocutor's "thinking face" (Stoenica & Pekarek Doehler 2020: 310), which allows GEB to display his non-understanding of David's informing. GEB's facial expression seems to function as an "open class repair initiator" (*ibid.*) that leads to a specification of the movie *Safe* with the relative clause (line 03). The excerpt demonstrates, by emerging, that the use of the increment here depends on the contingencies of ongoing talk—namely, the interlocutor's facial expression.

Understanding grammar-in-use as emergent implies that it "never exists as such, but is always coming into being" (Hopper 1987: 148). In this sense, grammatical patterns are "movements toward structure which are often characterizable in typical ways" (*ibid.*). This means that grammar is a process of constant reshaping, reform(ulate)ing and resedimenting according to the contingencies arising from a dynamic context, which is why linguistic

---

[14] How increments contribute to the organization of interaction and how they evidence the online processing of speech has been extensively discussed in the volume on incrementation by Ono and Couper-Kuhlen (2007).

structures are and must be constantly negotiated. The driving force behind this constant change is that everyday talk is integral to human sociality. A grammar-in-use as a dynamic system relies on the idea that speakers do things with talk and that verbal resources evolve with the respective needs.

In interaction, interlocutors constantly create and respond to the contingencies that arise from preceding turns and sequences. Contingencies in interaction emerge by adapting talk to a recipient in conversation. If a speaker realizes while speaking that an interlocutor either does not have enough knowledge about the current topic or displays their lack thereof, the speaker can, in response, add information. In his article on the "interactional construction of a sentence," Goodwin (1979) illustrates how speakers respond to the "characteristics" of an interlocutor: "[A] speaker in natural conversation has the capacity to modify the emerging meaning of his sentence as he is producing it in accord with the characteristics of its current recipient" (*ibid.*: 104). This is nicely shown in Goodwin's classical example (*ibid.*: 99), as follows:

```
John:    . . , . . . . . .⌈Don, ,      ⌈Don
         I gave, I gave u⌊p smoking ci⌊garettes::. =

Don:     = Yea:h,

John:    . . . .⌈Beth              . . .⌈Ann
         I-uh: ⌊one-one week ago t'da:⌊y. acshilly
```

The speaker's gaze at his (unknowing) interlocutor Don indicates that the turn is incremented "in accord" with Don's epistemic state. After Don's signal of recipiency at line 02, John orients to Beth, his wife, who is a knowing recipient about the event but not about the anniversary (line 03). Finally, he turns to Ann, who is, again, an unknowing recipient. This example illustrates that an utterance may be carefully tailored to its recipients.[15]

IL often resorts to emergent grammar, or "a grammar which is not abstractly formulated and abstractly represented, but always anchored in the specific concrete form of an utterance" (Hopper 1987: 142). Following emergent grammar in anchoring linguistic research in spoken language, the present work further explores mainly the following three points related to direct reported thought:[16]

(1) In the direct reporting format, quotative and quote (be it clausal or not) are assembled "on the fly," without syntactic marking through a complementizer, thus eluding traditional grammatical descriptions of clause-combining. The recurrent use of quotatives framing direct reported thought may point toward an ongoing routinization of the quotative in everyday talk.

(2) Bodily reenactments introduced by the quotative seem to be as emergent as language. In my data, they are also recurrently used in (or as) quotes. Depending on the activity speakers are involved in, such as a telling, a bodily movement may better suit the activity (i.e., staging a narration) or action (i.e., taking an affective stance) than may language. The contingencies of

---

15   For an elaborate analysis, see Goodwin (1979: 99ff).
16   Note again that, throughout the present work, *direct reported thought* designates the grammatical format consisting of a quotative and a quote. This is not to be confused with what I call *purported thought*, which describes the content of the quote itself, or the reported material introduced by the quotative.

the interaction may thus also influence the resources speakers use to quote their purported (past) thoughts.

(3) When the quotative is in present tense, the reported thought may be uttered in a way that reflects the on-line emergence of the thought while speaking, thus staging a thinking process as emergent.

Before a detailed discussion of clause-combining, or, more generally, bipartite structures in interaction, it is necessary to first explain the basic concepts underlying the idea of a temporally unfolding grammar in everyday talk. I will therefore address, in the next two sections, the notions of on-line syntax and projection.

2.2.   Projection

Projection can be described as "the fact that an individual action or part of it foreshadows another" (Auer 2002: 1). Because interlocutors usually know how sequences, turns, and clauses are routinely organized, it is also possible for interlocutors to guess what will be projected. This knowledge is not actively or analytically used; it is rather the fact that there is a shared understanding and prior experience of how interaction works, which is why people (of the same speech community) have some knowledge about organizational mechanisms in speech in interaction.

In his seminal paper, Auer (2002) suggests a "mechanism of projection in interaction and grammar" (*ibid*.: 1), thereby suggesting a separation into grammatical and interactional projection.[17] He argues that "grammars of human languages provide interlocutors with sedimentated and shared ways of organising projection in interaction" (*ibid.*). The idea of grammar as "sedimentated and shared" organizational principles has much in common with Hopper's *emergent grammar* (see Section 2.1), in which grammatical structures are similarly understood as "combinations of […] prefabricated parts" (Hopper 1987: 144), which are built through repetition, forming "a kind of pastiche, pasted together in an improvised way out of ready-made elements" (*ibid.*). Projection completes the notion of emergent grammar because it suggests explanations for *how* (i.e., according to which mechanisms) these patterns may be pasted together.

A major part of grammatical projection is, according to Auer (2002: 9), syntactic projection. This projection of grammatical trajectories is only possible because syntactic structures do not work arbitrarily but along certain principles; after all, grammar is a major resource that speakers have for interacting. Syntactic patterns are mainly assembled through the four principles of "government (dependency), constituency, adjacency and serialization" (*ibid*.: 10). For the analysis of direct reported thought, the notions of government and constituency are of specific interest:[18]

---

[17]   Auer's (2002) *interactional projection* is similar to Schegloff's (1980) concept of *action projection* (see *infra*) in that they both allow speakers to project an upcoming action.

[18]   Adjacency regulates the distance between constituents of a structure, such as where a reflexive pronoun must be placed regarding the verb. *Serialization* means that there are certain rules concerning word order, such as the placement of particles. Because these principles play a minor role in direct reported thought, I do not treat them in detail here.

(1) *Government* (*dependency*) describes the relation between syntactic heads and the lexical items that depend on them. For instance, a preposition in German decides on the case of its obligatory noun phrase. Government is specifically relevant for direct reported thought because the quotatives *ich denke (mir)* ('I think (to myself)') and *je me dis* ('I say to myself') in present and past tense consist of a complement-taking predicate (CTP) that usually requires a complement. The quotative, especially in initial position, thus grammatically projects the quote (also) for dependency reasons.

(2) *Constituency* regulates (clause-)combining possibilities in a language from simple to complex combinations. Auer (2002) explains that, although a noun phrase (like 'the woman') can form a part of a more complex noun phrase comprising a relative clause (like 'who sings'), it cannot do so with an inflected (predicate) phrase—at least, not in German (*cf. ibid*.: 10). This mechanism is relevant for direct reported thought because it allows the quotative to project certain structures but not others. For example, the clause *ich denke* ('I think') can grammatically project either a complement clause, a pronominal adverb like *daran* ('of it'), or a prepositional phrase (*an einen guten Kaffee*, 'of a good coffee').[19] When *ich denke* is used as a quotative, the pronominal adverb and the prepositional phrase are no longer projectable nexts.

In a nutshell, within clauses or constituents, speakers can anticipate what comes next because they have knowledge of—and can rely on—certain linguistic structures of word order and the building of clauses and constituents on several levels of grammatical detail. Auer (2002) argues that it is this order within syntactic patterns that enables interlocutors to perceive projection or project subsequent talk. For direct reported thought in interaction, Auer's principles induce:

- for the quotative: that it is assembled in an 'expectable' way, meaning that word order, the placement of particles or reflexive pronouns, etc., are shared knowledge between interlocutors and that they are projected accordingly. For instance, *und dann* ('and then') projects a verb as the next constituent; after that, *und dann dachte* ('and then thought') projects a personal pronoun or a noun phrase; *und dann dachte ich* ('and then thought I') leaves the possibility of adding a reflexive pronoun *mir* ('to myself') or the quote. If the reflexive pronoun is added, *und dann dachte ich mir* ('and then thought I to myself'), other particles or adverbs can still follow (e.g., *noch* ('still'), *immer* ('always'), or *so* ('like')) before the quote is finally introduced.
- or the bipartite structure [quotative + quote]: that quotatives may grammatically project an empty slot that is expected to be filled. However, even if the quote is grammatically projected—for example, as the complement of a CTP—it may not fit the 'traditional' grammatical requirements of the projected complement. This means that, formally, the quote does not have to be a noun phrase or a complete syntactic unit; it can also consist of response cries or embodied conduct only. In this sense, a quote, regardless of its grammatical form, always fills an empty argument position of a predicate. However, this conceptualization entails an extension of Auer's principle of constituency (when CTPs are used as quotatives, they can also project non-lexical vocalizations and body movements) and a restriction of this principle (when CTPs are used as quotatives, they

---

[19]   On the comparability of these types of complementation, see Thompson (2002).

do not project pronominal adverbs or prepositional phrases as possible units for the empty argument position).

Grammatical projection consists of more than syntactic projection. Prosodic projection (Mushin & Pekarek Doehler 2021: 13) also plays a crucial role in interaction[20] even though it may be more ambiguous in what it projects than are syntactic patterns. For instance, with turn-final intonation, speakers can indicate an incipient end of a turn (Schegloff & Sacks 1973) but also further talk (Barth-Weingarten 2009).[21] Prosodic cues are often inseparable from those grammatical; so-called *projector-constructions* (Günthner 2011a; Pekarek Doehler 2011), for instance, usually syntactically project a constituent *and* are pronounced with continuative prosody. In such cases, it is impossible to determine whether grammatical or prosodic cues lead recipients not to interrupt the current speaker after the projector-construction.

The interplay of several projective forces is common in interaction, which is why, generally, the distinction of grammatical and interactional projection also remains complicated. Projection in interaction is similar to Schegloff's (1980) notion of *action projection*:

> First, a speaker projects the occurrence of some type of turn or action by mentioning either what he or she will do (for example, 'Let me ask you a question' […]) or what will be involved for the recipient (for example, 'Tell me something' […]). Second the projected turn or action does not occur in the same talk unit (for instance, the same sentence) but is replaced there either by the name of the action […], by a pronoun ('Lemme ask you this'), or by a 'dummy term' ('Lemme ask you something'). (*ibid*.: 107)

This conceptualization of projection mainly concerns action formation because speakers announce with a preliminary what action is about to follow. For instance, 'Did I tell you?' projects a telling, and 'Can I ask you something' projects a question. In contrast to Schegloff (1980), Auer (2002) describes less action-bound and more interaction-organizational matters.

Projection in interaction allows speakers to manage interaction-organizational mechanisms beyond the boundaries of a sentence and is therefore "based on interactional knowledge" (*ibid*.: 9) of interlocutors. The speakers rely on recurrent interactional structures that indicate how subsequent speech will be shaped: "[W]e recognise, in its context, the type of a particular utterance, and […] know how such a type of activity is handled in sequential terms" (*ibid*.). To give a simple example, when somebody asks a question, as a default mechanism, a response is expected. If a response does not follow the question, recipients are expected to account for this noticeable absence of a relevant next.

In his article, Auer (2002) acknowledges that the distinction of interactional projection and grammatical projection may be ambiguous (*ibid*.: 29). One crucial distinctive feature is, however, that a grammatical projection cannot project further than one syntactical unit, whereas interactional projection allows speakers to also project multi-unit turns (i.e., several syntactical units), like tellings.

---

[20]   Auer (2005) also mentions semantic projection. He describes that certain syntagmas assume, for instance, the number (and, in German, the grammatical case) of items, such as the preposition *zwischen* ('between') projecting two following noun phrases in dative case. Auer notes, however, that semantic roles may "differ in transparency" (*ibid*.: 5, my translation). He suggests that "*(dass) dem Linguisten die neue Grammatik -> gefällt* [lit. that the linguist the new grammar -> likes, SF] [is] more strongly projective than *(dass) Fritz Maria?->gefällt, liebt,...* [that Fritz Maria -> likes, loves, SF]" (*ibid*., original emphasis).

[21]   Prosody as a means of organizing turn-taking or structuring a multi-unit turn—for instance, as a list (Selting 2007)—may be at the interface of grammar and interaction (see also Sandager Sørensen 2021).

Applied to direct reported thought, the distinction of grammatical and interactional projection entails the following: From the vantage point of grammatical projection, a prepositioned quotative that contains a CTP projects an empty slot.[22] However, because quotatives can also be followed body quotes, which cannot be syntactically projected by a CTP, the notion of interactional projection also applies. Even if the quotative has clausal status, such as *da dacht ich mir* ('then I thought to myself'), it may be analyzed, in the specific activity context of tellings, as a device that does not project anything on a grammatical (clausal) level. Instead, the quotative simply "projects […] a rather precisely defined constituent" (Auer 2002: 16) on an actional level—namely, a quote. Grammatically, the quotative would then become a "pre-front field constituent" (*ibid.*), thus losing its status as CTP and functioning similar to a preliminary (Schegloff 1980). Seen like this, the quotative does not project a specific *syntactic* format and rather an "open slot" that speakers fill subsequently with the "rather precisely defined constituent" (see *supra*) of the quote, either verbally or through bodily means.

If, like Auer (2002) states, "syntax can be seen (among other things) as the historical result of a sedimentation and (partly normative) regularisation of certain interactional projection techniques" (*ibid.*: 36), we need to define what these "other things" encompass. In the specific case of direct quotation, it seems necessary to also include embodied conduct or non-lexical vocalizations since they are recurrently interactionally projected by quotatives.

Of course, the question of grammatical and interactional projection is not binary, rather it is a continuum with fuzzy boundaries. Throughout my qualitative analyses, the two projections allow me to explain the grammatical and interactional trajectory of direct reported thought in everyday talk. They provide a framework accounting for emergent patterns that 'traditional' grammar fails to account for, especially concerning the combining of clauses with non-linguistic constituents.

In the next section, I outline general points concerning clause-combining within the grammar-in-interaction framework, especially in IL.

2.3.   Clause-Combining in the Interactional Linguistics Framework

The study of clause-combining first calls for a clear description of the term *clause* and how it is defined in the present book. The clause[23] as a potential locus of action has received much interest in grammar in interaction (Laury *et al.* 2019). Couper-Kuhlen and Selting (2018) define a clause as grammatical unit "formed with a predicate, most often a verb or a verb complex, and the grammatical constituents or phrases that need to accompany it in order for it to be interpreted as complete" (*ibid.*: 359). In other words, a clause encompasses at least one predicate and its obligatory arguments; for example, the subject pronoun *she* and the object *the bird* in the clause *She saw the bird.* Concerning complement-taking predicates like *denken* ('to

---

[22]   This is not the case for the French *je suis là / j'étais là*, which is syntactically complete and does not project anything on a purely grammatical level.

[23]   I am aware that the term *clause* is problematic, as it has been established for English and may not always fit the German or French syntactic unit I refer to (for a discussion, see Couper-Kuhlen & Selting 2018: 359ff). Because an elaboration on the non-equivalence of the English *clause*, the German *Satz*, and the French *phrase* (both: 'sentence') goes beyond the scope of the present study, I will restrict myself to the definition by Couper-Kuhlen and Selting (2018: 539).

think'), this would mean that *ich denke* ('I think') forms a clause only when its empty argument position is filled, such as in a complement clause.[24]

To be able to analyze all facets of direct reported thought, I need to widen the above definition of *clause*. The pattern [quotative + quote] includes quotes that can take various forms, from a complete clause to a body movement. Consequently, the obligatory "grammatical constituents or phrases that need to accompany it [the predicate] in order for it to be interpreted as complete" (see *supra*) must include non-verbal constituents. In other words, whereas the analyzed quotatives are defined as clauses, the projected constituent (i.e., the quote) may not, since it is not restricted to linguistic material. This way, it is possible to consider all kinds of combinational patterns of the quotative (i.e., a matrix clause).

When clauses are combined, they form clause-combining patterns. The analysis of complex syntactic patterns is a major part of each language's grammar. *Clause-combining* is an umbrella term for a specific type of bipartite structure, including coordinative and subordinate structures. Subordinate structures can be further divided into [matrix clause + adverbial clause], [matrix clause + relative clause], and [matrix clause + complement clause]. In the present work, only the pattern [matrix clause + complement clause] is relevant.[25]

For a long time, clause-combining patterns have generally been described as being limited to the two formats of coordination and subordination (Lehmann 1988; Quirk *et al.* 1985).[26] Whereas *coordination* is a non-hierarchical clause-combining pattern, *subordination* is hierarchical, as the subordinate clause is dependent on the matrix clause (for a cross-linguistic analysis and a discussion of subordination in spoken language, see Laury & Okamoto 2011b).

The clear separability of the two clause-combining formats has, however, been questioned, especially with the investigation of spoken language. Haiman and Thompson (1984) suggest regarding coordination and subordination as extremes of a continuum of seven formal properties that "underlie the compositeness of the notion 'subordination'" (*ibid.*: 511; for German, see also Auer 1998: 16). Plank (1986) demonstrates the same fuzziness regarding clause combination for the direct-reported-speech format.

In typology, Foley and Van Valin Jr. (1984), and Matthiessen and Thompson (1988) also argue for a revision of the bipartite and hierarchical approach to complex syntactic structures. In both papers, the authors argue for a study of complex clauses in their *contexts of use*. Elaborating on Givón's (1980) semantically motivated binding theory,[27] Foley and Van Valin Jr. (1984) also stress the strong influence of the semantics of the main verb on the degree of

---

[24] Following Thompson (2002), just because *ich denke* can also form a complete clause with a prepositional object or a prepositional adverb does not mean that these patterns are used for the same actions.

[25] The combination of the [*je suis là / j'étais là* + clause] is not a combination of matrix clause and complement clause (for a discussion, see Section 6.3.2.2).

[26] Lehmann (1988) does, however, not exclude "the possibility of more than two clauses being linked at the same level" (*ibid.*: 2). A similar observation has been made already by Brøndal (1943), who states that direct reporting formats combine two clauses that are "linked by no proper syntactic relationship" (*ibid.*: 78, my translation).

[27] In his seminal paper, Givón (1980) established a "binding hierarchy" based on the semantics of the main verb. For example, if the main verb is a manipulative verb, "The more emotionally involved the subject/agent of a manipulative verb is with the outcome encoded in the complement clause, the lower the manipulative verb will be on the binding scale" (*ibid.*: 337). Givón already notes that the binding scale is "non-discrete and scalar rather than discrete and binary" (*ibid.*: 335).

"tightly linked syntactic configurations" (*ibid*.: 271). The semantics of the verb is, at least for spoken German, crucial for the (non-)use of a complementizer. For instance, clausal complements of verbs of epistemic commitment, like the German *wissen* ('to know'), are less tightly linked syntactically than are verbs of volition like *wollen* ('want') because the first can have a complement without a complementizer whereas the second cannot. Without being able to go further into detail, the semantics of *denken* are probably one reason why it systematically introduces its complement clauses without complementizers (see also Auer 1998; Givón 1980).

In Interactional Linguistics, Couper-Kuhlen and Selting (2018) regard the combination of clauses as *sentences* (*ibid*.: 361). In the present study, I conceptualize a combination of a clausal unit with its grammatically projected constituent simply as a *bipartite structure* or *bipartite format*. There are two main motivations behind this conceptualization:

(1) The present work understands grammar, and thus also clauses, clause-combining, and bipartite structures, as "emergent, contingent, and co-constructed" (Laury & Ono 2014: 564). The notion of *sentence* may suggest that larger grammatical units containing more than one clause are "pre-planned" (*ibid*.: 563) and that speakers and recipients orient to such a prefabricated complex unit in social interaction. Such a pre-planning would contradict the notion of bipartite structures as emerging in time and in accordance with the local contingencies of an interaction.

(2) Recipients orient to the clausal unit (Thompson & Couper-Kuhlen 2005) for action projection (Auer 2002; Günthner 2011a; Hopper & Thompson 2008; Pekarek Doehler 2011, 2021a) or turn-taking (Sacks *et al.* 1974). Grammatical and interactional projection also build, among other units, on clauses to make next elements (grammatically or interactionally) expectable for recipients (Auer 2002, 2005). This is why, according to Laury and Ono (2014), the notion of clause also must encompass "conventionalized (formulaic)" (*ibid*.: 563) clauses. For projection to be successful, it is necessary that "the set of means used for projecting continuation [is] conventionalized, to the degree that it is used by speakers and recognized by the other participants, and in that sense, grammatical" (*ibid*.: 563). Laury and Ono (2014) explicitly state that these projected units can be verbal or non-verbal and that these different modalities "[interact] rather closely" (*ibid*.: 586).

To account for the wide array of clause-combining types that the direct reported thought format encompasses, I adopt a wide conceptualization of clause-combining. I thereby follow Gast and Diessel (2012), who define clause-combining as "the combination of a clause with some other constituent" (*ibid*.: 4). Other than a continuum ranging from subordination to coordination, this definition also accounts for instances in which a CTP (i.e., the quotative) is combined with body movement. In this clause-combining type, the grammatical projection of the CTP is not fulfilled by what fits the traditional description of a complement (clause) but by what Lerner subsumed under the term "utterance type": "the placement of the quote marker provides a *form* of an already projectable utterance type" (Lerner 1991: 446, my emphasis).

The form of the projected utterance type, here the quote, is adapted to the interactional needs of interlocutors in the ongoing conversation. With clauses forming the "primary vehicle for implementing social actions" (Thompson 2019: 254), speakers may, when combining clauses, adapt the trajectories of these malleable structures to the local contingencies. Consequently,

and in line with the emergent grammar approach,[28] I see clauses and clause-combining as "pasted together in an improvised way out of ready-made elements" (Hopper 1987: 144). Depending on the contingencies of the ongoing interaction, speakers may complete the "ready-made element" of the CTP-quotative with whatever unit or element fits these contingencies. The predominant prepositioning of the quotative[29] (on the preference for post-positioned complements, see Schmidtke-Bode & Diessel 2017) contributes to the possibility of the quotative routinely accomplishing this projective work.

A notion that encompasses the characteristics of quotatives and their projective potential presented above is the notion *social action format* (Fox 2007; Schegloff 1996; on the emergent grammar of clause-combining, see Maschler *et al.* 2020). Fox's (2007) concept of *social action format* is based on Schegloff's (1996) notion of *positionally sensitive grammar* and describes recurrent grammatical formats that speakers use to accomplish distinct actions in specific sequential contexts. In their paper on requests, Fox and Heinemann (2016) define social action formats as

> relations between linguistic form and social action: that is, how speakers design their utterances to be recognizable as doing a particular action (action formation), and how recipients recognize that action, or ascribe that action to the utterance (action ascription […]). Crucial to this line of research has been the notion of format; today the term is most commonly used to capture a recurrent linguistic practice or formulation for accomplishing a particular action. […] Social action format was meant to capture growing evidence that actions are often accomplished in recurrent and highly patterned grammatical 'packages' that can be discovered using the data and methods of Conversation Analysis and Interactional Linguistics. (*ibid.*: 2f)

As Fox and Heinemann state, the concept of social action formats is closely related to action formation and action ascription (Levinson 2013), notions that focus on the recognizability and identifiability of actions in everyday talk. Well-researched social action formats investigated specifically under the aspect of action formation and ascription include requests (Fox & Heinemann 2016; Goodwin 1980), assessments (Goodwin & Goodwin 1987; Lindström & Mondada 2009), and storytellings (Goodwin 1984; Mandelbaum 2013).

Due to their potential high degree of variation, social action formats can be difficult to pinpoint. The present work sets out to explore if quotatives that introduce reported thought could be defined as social action format. My data suggests that a recurrent format to report thought exists: A prepositioned quotative that routinely projects—interactionally and/or grammatically—a quote. While the quotative is a verbal unit, the quote is not restricted to verbal material. On the contrary, the quote is recognizable through features like prosodic or bodily depiction. It is this grammatical format specifically, the bipartite structure [quotative + quote], that the present book sets out to explore in order to determine which social actions speakers implement with it in everyday talk.

---

[28]   As well as usage-based approaches (Bybee 2006, 2010; Bybee & Hopper 2001).
[29]   This may be surprising because, at least in written language, the quotative can be positioned flexibly (Sams 2008: 98ff on English).

2.4.   The Grammar-Body Interface

With the ability to video-tape interactions, research on grammar in interaction also began to investigate resources other than speech. Today, there are numerous studies investigating the functional use of linguistic structures while considering the interlocutors' embodied conduct (De Stefani & Mondada 2021; Horlacher & De Stefani 2017; Keevallik 2015; Mondada 2007, 2011, 2016; Pekarek Doehler 2019; Pekarek Doehler *et al.* 2021; see also the 20 papers collected in Pekarek Doehler *et al.* 2022; Schoonjans 2013, 2018; Stukenbrock 2016, 2018).

Some of the first articles demonstrating that linguistic and embodied conduct both contribute to the accomplishment of social interaction were by Kendon (1967), Kendon and Cook (1969), and Goodwin (1981). Kendon (1967) discovered that gaze plays a role in turn transition and that, to enable turn allocation in dyadic conversations, the recipient must look much more at the speaker than vice versa.[30]

Goodwin (1981) shows that one way for participants to indicate their "hearership" (*cf. ibid.*: 9) is to gaze at the speaker. There are two main findings from Goodwin's early work on the contribution of gaze to social interaction. The first is that gaze is used to reach a state of mutual orientation at the beginning of a turn; if a recipient does not gaze at the speaker, there are recurrent verbal techniques to solve this interactional problem, such as pausing or restarting, until the recipient gazes at the speaker (*ibid.*: 55ff). The second is that gaze is involved in recipient design; it allows speakers to show who their talk is addressed to by gazing at the respective participant (*ibid.*: 146ff).[31]

With the growing investigation of multimodal resources (Mondada 2016), it became clear that speakers and recipients use numerous resources[32] to coordinate and organize interaction and accomplish social action. A few examples from the literature illustrate this: Gaze is a powerful tool for turn-allocation in triadic conversations (Auer 2021) as well as for obtaining recipients' attention and/or affiliation during tellings (Thompson & Suzuki 2014). Body posture can introduce the (visual and bodily) reorientation of a group (De Stefani & Mondada 2013), and what Schegloff (1998) called "body torque" may be a means to project transition to another activity. Pointing allows recipients to announce themselves as "possible next speakers" without interrupting ongoing speech (Mondada 2007), or to displace an interlocutor's perception from the actual space of the ongoing conversation to an imagined space (Stukenbrock 2014). The palm-up open hand gesture has been demonstrated to contribute to action formation in that it occurs in responses where it is used as a "resource for reasserting and pursuing a prior position,

---

[30]   In their experimental study two years later, Kendon and Cook (1969) showed that "[a]mount, length and frequency of gazes and actions are related" (*ibid.*: 490), confirming Kendon's (1967) qualitative results.

[31]   Concerning gaze, recent studies, especially eye-tracking studies, have revealed new information. Auer (2021), for instance, extends Kendon's (1967) findings on turn-allocation to triadic interactions whereas Pfeiffer and Weiß (2022) discuss the role of gazing for recipiency in storytellings (see Section 3.1.3).

[32]   I understand *resource* as described by Couper-Kuhlen and Selting (2018), as "a form-based entity, including *verbal* forms such as phones and other sound objects, morphs, words, phrases, clauses, sentences, and recurrent larger units of talk; combinations of forms in (construction) formats; and *non-verbal* forms or devices such as prosodies, facial/bodily gestures, and body position and movement, to the extent that they accompany language and interface with it systematically in the communicative process." (*ibid.*: 356, original emphasis). Consequently, when referring to *multimodal resources*, I include both non-verbal and verbal resources under this umbrella term.

now laminated with an embodied claim of 'obviousness'" (Marrese *et al.* 2021: 1). Facial expressions may influence the "projected course of the talk" (Kaukomaa *et al.* 2015: 319) by displaying an emotional stance, while turn-opening frowns seem to allow speakers to "foreshadow utterances that somehow deviate from the recipient's routine expectation" (Kaukomaa *et al.* 2014: 132). All these studies exemplify that embodied conduct takes a major part in everyday human sociality. And just as language does, bodily resources may also interactionally project what comes next or make a next element or turn conditionally relevant.

Several studies have shown how a multimodal analysis can help us understand grammar-in-use. Stoenica (2020) shows in her study of French relative clauses how the recipients' multimodal conduct can initiate other-repair. She demonstrates, for instance, that when a reference is not understood by a recipient, the recipient may initiate repair through gaze or a facial expression during the pause following the reference. As repair, speakers may then extend their turn with a relative clause (see also Excerpt 2.1).

In his study of *if*-clauses, De Stefani (2021) demonstrates that clause-combining, bodily conduct, and action projection are closely intertwined. Investigating *if*-clauses in guided tours, he shows that guides use the format to reorient the participants' attention. In this use, the apodoses of the 23 investigated *if*-clauses are produced with different degrees of syntactical integration. Independently from the syntactic characteristics, speakers make systematic use of pointing gestures. The participants' subsequent response to the *if*-clauses combined with a gesture is a reorientation toward the next object of interest. This means that recipients respond to a grammatical format (= first pair part) with an embodied action (= second pair part). De Stefani (2021) introduces the notion of "*embodied action projection*" (*ibid.*: 16, original emphasis) to account for the features of the empty slot projected by the *if*-clause. This conceptualization may also be interesting for direct reported thought in which the quotative may also project embodied actions or in which embodied actions project a stance early on, before a verbal quote.

I will briefly illustrate with two examples how bodily resources and the pattern [quotative + direct reported thought] are intertwined. In both excerpts, speakers use direct reported thought to display their affective thereby exploiting mainly bodily resources. In Excerpt 2.2, Elinda tells her friend about almost fainting during a gym class, and in Excerpt 2.3, Jana tells that, in her former apartment, she could smell that the prior tenants were smokers.

Ex. 2.2: tout noir / all black (Pauscaf 20, 32min15)

```
ELI: Elinda, EKT: Ekti

01 ELI: et§ je suis <relevée> j'ai vu ↑TOUT noir-
        and I came up        and I saw all black
   eli   §both hands flat on the table-->>
02      j'étais là; #(1.1)#$
        I was there
   eli                    $gz at ekt-->>
   fig          #fig1 #
```

fig1

```
03 EKT: >ah MERde-<=
         oh shit
```

Ex. 2.3: muffig / musty (KAFKU_03, 05min33)

```
JAN: Jana, MAR: Marius, NIN: Nina

07 JAN: nd ich lag dann (.) [mit dem kopf an der  ] WAN:D,
        and I lay then   with my head turned toward the wall

08 MAR:                     [(ah scheißdreck)     ]
                             ah crap

09      (0.5)

10 JAN:  >und ich dAcht_immer# °H=°H<
          and I always thought ((sniffs))
   fig                        #fig1
```


fig1

```
11      des# riecht [aber] do(ch) irgendwie [SELt[sam.]
        it   smells PART   PART   somehow    weird
   fig  #fig2
```


fig2

```
12 MAR:             [hm- ]                        [hmhm,    ]
13 NIN:                                          [hm  ]hm-
```

In the first excerpt, Elinda depicts her body position and expresses her bafflement through a facial expression (Excerpt 2.2; fig1); In the second excerpt, Jana reenacts her sniffing at the

wall while taking a negative stance through a facial expression (Excerpt 2.3; fig1 and fig2). Both occurrences have the same bipartite structure. The verbal quotative is followed by a bodily depiction and a facial expression that display the stance that the respective speaker takes. In the German excerpt, a verbal explanation follows. In the French excerpt, however, the bodily depiction seems sufficient since the recipient orients to the speaker's stance with affiliation (line 03).

The staging (Yule & Mathis 1992) of the speakers' experience happens on two (or even three) levels simultaneously: Body movements, facial expressions and, in the German excerpt, an additional verbal utterance. These multimodal resources allow speakers to simultaneously co-index their affective stance on multiple levels. In line with Goodwin's (2013) notion of *laminated actions*, I analyze every resource as delivering a different component of the speakers' stance.

The two examples demonstrate the interwovenness of verbal and bodily resources. Through the simultaneous use of language and body, speakers may efficiently perform parallel actions: depicting a movement or event (e.g., through gesture or posture) *while* displaying their stance (e.g., through a facial expression). The quotative is thereby used as a device that projects these actions and makes them recognizable as quotes for recipients.

There have been suggestions to further the inclusion of bodily resources in grammar. Bodily conduct as part of grammatical projection has been much debated since the beginning of the new millennium. Keevallik (2018) argues that embodied conduct is integrated into grammar. With an Interactional Linguistic approach, she suggests that traditional grammatical categories (that are formulated on the basis of written language) should consider structures from interaction, including bodily conduct.[33] According to Keevallik (2018), "both social action formation and the emergence of grammatical structure are multimodal processes" (*ibid.*: 2).

In response to Keevallik's (2018) article, Couper-Kuhlen (2018) puts into perspective the claim of integrating the body into a grammatical system. While agreeing with many of Keevallik's suggestions about grammar and embodiment (*ibid.*: 22f), she also identifies problems, such as whether embodied behavior "[s]hould […] figure in the description of grammar at all" (*ibid.*: 23). Her main counterargument is that, even though body movements are grammatically projectable within an existing grammatical system, they do not themselves constitute "'grammatical' elements" (*ibid.*); instead, she proposes conceptualizing embodied elements as part of a syntagma (*ibid.*: 23) in which verbal elements can be replaced by embodied elements. This point of view may be regarded as verbo-centric, as it sees the verbal element as the default unit that can be replaced by bodily units. However, frequency arguments may speak in favor of such a syntagmatic conceptualization.

Another argument put forward by Couper-Kuhlen (2018) is that embodied elements are "not wholly linear but can overlap, if not fully coincide, with their frames" (*ibid.*: 23), as is often the case for direct reported thought. This contradicts the feature of unidirectionality of grammar in interaction (Fox 2007)—indeed, compared to verbal elements, embodied elements are not mutually exclusive; a speaker cannot produce two words at the same time, but they can, for

---

33   Presenting examples from several languages, she demonstrates that embodied conduct can carry out a multitude of actions that are closely tied to the organization of sequences, such as the initiation of sequences or the response to a first pair part.

instance, report a prior thought while depicting it with gestures and displaying their stance with a facial expression. Thinking this observation through, speakers cannot avoid simultaneously gazing somewhere, moving (or not moving) their hands, having a certain body posture, etc., which recipients may orient to as possibly relevant in the interaction. Thus, for analysts, it may not always be clear which resources are actually meaningful in interaction.

Keevallik (2018) regards quotations as structures in which one part projects embodied displays, while also suggesting that "these could be systematically described in a comprehensive grammar" (*ibid*.: 10). Indeed, the quotatives I investigate introduce utterances of which the degree of embodiment occurs as a continuum ranging from *no embodied conduct* to *embodied conduct only* (*cf*. also Keevallik 2013). In the present work, I do not aim at integrating the different types of embodied conduct into a detailed grammatical description of direct reported thought; instead, I conceptualize the bipartite structure of quotative and quotation in line with Couper-Kuhlen's (2018) suggestion of a syntagmatic relationship between verbal and bodily quotes, thereby acknowledging that "[g]rammar must be able to account for the slot but not necessarily for what fills it" (*ibid*.: 24).

2.5.   Conclusion

The basic principles that shape the system of a grammar-in-use as I apply it in the present work are all linked through the central role of *temporality*. The notions of emergence and projection allow me to approach direct reported thought from a point of view that considers the adaptability of grammatical patterns to interactional needs. Integrating bodily resources into the grammatical and interactional analysis of the complex syntactic pattern [quote + quotative] allows me to analyze the laminated character of actions (Goodwin 2013)—an undertaking that addresses an important research gap in research on reporting formats and integrates itself into a growing body of research on the grammar-body interface.

As Fox (2007) concludes in her article on the seven principles shaping grammar:[34] "[G]rammar is organized by dynamic and emergent practices; it is publicly available embodiment of unfolding actions situated in turns and sequences, it is contingent, strongly shaped by interaction, which is its birthplace and its natural home" (*ibid*.: 314). This book takes a close look at such "unfolding actions" and how they are, grammatically, interactionally, and multimodally, implemented through the direct reported thought format.

---

[34]   These principles are frequency, collocations, that these collocations occur in turns and sequences, unidirectionality, interactional construction, and embodiment. For a discussion in light of the results of this book, see Chapter 8.

3.   Direct Reported Speech and Thought

> "It makes sense that narrative—a ubiquitous form of human thought—would, mimicking memory itself, focus on the meaningful and leave out the meaningless. What I am indifferent to I mostly forget. The stories of memory and fiction are also made by absences—all the material that is left out." (Hustvedt 2011: 189)

When telling a story in everyday talk, speakers typically describe events and their circumstances, and they report their own and others' speech, thus making their narrations into something worth telling and worth to be listened to. Especially stories that affect the speaker directly are infused by the speaker's personal attitude toward what happened. This personal attitude transpires through lexical choices, prosody, and the speaker's bodily conduct.

Direct reported thought remains a niche research topic in linguistics (but see Casartelli *et al.* 2023) and even more within the CA and IL frameworks (but see Haakana 2007). Even concerning direct reported speech, "it is only in relatively recent years that research has focused on what reported speech does in interaction" (Clift & Holt 2007: 3), thus addressing questions of action formation and action ascription (Levinson 2013). This is even more surprising given that direct reported speech has a long scientific history in other disciplines, such as literary theory (Bakhtin 1981; Vološinov 1973), sociology (Goffman 1959, 1981), and philosophy (at least since Plato's *Republic*, where *oratia recta* is used as an argumentation technique (for a bibliographic overview see Güldemann *et al.*).[35]

According to Clift and Holt (2007), studies on direct reported speech have been concerned with mainly three research angles: its form, the authenticity, or truth value, of what is reported, and the actions that direct reported speech accomplishes, with the last angle being the most recent (*ibid.*: 3).

The amount of research in these three domains varies between French and German. In both languages, reported speech and thought have been studied in different ways. French linguistics has been primarily preoccupied with the syntactic features of direct reported speech and thought in written and spoken language (Blanche-Benveniste 1988; Mosegaard Hansen 2000; Rosier 1999). The most extensive work on direct and indirect reported speech and thought was done by Jacqueline Authier-Revuz[36] (Authier & Meunier 1977; Authier-Revuz 1984, 1978, 1979, 1992, 1993, 1987, 2019), who not only addresses grammatical challenges of the reporting format in general (Authier-Revuz 1984, 1979; see also Blanche-Benveniste 1988; Blanche-Benveniste 2010; De Cornulier 1978) but also develops a comprehensive theoretic approach to what she calls *la représentation du discours autre* ('the representation of the discourse of the Other') (Authier-Revuz 2019), thereby building on Bakhtin's reflections on dialogue (Bakhtin 1981). One of the first studies of reported speech in French everyday talk was carried out by Vincent and Dubois (1997) on Canadian French.

---

[35]   Direct reported speech has also been discussed under several names like, among others, *oratio recta*, direct quotation, constructed dialogue, inner speech, etc.

[36]   But see also the two doctoral theses by Gachet (2012) and Moreno (2016).

More recently, several pragmatic studies working with oral data have analyzed quotatives (Dostie 2020; Cheshire & Secova 2018; Rendulić 2015), along with some sociolinguistically oriented work, like Secova (2015), Neveu *et al.* (2014), or Moreno (2016). Demers (1998) has focused on prosody in reporting formats and found a recurrent pitch change between quotative and quote in Canadian French interactions. Many studies have also investigated specific quotatives in spoken language and even in interaction, including Abouda and Rendulić (2017) on the distinction between *faire* ('to do') and *dire* ('to say'), Blondeau and Moreno (2018) on *comme* and *genre* ('like') in Canadian French, and De Brabanter (2018) on *genre* and *dire*. Only few studies of French have investigated direct reported speech from a CA or IL perspective, thus including questions of sequentiality and action formation (Bangerter *et al.* 2011; Berger & Pekarek Doehler 2015). That speakers may also resort to bodily resources to stage direct reported speech or thought has, to my knowledge, only been mentioned in French and has been restricted to prosody (Secova 2015).

German linguistics has approached direct reported speech from a variety of perspectives, also concerning its methodological orientation (for an overview over the phenomenon's theoretical complexity, see Brendel *et al.* 2007). Other than research focusing on syntactic and semantic features of direct reported speech based on written language (for a semantic approach, see Discherl & Pafel 2015; on the contrast between direct and indirect speech, see Kaufmann 1976; on syntactic integration in reported speech, see Leistner 2016; for a focus on the word choice for quotatives, see Michel 1966; for a focus on deictic and syntactic matters, see Plank 1986: 301ff; on the formal features of the in-between of direct and indirect speech, see Roncador 2012; on syntactic and verbal matters, see Vliegen 2010; for a semantic and logic account, see Wierzbicka 1974), there are also studies in the domains of corpus analysis (Brunner 2015)[37], discourse analysis (Brünner 1991; Hausendorf & Quasthoff 1995; Quasthoff 2008, 1980; Redder 1994), functional pragmatics—which are mostly based on written language as well[38] (Baudot 2002; Weinrich 1993; Steyer 1997)—and CA and IL (Golato 2002a, 2012; Günthner 1997, 1999b, 1999a; Katelhön & Moroni 2018; Wessels 2019). Some studies include embodied issues, such as Ehmer's (2011) dissertation on speech imagination and animation, Streeck's articles on body quotes (Streeck 2002, 1988), or two recently published special issues (König & Oloff 2018a; Zima & Weiß 2020). All these studies either exclusively analyze direct and indirect reported speech or treat reported speech and thought equally, often without mentioning differences between them.

Two book-length studies with a constructional and discourse-analytic approach, working primarily with written data, investigate reported speech and thought but treat reported thought either marginally or as the same phenomenon as reported speech (Marnette 2005; Vandelanotte 2009). A few articles, however, investigate direct reported thought as a phenomenon distinct from direct reported speech (Acuña Ferreira 2021; Barnes & Moss 2007; Dostie 2020; Haakana 2007; Kim 2014; Macaulay 2001; Maynard 1996; Nguyen 2015; Nishiyama 2019; Park 2018;

---

[37] Between 2017 and 2020, there has also been a DFG-project on the corpus analysis of direct reported speech (https://www.ids-mannheim.de/lexik/redewiedergabe/) under the lead of Dr. Annelen Brunner (IDS Mannheim) and Prof. Dr. Fotis Jannidis (Unversität Würzburg).

[38] Studies working on oral data only started to emerge around the 1990s, such as Brünner (1991), Kallmeyer *et al.* (1994), or Schwitalla (1986). They all have a sociolinguistic or ethnographic dimension, such as diatopic or diastratic variation.

Sams 2010; Vásquez & Urzúa 2009; Wessels 2019), including Jefferson's (2004) paper on the pattern *at first I thought*. Among these studies, only Marnette (2005) and Dostie (2020) investigate French, and Wessels (2019) investigates German.

Additional studies on French and German treat specific grammatical formats or conversational practices including reported thought.[39] Some work has focused on verbless devices to introduce thought, such as the English *like*,[40] the German *so* (Golato 2002b; Streeck 2002: 590) and the French *comme* or *genre* (Cheshire & Secova 2018; Dostie 2020; Rendulić 2015: 178ff). These verbless quotatives, however, have been excluded in the present work.

Most literature on complement-taking predicates with the semantics of thinking has so far been concerned with their syntactic status as parentheticals or discourse particles, or with their function as hedges. Kärkkäinen (2010), for example, analyzes *I thought* as an epistemic phrase. Thompson (2002), like numerous other studies, investigated hedges or epistemic parentheticals (Aijmer 2011; Kärkkäinen 2003; Kearns 2007; Schneider 2007; Thompson & Mulac 1991b; Traugott 1995).[41]

In her 2012 paper, Kärkkäinen analyzes *I thought* as a "conversational format for taking a stance" (*ibid*. 2197). To do so, speakers often use *I thought* as a quotative introducing direct reported thought (*ibid*.: 2200). She states that "utterances involving *I thought* appear at points in longer tellings where speakers stop to evaluate what they are saying" (*ibid*.: 2197). Similar observations have been made for German by Deppermann and Reineke (2017) in one section of their article on the German *ich dachte* ('I thought'). The authors describe *ich dachte* as a device for taking an affective stance with similar formal characteristics: [*ich dachte* + interjection], occurring in tellings, mostly after the story's climax (*ibid*.: 342; see also Wessels 2019 for a similar practice with *ich dachte* from a sociolinguistic perspective). However, they do not conceptualize the format as a quotative introducing direct reported thought.

The results of these studies demonstrate that Couper-Kuhlen and Selting's (2018) claim that "there are good reasons for investigating reported thought in its own right" (*ibid*.: Online Chapter D, 22) is justified and that more research in the field is necessary. Neither in French nor in German has reported thought been studied as a distinct conversational practice (but see Fiedler 2023), let alone cross-linguistically—a research gap that the present book seeks to fill.

In the remainder of this chapter, I first outline general characteristics of direct reported thought by comparing them to the formal and functional features that have been established for direct reported speech: syntactic features, verbatimness, and the delimitation of quotes (Section 3.1.2) as well as bodily features (Section 3.1.3). In Section 3.2, I focus on the interactional workings of direct reported thought. The sequential environment of direct reported thought (Section 3.2.1), the review of Goffman's concept of *footing* (Section 3.2.2) and the potentially

---

[39] *Je me suis dit* is mentioned by Bangerter *et al.* (2011: 207f) and Cheshire and Secova (2018: 26ff). On *j'étais là*, see Secova (2015); on *ich dachte / ich habe gedacht*, see Deppermann and Reineke (2017); on *ich dachte so*, see Wessels 2019.

[40] Buchstaller notes on English *like* that it may have started as a "marker of reported thought" (2014: 110).

[41] Kärkkäinen (2010) argues, following Traugott (1995), that *I think* is becoming a construction marking the "starting-point of a perspective" (Kärkkäinen 2010: 229). This might first sound parallel to what I observe regarding quotatives introducing direct reported thought that allow speakers to display their stance. Kärkkäinen's (2010) analysis, however, is concerned with sedimented forms of *I think* and *I thought* that are inserted into utterances in variable positions as epistemic parentheticals—exactly those functions that are excluded in the present work.

distinct interactional use of direct reported thought (Section 3.2.3) show that a distinction between direct reported speech and thought may be necessary. I close the chapter with a conclusion (Section 3.3).

3.1.   Formal Features of Reported Speech and Reported Thought

Both direct reporting formats (speech and thought) have often been treated as equivocal. Because of the limited research on direct reported thought as a distinct phenomenon, the theoretical basis, and the formal features I will present must be mainly based on studies of direct reported speech;[42] if not pointed out otherwise, the features apply to both formats and have been treated as such in the quoted literature.

3.1.1.   Cross-Linguistic Syntactic Features

Research on reported speech has focused on several crucial formal features, including the differentiation between direct and indirect speech. Reported thought can, like reported speech, occur in the direct and indirect format, which is why most research has presented the features established for direct reported speech as equally applicable to direct thought. Because the formats introducing direct thought in my data also match the findings on direct reported speech, I will similarly apply formal criteria from reported speech to reported thought.

In my data, the German and French quotatives I investigate introduce, with few exceptions in French, a *direct* quote.[43] This is why, in what follows, I will focus on direct quotation only. Two prototypical examples of direct quotation are represented in the following two excerpts, here with *je me dis* ('I say to myself') and *ich denke* ('I think'):

Ex. 3.1: se calmer / calm down (Pauscaf 27, 10min44)

```
07 DAM: =je me dis oh nON;
           I say to myself oh no

08         ça ↑va ou quoi-
           are you okay PART
```

---

[42]   This treatment as equivalent in most of the previous literature has the effect that direct reported speech as research focus often includes direct reported thought and a separation of both formats is hardly possible. If not declared otherwise, when referring to terminology in the literature, I take over the terminology that the respective author has used to prevent me from wrongly attributing a larger or more narrow meaning to the term. When referring to the phenomenon in the present work and to my own research results, I use the term direct reported thought explicitly. The general use of direct reported thought might be underspecified for French *je suis là* ('I am there') and *j'étais là* ('I was there'), which seem to leave open whether what is reported has been said or thought. This issue will be discussed in the respective chapters (see Chapters 6 and 7). For the sake of readability, I will for now stick to the term direct reported thought.

[43]   For German, Auer (1998) has made a quantitative analysis of *(sich) denken* in a corpus of spoken and written language. According to his numbers, *denken* introduces dependent main clauses—the format that corresponds most to direct reported thought when the quote consists of a clause—in 53% of the 81 occurrences in his corpus of spoken language (*ibid.*: 6). This may be because, as he points out himself, his oral data also consists of formal speech situations in which speech has the style of conceptual written language (*ibid.*).

Ex. 3.2: stuttgart (FOLK_E_00228, 36min52)

```
18 SAH: (.) °h da denk ich HÖ (.) fünfhundertfünfzig für
            there I think huh  five hundred fifty for

        stuttgart sin ganz oKAY-=
        stuttgart are rather okay
```

The main syntactic feature of direct quotation in French and German that all occurrences of direct reported thought have in common is that the bipartite structure lacks explicit syntactic linkage with a complementizer. French and German syntax is divergent concerning clause-combining. The lack of the complementizer entails different clausal patterns that are, especially in German, complex and must be discussed separately. I will therefore address the language-specific features in the respective chapters for each language.[44]

I will nevertheless name formal features of direct reported thought and that go beyond a mere syntactic description and are valid for both French and German. I am aware that the following description is simplistic. However, because the features must be cross-linguistically valid, I restrict myself to these few universal formal characteristics, which have similarly been documented for direct reported speech:

- no complementizer between the quotative and the quote,
- no change of deixis[45] (Holt 1996; Plank 1986),
- introduction of the quote with *verbum dicendi*, *verbum sentiendi* or other quotatives not containing a complement-taking predicate (on the German *und ich so* ('and I like'), see Golato 2000 ; on the French *j'étais là* and *j'étais comme*, see Dostie 2020 and Cheshire and Secova 2018),
- no temporal or modal shifts of the verb in the quote,[46]
- dramatizing effects through prosodic and embodied reenactment (Couper-Kuhlen 1996; Klewitz & Couper-Kuhlen 1999; Sidnell 2006)

These formal features are mostly established based on prototypical direct quotes and have been criticized in French and German for being too restrictive and oriented to written language (for French, see Authier-Revuz 1979; Authier-Revuz 1984, 1992; for German, see Fabricius-Hansen 2019; Günthner 1999b, 2000; Imo 2007a). Indeed, such a categorization is hardly possible for spoken language where multiple resources like prosody or bodily conduct come into play. Restricting the analysis to such a prototypical format of direct reported thought would

---

[44] See Section 6.2.3 (German) and Sections 6.3.1.2 and 6.3.2.2 (French). The features that I present in these sections hold for the past and present forms of the respective quotatives.

[45] For German reported speech, Plank (1986) establishes a 'deictic hierarchy' of eight deictic categories: epistemic distance, time (verbal tense), space, time (adverbial), speech act roles, social distance, relational expressions, and speaker indexicality (*cf. ibid.*: 296). According to his study, if these eight categories refer to the reported situation, one can call a quote 'direct speech'. One of the characteristics of direct speech is, accordingly, "the highest possible degree of accuracy of rendering compared to the original speech" (*ibid.*, my translation). For the present work, applying a CA and IL methodology, it may be more apt to say that direct reported speech *displays* the highest degree of accuracy. Everything beyond a display lies beyond the analytical possibilities of CA and IL. Also, regarding several other categories, for example prosody, Plank's (1986) conclusion needs to be taken with caution (Couper-Kuhlen 1996: 366f).

[46] A recurrent phenomenon is, however, a shift to the narrating present tense in the quote (Günthner 2009b).

neglect the formal variety that speakers use in everyday talk—especially in German, a language with high grammatical flexibility (Günthner 2000: 291; Imo 2007a).

One of the main characteristics of direct quotation has been its verbatim reproduction of former talk/thought (Coulmas 1986; Banfield 1982; Leech 1971). In the 1990s, the verbatim character of direct reporting formats started to be questioned more widely (Clark & Gerrig 1990). Even earlier, Deborah Tannen already introduced the notion of *constructed dialogue* (Tannen 1986; see also Rühlemann 2020), in which she denounces the term *reported speech* as "misnomer" (*ibid*.: 311). As Günthner (2000: 281) relevantly points out, the characteristics of direct (and indirect) formats have been established based on constructed or literary (written) examples, which cannot be paralleled with spoken language, in which boundaries are often fuzzier (see also Imo 2007: 48ff and the chapters on *sagen* ('to say') and *meinen* ('to mean')).

Specifically relevant for the present work is Golato's (2012) observation that, "from a conversation analytic perspective, it is irrelevant whether direct or indirect quotations are entirely faithful to an original utterance" (*ibid*.: 6). What is important, however, is that speakers are "*doing being faithful*" (*ibid*.: 6f; see also Bergmann 1987).

In addition to staging reported thought as faithful, speakers must make direct reported thought recognizable as such for their interlocutors. In my data, this is first and foremost done with the social action format of the quotative. Quotatives can, however, mark only the beginning of a direct reported thought, not the end. How the delimitation of quotes is made recognizable through linguistic and bodily means, especially gaze, and how the inclusion of bodily resources complicates the separation of quotative and quote will therefore be addressed in the next section.

3.1.2.   Delimitation of Quotes

3.1.2.1.   Delimitation of the Beginning of the Quote[47]

A direct quotation of thought, when marked as such by an enquoting device, usually comprises two parts: the quotative and the quote. In the present work, the quotative, or the social action format introducing[48] the quote, is either a complement-taking predicate or a syntactically complete clause (like the French *j'étais là* 'I was there')[49]; and the quote is the direct reported thought itself, often prosodically and embodiedly staged and thus distinguishable from the introducing quotative. The following example illustrates the bipartite structure of the pattern [quotative + direct reported thought]:

Ex. 3.3: bier / beer (DejaVI_001, 04min10)

```
04 SON: <<laughing>ich da°ch° grad so,>
                   I thought PART like
```

---

[47]   In contrast to Sidnell (2006), I do not use the terminology of 'left and right boundary of the quote' because of its too textual connotation. 'Left' and 'right' echoes too much the written conception of language, which I try to overcome in the present work.

[48]   I am aware that quotatives can also be placed in the middle or at the end of a quote. The position of the quotative before the quote is, however, the most frequent one in my data (see Section 2.3).

[49]   As I explain in more detail in Chapter 5, verbless construction such as *genre* ('like') or *und ich so* ('and I like') were excluded in the present work.

```
05        oke also w(h)enn des des EINzige is w(h)as_er m(h)acht,
          okay well if    that's the only thing that he does
```

In Excerpt 3.3, line 04, the quotative, comprises the complement-taking predicate *denken* ('to think') as well as the deictic *so* ('like'). The quote is realized as a new intonation unit (line 05) and starts with a response cry (Goffman 1981), followed by the first part of a conditional clause of which the second part remains implied.

Whenever the whole pattern comprises a quotative, formally speaking, the delimitation of quote and quotative seems easy. There are, however, a few characteristics that complicate the separation of quotative and quote.

(1) In my data, direct quotes often include initial response cries, such as swear words, *oh*, *huh*, etc., or *particules d'amorces* (Guérin & Moreno 2015; for an explanation, see Section 6.3.1.2), such as *yes* or *uhm*. These small items are often difficult to attribute to either the quotative or the quote because they are prosodically attached to the quotative but often belong either to the quote or to neither the quote nor the quotative. Consider Example 3.2 again:

Ex. 3.2: stuttgart (FOLK_E_00228, 36min52)

```
18 SAH: (.) °h da denk ich HÖ (.) fünfhundertfünfzig für
            there I think huh  five hundred fifty for

        stuttgart sin ganz oKAY-=
        stuttgart are rather okay
```

The response cry *hö* ('huh') is prosodically attached to the quotative *da denk ich* ('there I think') but separated from the rest of the quote by a micro-pause. On a purely prosodic basis, *hö* would belong to the quotative whereas *fünfhundertfünfzig für stuttgart sin ganz okay* ('five hundred fifty for stuttgart are rather okay') would form the quote. However, a closer analysis clarifies that *hö* allows the speaker to (also through bodily means, not in transcript) first display her affective stance before moving on to a more explanatory part of her quote that displays the stance—again—verbally.

This phenomenon has already been described in studies working with oral data (Deppermann & Reineke 2017; Wessels 2019; Fauré 1997; Fauré & Vérine 2004; Guérin & Moreno 2015: 73ff). These studies show that response cries may allow the speaker to already display their stance toward the subsequently projected quote. Recipients thus get an early hint about the speaker's stance. In my data, such early indicators do not have to be response cries; they can also be bodily displays.

(2) On a prosodic level, the quotative and the quote may, in some cases, be clearly separable. In direct quoting formats, speakers can mark the beginning of a quote through a change of voice quality, volume, pitch register, or pitch range (Demers 1998). Such prosodic changes are, however, "often dependent on the presence (or absence) of subtle prosodic and paralinguistic effects" (Couper-Kuhlen 2007: 107). This results in the question of whether recipients perceive these subtle changes. Notwithstanding this analytic problem, which lies outside the scope of the present work, Couper-Kuhlen (2007) shows that such prosodic features are recurrent and that they are as crucial for direct reported speech as they are for direct reported thought (*ibid.*: 107ff).

In their article on the importance of prosody to mark quotes, Klewitz and Couper-Kuhlen (1999) first assume that these prosodic changes work similar to quotation marks in written language, a hypothesis that they later refute when analyzing authentic conversational data. One important result of their study is that prosodic marking does not replace the quotative when the latter is absent (*cf. ibid*.: 469). Pitch changes can, however, indicate boundaries when the changes occur only at the beginning and end of the quote, a phenomenon they call prosodic "flagging." The distribution of such pitch changes strongly varies in spoken language and is not systematically used.

One systematic prosodic indicator of the beginning of quotes that Klewitz and Couper-Kuhlen (1999) reveal is the rhythmic pause (*ibid*.: 475f), allowing the speaker to let the quote follow with a "suspenseful delay" (*ibid*.: 375). One aspect that can be added to Klewitz and Couper-Kuhlen's (1999) study—as seen in the last example—is that rhythmic pauses may also occur *after* a response cry or an interjection, which recur between quotative and quote and are still prosodically attached to the quotative. Additionally, as my multimodal analysis will show, what may be analyzed as rhythmic pause with audio data only may actually be a filled pause comprising the speaker's bodily conduct.

(3) Gestures and facial expressions often accompany direct quotation (Sidnell 2006)[50] and can even constitute a quote on their own (*cf.* Streeck 1988, 2002). Thus, similar to prosodic marking, the bodily depiction (see Section 3.1.3 for an explanation) may not be limited to the quote only. Previous research has shown that gestures or body movements precede speech onset, as, for example, when speakers reorient to a new object during a guided tour (De Stefani & Mondada 2013) or when they take an affective stance, in which case facial expressions may precede talk (Ruusuvuori & Peräkylä 2009). In addition to preceding the quote, body movements may also continue during the quote.

The use of embodied resources before the speech onset may also be of importance for the format of direct reported thought. Similar to fuzzy prosodic boundaries, these bodily characteristics of direct reported thought make a clear delimitation of quotative and quote difficult, as in Excerpt 3.4, where the speaker changes facial expression already during the quotative *j'étais là* ('I was there'), as fig1 and fig2 in the following illustrate:

Ex. 3.4: un doute / a doubt (Pauscaf 20, 41min00)

```
01 ELI: d'ailleurs ma meilleure amie,
        by the way my best friend

02      elle pensait qu'elle avait un DOUte pendant un moment-#
        she thought that she had a doubt for a moment

   fig                                                    #fig1

03      j'étais là #(0.5) faudrait peut-être saVOIR;
        I was there      might be good to know

   fig            #fig2
```
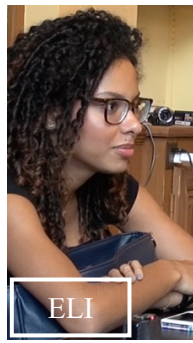
ELI fig1 fig2
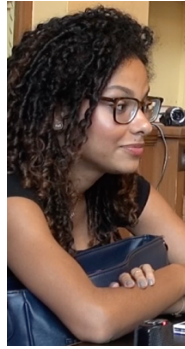
```
04          parce que (.) chais PAS si ça va être pareil; h°
            because I don't know if it's going to be the same
```

Before verbalizing her negative stance toward a friend who is not certain about her sexuality, Elinda's facial expression first indicates her negative evaluation of the matter (see fig1 and fig2). This means that, even before the verbal onset of the quote, the recipient, Ekti, can already know that a negative evaluation is about to follow. An analytical challenge lies in the belonging of the embodied conduct: Is it part of the quote, even though it partly overlaps with the quotative? Is it actually projected by the social action format in this case? Regardless, a partial overlap questions the bipartite structure of the quotation pattern.

These three points demonstrate that a clear delimitation of quotative and quote is complicated. What I observe in my data, however, is that recipients do not seem to struggle with recognizing the quote projected by the social action format; alternatively, it simply may not matter for them to know. Notwithstanding prosodic, syntactic, or embodied fuzziness, recipients do seem to know when a quote begins, if necessary.

So far, I have only treated the beginning of the quote, the determination of which I argue is interactionally less important than previous literature suggests. As I will show in the next section, the contrary is the case for the end of the quote.

3.1.2.2.   Delimitation of the End of the Quote

The end of a quote is interactionally relevant because it can coincide with a TRP and thus a possible turn-taking. To guarantee smooth turn-taking and avoid lapses, it is important for recipients to understand when a quote ends. That this may be so is the logical conclusion from the CA principle of "one speaker at the time" and the minimization of pauses, which accounts for most cases, as Sacks *et al.* (1974) argue in their seminal paper on the turn-taking system. However, quotation has one characteristic that makes the end of a quote difficult to determine: speakers may "fade out" a quote without specifically marking its end. This can induce misunderstandings, but it may also be used as an interactional tool.

The end of the quote often remains fuzzy, thus leaving space for interpreting whether a TCU, turn, or sequence is really closed. Verbally, no clues indicate where a quote ends other than, retrospectively, through the following turn. Speakers thus must resort to other resources. As mentioned previously, prosodic flagging may help recipients pinpoint the end of a quotation (Klewitz & Couper-Kuhlen 1999). However, such clues, like pitch changes, do not occur systematically. On a multimodal level, Sidnell (2006) argues that the end of a quote is marked by a systematic gaze returning to the recipient. The systematicity of this gaze pattern has been

revised and refined later by Pfeiffer and Weiß (2022), who show that the gaze back at the recipient can also occur earlier—namely, already during the reenactment—which defies marking a turn ending (*ibid*.: 26).

Thompson and Suzuki (2014), building on Sidnell's (2006) observations on gaze "parsing the larger telling into interactionally relevant units" (*ibid*.: 394), investigate multiparty conversations. They find that gazing at the recipient may not necessarily be an indication of a shift in footing back to the current participation framework but that it may also make the recipient from the ongoing conversation a character of the telling (*ibid*.: 833), an observation that Pfeiffer and Weiß (2022) also make for dialogical speech situations. The difficulty of analyzing gaze during quotatives thus consists in determining the "purpose" of a gaze: does a gaze at the recipient mean that the quote is about to be closed, or does it make the recipient a character of the telling sequence without affecting turn-taking? Because recipients rarely metapragmatically address gaze, as an analyst, I can only use the recipient's subsequent response as an indication of whether gaze has been used to yield the turn. The next turn proof-procedure[51] thus does not account for occurrences in mid-turn position.

This overview demonstrates that, while remaining difficult, the delimitation of the end of a quote is primarily done with embodied and prosodic resources. So far, I have discussed only bodily features in their role in the turn-taking machinery. Generally, body movements often occur during quotation (of thought and speech), thus contributing to staging and dramatizing effects. How speakers use prosody, facial expressions, gaze, and gestures more generally while quoting their own thoughts will be addressed in the next section.

3.1.3. The Multimodality of Direct Reported Thought

Reenactments are often used for staging purposes while telling a story (Clift & Holt 2007; Holt 1996), enabling the recipient to experience events similarly to the speaker. However, a systematic investigation of the interplay of storytellings and bodily resources still needs to be done (but see Bavelas *et al.* 2014; Blackwell *et al.* 2015; Couper-Kuhlen 1996, 1999; Günthner 2002; Klewitz & Couper-Kuhlen 1999; Park 2009; Ruusuvuori & Peräkylä 2009; Sidnell 2006; Stec *et al.* 2015; Thompson & Suzuki 2014). Only recently, König and Oloff (2018b) and Zima and Weiß (2020) focused on narrations and multimodality in their edited special issues. What has not yet been specifically addressed is how body movements may be integrated in a grammatic description of the social action format of the quotative and the quote that it projects. The present section will investigate this question while giving an overview of the bodily resources that speakers resort to most frequently in my data when quoting thought. Because of a lack of literature on direct reported thought and multimodality,[52] I will establish the state of the art based mostly on research results concerning direct reported speech.

A milestone in research on the relationship between quotations and embodiment was Clark and Gerrig's (1990) study on quotation. From a semiotic perspective, and analyzing authentic interactions of everyday talk, they distinguish between *demonstrations* and *descriptions*. According to them, quotations are a way to *demonstrate* what happened (instead of *describing*

---

[51]   For an explanation of the notion, see Chapter 5.
[52]   On multimodality and its analytical challenges, see Mondada (2016).

what happened). They focus on the mimicking character of quotations, showing that everyday talk consists of more than "serious" descriptions.[53] Instead, speakers use demonstrations to depict scenes with their bodies (Couper-Kuhlen 1996).[54] Later in their article, Clark and Gerrig (1990) also point out that during (embodied) demonstrations, "speakers can be committed to depicting any combination of aspects [i.e., bodily or verbal aspects, SF]. No one of them is privileged" (*ibid*.: 799).

Clark and Gerrig (1990) also specify two main differences between demonstrations and descriptions: Demonstrations are "nonserious rather than serious actions," and they "depict rather than describe their referents, though they depict only selected aspects of the referents" (*ibid*.: 764). The authors argue that the differentiation between demonstrations and descriptions allows speakers to modify the interlocutor's understanding of the respective utterances. One crucial feature of demonstrations is that they are always selective and "depict their referents from a vantage point" (Clark & Gerrig 1990: 767), which includes certain aspects of the depicted scene but excludes those which may not contribute to the action that the speaker accomplishes. Demonstrations thus allow speakers to report prior events while infusing them with their perspective by leaving out unimportant, unpleasant, or unnecessary parts—a feature of the direct reporting format that previous research has evidenced already for direct reported speech (Barnes & Moss 2007; Bavelas *et al.* 2014; Fox & Robles 2010; Günthner 1997; Ruusuvuori & Peräkylä 2009; Selting 2010, 2012, 2017).

Along similar lines, Sidnell (2006) describes a parallel phenomenon that he calls "reenactment." Reenactments are quotations in which both verbal and bodily resources are used by speakers and allow recipients to experience the reenacted scene. Similar to Clark and Gerrig's (1990) idea of quoting from a certain "vantage point," the notion of reenactments includes that the speaker's depiction "represents the events or actions from a particular perspective or position" (Sidnell 2006: 400). Again, Clark and Gerrig (1990) but also Sidnell (2006) describe only direct reported speech, but their arguments seem to be transferable to direct reported thought.

Different multimodal resources used simultaneously allow speakers to create what Bakhtin (1981) calls "layering of voices," or "polyphony." Günthner (1999a) has already shown that polyphony can be achieved through prosody during direct reported speech—a result that seems to be also applicable to direct reported thought. I argue that other bodily resources also allow speakers such a layering or, in Goodwin's terms, lamination (Goodwin 2000, 2013); Whereas verbal and bodily resources may be used to reenact the stance speakers take with a thought (e.g., a negative affective stance), gaze may allow the speaker to display, for example, the dispreferredness of the action that they are or were carrying out with the reenactment by gazing away (for empirical evidence, see Section 7.3.2).

---

[53]   By 'serious' actions, Clark and Gerrig (1990) mean actions—thereby following Goffman's (1974) division of serious and non-serious actions—that "'are said to be real or actual, to be really or actually or literally occurring'" (*ibid*.: 766), for example when two people are having a fight. In contrast, the authors continue, when two people only play to be involved in a fight, "their actions are non-serious" (*ibid*.). They point out that "non-serious actions are 'transformations' of serious actions" (*ibid*.).

[54]   Regarding prosody, Couper-Kuhlen (1996) concludes similarly that direct reported speech is a means to "imitate, and at the same time critically comment on, another speaker" (*ibid*.: 401).

Sidnell (2006) further develops these ideas from a CA point of view. Investigating video-taped interactions, he argues that the integration of several modalities "involves examining interaction *not* for gesture, gaze, or particular features of the talk per se but for the *activities* that the participants understand themselves to engaged [sic!] in" (*ibid*.: 380, original emphasis). Making the embodied nature of reported speech an issue of action formation allows him to pinpoint that such reenactments occur systematically in larger tellings.

Of particular significance to him is how multimodal resources allow speakers to delimit different parts of the telling. Sidnell (2006) argues that

> recipients must be able to determine whether the talk of the moment and associated gestures are anchored in the immediate context of the local telling or alternatively, in the events being talked about. (*ibid*.: 382)

That a recipient needs to know in which habitat gestures are anchored is crucial for the present work and has not yet been discussed for direct reported thought. When an utterance has been "only" thought, Sidnell's clear-cut difference between an anchoring in the past or in the ongoing conversation becomes fuzzy.[55] When, additionally, speakers use direct reported thought to take a stance without a clear marking of the footing, it becomes even less clear which habitat multimodal features, such as verbal conduct, facial expressions, or gaze, belong to.

In addition to CA studies, experimental studies have revealed important features on multimodality during tellings. In their article, Bavelas *et al.* (2014) present two significant results of their experimental study: (1) speakers use more direct quotations and facial expressions when they are in a dialogical communication setting (compared to monologues), and (2) this higher use of demonstrations is independent of whether the interlocutors can see each other. In a second experiment in the same article, they confirm the same significant results for gestures and what they call *figurative references* "which demonstrate a property of the referent by using the features or characteristics of something else" (*ibid*.: 623), for instance, 'puppy-dog eyes'. Their data is, as they indicate (*cf. ibid.*: 647f), limited regarding generalizability for naturally occurring data. This is due to three characteristics of their data: Not only did the recordings take place in a laboratory-like setting, interactants also did not talk about stories that they had experienced themselves but instead talked about a film.[56] This implies that, compared to my data, storytellings in experiments do not emerge from natural conversation, and speakers are likely less emotionally involved in tellings they have not experienced themselves.

The recurrence of bodily conduct during tellings is in line with what Streeck (2002) investigates in his study on the English *like* and the German *so*[57] ('like'; on the German *so* as a quotative, see also Golato 2002b). He describes that these two words are systematically followed by embodied conduct, which he calls "body quotation" (Streeck 2002: 581; see also Streeck 1988). For the German *ich so*, he states that it "occurs only with embodiments; it has

---

[55]   Such a 'clear-cut' difference is probably to be doubted anyway. As I have already pointed out on several occasions in the present work, it is preferable to speak of a continuum reaching from one extreme to another, in this case from talk being anchored in the telling or in the ongoing situation of the conversation.

[56]   Note that this is also partly the case for the data that Pfeiffer and Weiß (2022) use for their eye-tracking study. Participants mostly talk about a recent film or book they have seen or read. This might affect the degree of emotionality and affect that speakers display during tellings.

[57]   On *so* and its difficult categorization, see Auer (2006).

apparently not yet been generalized to other forms of reported discourse" (*ibid.*: 592), like its English (interactional) "equivalent." This highly generalizing assertion must be treated with caution, especially now, 20 years later, when the quotative may have routinized even further into a device introducing all kinds of quotes. It may be more consensual to say that there is a tendency for certain quotatives, such as the German *ich so*, to introduce dramatized reported events.[58] However, Streeck's cross-linguistic article is an important starting point for the present work, as relating the characteristic of the embodiedly reenacted projected material to a specific grammatical format (see also Fox & Robles 2010) is a crucial step in my analysis of the quotatives that introduce direct reported thought in French and German.

Multimodal conduct and action formation

How bodily conduct can contribute to action formation and differentiate two actions has been studied by Goodwin and Goodwin (1986) and Pekarek Doehler (2019). In their multimodal study on word-searches, Goodwin and Goodwin (1986) show that the understanding of all modalities depends on the activity context the modalities occur in. Their most prominent example for this activity-dependence is word searches. They convincingly show that a word search coupled with a speaker's gazing away conveys an individual search whereas gazing at the interlocutor allows the speaker to seek assistance. Similarly, Pekarek Doehler (2019) shows that in the same sequential position, *chais pas*, a strongly reduced form of the French *je ne sais pas* ('I don't know'), primarily carries out two actions, which are coupled with a distinct gaze conduct of the speaker. In sequences where Speaker A utters a guess, Speaker B does not respond, and Speaker A then produces a *chais pas*, speakers either avert gaze when using *chais pas* for "*backing away from the just-produced sequence-initial action, canceling the relevance of the projected response, and thereby making a bid for aborting the sequence*" (*ibid.*: 19, original emphasis) or continue gazing at the recipient and produce a "response-inviting gesture," "*pursuing response while relaxing the preference for a precise type of response*" (*ibid.*, original emphasis).

Goodwin and Goodwin's (1986) and Pekarek Doehler's (2019) results clarify that verbal information is only one part of the puzzle. Bodily resources provide, together with verbal material, information about actions, response relevance, and preference. This makes them a relevant part of the interaction-organizational infrastructure of everyday talk that participants may orient to. Although the linguistic resources seem to be multi-functional in interaction (a word search or *chais pas* ('dunno')), multimodal resources may help speakers clarify the meaning of their talk and make only one of the two (or more) actions interactionally relevant.

Multimodal resources include verbal and prosodic resources, facial expressions, gaze, gesture, and body posture. Not all these resources have been equally the object of previous research, especially concerning their use in direct reported speech, let alone direct reported thought. In what follows, I will briefly present research on the use of individual resources and their role when speakers report thought.

---

[58] For tellings in German, several dramatizing, or "performing," resources have been described by Günthner (2009a) and Imo (2009).

Prosody

On a prosodic level, Günthner (1999a), taking up Bakhtin's concept of polyphony,[59] argues that speakers use "prosody and voice quality" (*ibid.*: 3) to perform multiple voices in reported dialogues (see also Couper-Kuhlen 1996, 1999; Klewitz & Couper-Kuhlen 1999). Indicating a shift in footing (Goffman 1979), speakers can take over multiple roles within a telling. Such changes in voice quality, pitch range, or volume do not function like quotation marks in written language (for a discussion, see Klewitz & Couper-Kuhlen 1999). Especially regarding prosody, it is difficult to systematically assign a single feature to the same emotion or stance.[60] This has been evidenced by Günthner (2000) concerning what she calls "*vorwurfsvolle Stimme*" ('reproachful voice'). In her chapter on the prosodic analysis of reproaches, she convincingly argues that why-formats, such as *Warum hast du mir denn nichts gesagt?* ('Why didn't you say something'; *ibid.*: 85ff), do not have *one* systematic prosodic clue that characterizes them as reproach; instead, several parameters, such as co-occurring strong accentuation, falling intonation, high volume, or, less frequently, an elongated vowel on the accented syllable (*cf. ibid.*: 148).

Facial Expressions

Facial expressions have been recurrently related to stance and assessments in prior literature (Kaukomaa *et al.* 2015; Peräkylä & Ruusuvuori 2006, 2012; Ruusuvuori & Peräkylä 2009). Ruusuvuori and Peräkylä (2009) argue that facial expression allow speakers to "stretch the temporal boundaries of an action" (*ibid.*: 393). Investigating 10 storytellings in which assessments are used to close the telling—a practice that similarly occurs in my data with direct reported thought—they find that facial expressions occur in three main configurations: (1) the facial expression occurs before the speaker verbally conveys their stance, (2) the facial expression and lexical material co-occur, or (3) the facial expression occurs after the closing of the telling. A possible fourth category is a constant shift between a twofold stance, with the facial expression foreshadowing these stance changes, thus giving the recipient the possibility to align. Ruusuvuori and Peräkylä (2009) conclude that "face is a resource both for the speaker and for the recipient, and it serves in the construction of meaning as well as in shaping the relationship between the participants" (*ibid.*: 392). As such, facial expressions are closely linked to intersubjectivity (Lindström *et al.* 2021; Raymond 2019).

Of particular interest for the present work is Ruusuvuori and Peräkylä's finding that facial expressions, when occurring before or after stance-taking, can be conceived as a prolongation of the action. What the present work specifically addresses is the exact timing of facial expressions within the pattern [quotative + direct reported thought]. Does the facial expression start during the quotative already? Does it occur during the "rhythmic pause" that Klewitz and

---

[59]   Bakhtin's concept of polyphony has been originally established for written texts and the analysis of discourse in novels. Günthner (1999a) applies the theory on spoken language. This application has been re-used frequently when analyzing reported speech.

[60]   As Brünner (1991) points out, there is, however, a shared knowledge between participants concerning "stereotypes" of which prosodic or embodied characteristics are related to which emotion, for example, fear (*ibid.*: 8). She criticizes, however, that while such stereotypes may be transmitted, "not least in actu through reported speech itself" (*ibid.*), they can be "empirically wrong" and "do not need to correspond the actual conduct" in the reported speech (*ibid.*).

Couper-Kuhlen (1999) describe in their investigation of audio-only data? And wouldn't that mean these rhythmic pauses are actually filled pauses? How do facial expressions contribute to the "laminated organization of human action" (Goodwin 2013: 11; see also Goodwin 2000)? The present work will investigate the timing of facial expressions when they occur with direct reported thought. If, as Ruusuvuori and Peräkylä (2009) argue for their fourth configuration, facial expressions help recipients understand *early on* the affect that they must align, or affiliate, with (*ibid*.: 392), a pattern with a facial expression preceding the affective stance-taking or occurring early during the quote would be expected. If the recipient's affiliation is not delivered during the stance-taking, then a facial expression may go beyond the quote, thus maintaining the relevance of a response.

Facial expressions that can "stretch temporal boundaries" must be separated from head movements that recurrently co-occur with certain lexical items, such as modal particles (Schoonjans 2018), thus forming something close to a multimodal gestalt or construction. In his corpus-based study, Schoonjans (2018) examines the German modal particle *einfach* and shows that it systematically co-occurs with a headshake (*ibid*.: 112ff). Comparing the meaning of a headshake with the meaning of the particle *einfach*, he concludes that there is a consistency of pragmatic function (*ibid*.: 112). This is confirmed by a quantitative and qualitative analysis. Despite the result that only in about 50% of all occurrences in his corpus the particle *einfach* co-occurs with a headshake, Schoonjans shows through a detailed qualitative analysis that the density of sequential and contextual features allows the conclusion of a recurrent multimodal gestalt [*einfach* + headshake]. Schoonjan's study makes clear, similar to Stukenbrock (2011) concerning the multimodal gestalt [*so* + pointing gesture], that there are instances in which specific head movements or gestures can be closely related with one lexical item. However, Schoonjan's (2018) work also clarifies that such co-occurring verbal and multimodal features reveal only tendencies and not a clear and/or exclusive conditional relationship between language and body.

Speaker's Gaze

Speaker's gaze does three main things during reported speech and thought: (1) organizational work, like turn yielding (Sidnell 2006), (2) it makes the recipients of the ongoing conversation characters of the telling (Thompson & Suzuki 2014) but only when the reported event is dialogic (Pfeiffer & Weiß 2022), and (3) it monitors the recipients' "appreciative recipiency" (Thompson & Suzuki 2014: 816).

It is difficult to differentiate between these three actions, as speakers can perform them with a single gaze. One major problem is that next turn proof procedure does not always work regarding bodily conduct because interlocutors do not necessarily address whether a gaze at the recipient has been understood as yielding the turn or monitoring recipiency. If a recipient response follows, it could have been both; if an alignment follows even though affiliation has been projected and the speaker continues their telling, the gaze at the recipient could have been only for monitoring. Another interpretation could be that the gaze is an unsuccessful attempt to yield the turn, in which case the speaker can pursue affiliation. A detailed sequential and multimodal analysis will address, for each analyzed excerpt, how gaze is used in telling

sequences with direct reported thought. I will show that, in some cases, gaze conduct may be closely related to the actions that speakers carry out with direct reported thought.

Recipient Gaze

Previous research on gaze has mainly investigated recipient gaze regarding participation (Goodwin 1979, 1981, 1984; Lerner 2003). In his article on gaze behavior in a telling sequence, Goodwin (1984) recognizes recipient gaze as contributing to the dynamics and the progressivity of a telling (see also König & Oloff 2018b: 281; Mandelbaum 2013). Later work on recipient gaze has focused on its contribution to the organization of social interaction and its interrelation with activity contexts (Rossano 2010, 2012; Stivers & Rossano 2010).

Rossano (2012) calls for an investigation of gaze regarding the activities that speakers and recipients are involved in. He shows remarkable differences in the recipient's gaze conduct (*ibid*.: 53ff) in the openings of two distinct sequential environments: extended telling sequences (ETS) and "adjacency pair-based sequences."[61] During the first TCU of extended tellings, recipients usually gaze at the speaker "as soon as an ETS is projectably underway" (*ibid*.: 53) and hold their gaze during the telling. This is usually not the case in first pair parts of an adjacency pair like requests. Rossano (2012) also calls for caution concerning Kendon's (1967) suggestion of the default gaze pattern that while listening, participants look at the speaker "with fairly long gazes, broken by very brief gazes away" (*ibid*.: 27). Such a pattern would entail two consequences, as Rossano relevantly points out: first, an activity-independence of this gaze pattern, which he disproves in his thesis (*ibid*.: 53ff), and second, that speakers do not need to deploy resources to mobilize recipient gaze because it would be established anyway (*cf*. Rossano 2012: 56).

In contrast to Rossano (2012), I investigate the middle and the closing environment of tellings, not their openings. While Rossano also shows that during tellings or after laughables, gazing at the recipient allows the speaker to pursue a prior absent response (*ibid*.: 207ff), he states that, for closings in general, "participant's gaze orients toward the accomplishment of the projectable course of action more than toward the completion of a single sequence" (*ibid*.: 280). Because he investigates extended telling sequences, participants orient to both smaller sequences and larger courses of actions. Rossano demonstrates that recipients (and speakers), after sustaining gaze over multiple sequences, withdraw gaze only when the end of the larger project (here, a telling) is reached. This finding is particularly relevant for the present work because direct reported thought often occurs in closing environments of extended telling or reasoning sequences. Gaze could therefore be a relevant means for recipients and speakers to reach completion and manage a possible turn-taking (also) through gaze.

Gesture

Most research on multimodal storytellings focuses on gestures contributing to the dramatization of reported events (Fox & Robles 2010; Ladewig 2014; Sidnell 2006; Stec *et al.* 2015). Gestures have, however, not yet been systematically investigated for direct reported speech, let alone for direct reported thought.

---

[61]   For a detailed explanation of the notion *adjacency pair*, see Chapter 5.

When analyzing tellings multimodally, gestures are often part of the analysis without being the focus of investigation. They have been described as iconically depicting parts of the telling (Ladewig & Hotze 2020) or as referring deictically to an imaginary space (Stukenbrock 2012, 2014, 2016, 2018). Within the framework of gesture studies, Ladewig and Hotze (2020) have shown that, in storytellings, gestures contribute more to the development of the narration than to interactional organization. With mainly iconic gestures, speakers "emphasize individual semantic facets of linguistic-gestural meaning in the course of the narration" (*ibid*.: 124, my translation). Sometimes, gestures can even fill verbal gaps when utterances remain syntactically incomplete (*ibid*.).

One of Ladewig and Hotze's (2020) findings on gestures during tellings is particularly interesting for the present work: their contribution to intersubjectivity (Lindström *et al.* 2021; Raymond 2019). The authors observe that the use of gestures increases around the story's climax (see also Selting 2017), after which speakers usually seek the recipient's response (Guardiola & Bertrand 2013). Before reaching the story climax, it is thus necessary to establish mutual understanding between interlocutors, as only when mutual understanding is established can a response be delivered by the recipient. Despite Ladewig and Hotze's (2020) claim that gestures contribute less to interactional organization, their increased occurrence correlates with a point in the telling that most evidently projects a response. The function of gestures may thus be indirectly related to the organization of turn-taking, as it makes, along with other resources, the gist of the telling (more) recognizable for others.

Selting (2010), in her interactional study of affectivity in complaint stories, reaches a comparable result. She observes a direct proportionality between the intensity and amount of verbal, prosodic, and bodily cues and the degree of the teller's "emotive involvement" (*ibid*.: 270). Speakers use "visible devices," such as facial expressions, gestures, posture, etc., to "construct and make recognizable climaxes with displays of affectivity" (Selting 2017: 29). One crucial result of Selting's interactional approach is that the multimodal reenactments seem to be interactionally motivated, as recipients must recognize which activity speakers are involved in during their telling. Exploiting a maximum of resources makes the recognition of complaints, amusing stories, assessments, etc., easier. Particularly important for the present work is Selting's observation that affectivity and multimodal resources are tightly related.

The last paragraphs have shown that embodied conduct does not simply accompany the speaker's verbal utterances but also adds a semiotic layer (Goodwin 1979, 1981) to the quote that is far from being redundant (De Ruiter 2007). Multimodal features provide important insights into the recipient's and speaker's orientations toward each other and the roles they assume within a telling, thus managing smoothly sequential implicativeness on different levels of footing. At the same time, the multiple resources enable speakers to "layer" their actions; while reporting thought, they may simultaneously display their stance through a facial expression. What Goodwin (2013) calls the "lamination of actions" may, in the case of direct reported thought, be "assembled through the dynamic interplay of different layers of diverse semiotic phenomena (talk, prosody, gesture, etc.)" (*ibid*.: 14).

3.2.   On the Interactional Workings of Reporting One's Own Thoughts

3.2.1.   The Sequential Environment of Direct Reported Thought: Storytellings and Sequences of Reasoning

In my data, direct reported thought is recurrent in storytellings. These storytellings can be emotionally charged tellings or rationalizations of past events, depending on the "occasions of their production" (Mandelbaum 2013: 492). Because this differentiation is relevant to the present work, I will call less emotional tellings in which speakers use the narrative format to make publicly available their rationalizations *reasonings*. The terms *tellings* and *storytellings* will be used for emotionally charged narrations in which speakers tell a story that may contain multiple characters and their reported speech. This is rarely the case in reasonings, in which speakers usually present their thinking process without staging multiple voices. That storytellings can, more generally, take different shapes ranging from emotional tellings to argumentative sequences has already been documented by Grosjean and Lacoste (1999). Building on this observation, Bangerter *et al.* (2011) show that, nurses use different kinds of tellings when reporting past events to their colleagues taking over their shifts. During their handover meetings, these nurses also use similar kinds of rationalizations to those in my data.

Telling-sequences are often shaped as "big packages" (Sacks 1992: 354ff, Vol. 2) that emerge in time and are influenced by contextual and interactive clues from the recipient as much as they are from the environment. Instead of being monological sequences of the speaker, tellings are co-constructed in that their progressivity is (or at least can be) influenced by recipients, who can signal attention, display their (non-)understanding or stance (Ruusuvuori & Peräkylä 2009), or respond to what their interlocutor narrates (Stivers 2008). As Mandelbaum (2013) puts it, "stories are interactive productions, co-constructed by teller and recipient and tailored to the occasions of their production" (*ibid.*: 492). The reciprocity of tellings is enhanced by the speaker's preference for the recipient's alignment and affiliation with the story (Selting 2010; Stivers 2008) and their working to achieve this goal, which can be performed verbally or through bodily means; tellings are often staged (Yule & Mathis 1992) and contain descriptions and depictions (Clark & Gerrig 1990), and speakers often make use of multiple bodily resources for dramatical purposes (Blackwell *et al.* 2015; Couper-Kuhlen 1999; Ruusuvuori & Peräkylä 2009; Sidnell 2006; Stec *et al.* 2015; Streeck 2002; Thompson & Suzuki 2014).

The sequential structure of tellings has already been discussed by Jefferson (1978). Her starting points are Labov and Waletzky's (1967) seminal paper on the narrative structure of tellings and Sacks' (1992) explanations in his lecture "the baby cried the mommy picked it up" that tellings are occasioned by context and interactively developed by the teller *and* the recipient. Jefferson (1978) states two important sequential features of tellings regarding their beginnings and endings: "Stories emerge from turn-by-turn talk, that is, are *locally occasioned* by it, and, upon their completion, stories re-engage turn-by-turn talk, that is, are *sequentially implicative* for it" (*ibid.*: 220, original emphasis). More recent literature has shown that embodied resources contribute to the sequence organizational machinery of tellings (see Section 3.1.3). Gaze, for instance, seems to be a powerful tool to mobilize response at the end of—but also during—a telling (Stivers & Rossano 2010; Stoenica & Fiedler 2021). For

Jefferson (1978), there are two main patterns in how tellings are occasioned: either they are "triggered" (*ibid*.: 220) by something that makes one of the participants think of the story, or they are "methodologically introduced" (*ibid*.), meaning that the telling is introduced with a clear account of why it occurs at this specific moment. This recurrent introductory practice for tellings is called *story preface* (Goodwin 2015: 197ff; Sacks 1978, 1986, 1989), and it allows speakers to indicate early on what kind of telling is about to follow or to provide temporal or local aspects through adverbs like *yesterday*, *when I went out last Saturday*, *at the cinema*, etc.,

> thus providing resources for prospective recipients to align as recipients of an upcoming extended turn, and to project what kind(s) of responses will be relevant to the story and when, not only while it is in progress, but also upon completion. (Mandelbaum 2013: 496)

Reasonings, even though often formatted as "big packages" (see *supra*), may also occur as short insertions after a telling. They are usually monologic reports of a speaker's thinking process. Speech or thought from third parties is rarely reported. Compared to tellings, reasonings are also less depictive; speakers use fewer resources to report events and less frequently share their affective stance toward what is reported. However, similar to tellings, speakers often frame their reasoning temporally and locally in cases in which the reasonings form a multi-unit turn and do not occur at the end of a telling.

When speakers arrive at the end of their telling (or reasoning), there are, according to Jefferson (1978), again two possibilities for recipients to continue: either the story is followed by "topically coherent subsequent talk" (*ibid*.: 228), such as a second story, or subsequent talk contains an account for its fittedness to the story, thus showing that the story has indeed been "implicative for subsequent talk" (*ibid*.). This second possibility of a response to tellings is tightly linked to the observation that, while telling a story, speakers usually convey their stance (Stivers 2008) through verbal and embodied resources. Mandelbaum (2013) relevantly points out that "[t]he teller's stance, and the extent to which this is conveyed, is a crucial resource for recipients, as it makes available the teller's expectations regarding how the events of the storytelling are to be responded to" (*ibid*.: 499). My analysis shows that direct reported thought may be one way to explicitly display stance, thus making publicly available these expectations.

Generally, research in CA and IL has shown that, "[f]or the most part, people tell stories to do something—to complain, to boast, to inform, to alert, to tease, to explain or excuse or justify" (Schegloff 1997: 97). Labov and Waletzky (1967) already show that evaluating prior reported events of a telling is part of the narrative structure of stories. Labov (1972) describes this phenomenon of conveying an attitude through a story itself as "internal evaluation" (for a similar observation, see Couper-Kuhlen 2007; Selting 2017). The recipient's response becomes especially relevant after the climax of the story and its 'internal evaluation'.

According to Stivers (2008), the speaker's evaluation or stance-taking during the telling makes mainly two kinds of responses at different positions relevant: alignment (during the telling) and affiliation (in response to the telling). Through their alignment, a recipient "supports the structural asymmetry of the storytelling activity" (*ibid*.: 34)—meaning, they display their understanding of the suspension of the usual turn-taking rules (Sacks *et al.* 1974: 704; see also Chapter 5) in favor of the speaker's multi-unit turn (i.e., the telling). Alignment is displayed through nods or backchannels. However, Stivers (2008) shows that nods in mid-telling position can also be understood as affiliative: "[W]hen a recipient claims that he or she has achieved

some measure of access to the teller's stance, that recipient also suggests that the telling is on track to receiving affiliative uptake at story completion" (*ibid.*: 53).

In response to a telling conveying a stance, the preferred response is affiliation (see also Lindström & Sorjonen 2013; Stivers *et al.* 2011). By affiliating with the teller's stance, a recipient "displays support of and endorses the teller's conveyed stance" (Stivers 2008: 35), thus emotionally participating in what is told. The following table, taken from Steensig (2020: 249) shows both types of responses:

Table 1 Overview of alignment and affiliation features

Alignment: structural level	Affiliation: affective level
Facilitate and support activity or sequence	Display empathy
Take proposed interactional roles	Match, support, and endorse stance
Accept presuppositions and terms	Cooperate with action preference
Match formal design preference	

Previous research on direct reported thought in storytellings has demonstrated that quotations of thought allow the teller to display their stance (Couper-Kuhlen 2007; Haakana 2007; Selting 2017). In these cases, direct reported thought often occurs after the climax, as "internal evaluation" (see *supra*). This position in the telling is different from what has been shown for direct reported speech, which recurrently forms story climaxes (Holt 1996).

In addition to the results of prior literature on direct reported speech and thought, the present book also exploits Goffman's concept of footing to analyze the interactional workings of direct reported thought. In the next section, I first address the characteristics of the notion of footing before connecting these features to my phenomenon of interest.

3.2.2.   Goffman's Concept of Footing: Impacts on Self-Quotation[62]

Generally, each time a speaker changes from someone else's speech or thought[63] to their own thought and when they shift from their past self to their self in the here-and-now, they perform a change of footing. In an early paper, Goffman (1974) suggests that a speaker can be separated according to different parts between which they may change: author, animator, and principal. He defines the roles as follows: the person producing sounds and bodily conduct is the *animator*, the person having originally produced certain (purportedly thought) words or movements is the *author*, and "someone whose position is established by the words that are

---

[62]   One formal aspect has not yet been addressed specifically: With a speaker having epistemic primacy over their thoughts, direct reported thought is, in most cases, a self-quotation (Golato 2002b; Lamerichs & Te Molder 2009), which holds specific features regarding the conceptualization of the reporting and reported speaker. Previous literature on self-quotation as a distinct phenomenon clarifies that an interactional investigation of self-quotation is still to be conducted (but see Golato 2002b). The present work regularly touches on specific functional aspects of self-quotation of thought, in which reporting and reported speaker are "the same."

[63]   I am aware that Goffman established the concept of footing on the basis of (reported) *speech*. Because the present work treats reported *thought* and because footing is as valid for direct reported thought as it is for speech, I will consistently use thought as reference medium, thus applying Goffman's concept to the research topic in the present work. This transfer excludes, of course, citations from Goffman's original texts.

spoken, someone whose beliefs have been told, someone who is committed to what the words say" (Goffman 1981: 144) is the *principal*. The flexibility of these parts and a shift between different animators, authors, and principals allow speakers to report and enact thought (on the relation of quotation and Goffman's participation roles, see also Golato 2002b). A certain footing can be held over several turns, as it is possible for recipients to react to the speaker's telling without interrupting the speaker's footing despite backchannels from the recipient (Goffman 1981: 155).

The analytic challenges of footing have been addressed in Goffman's 1974 study. He argues that when a speaker repeats former talk and/or is non-serious,

> he means to stand in a relation of reduced personal responsibility for what he is saying. He splits himself off from the content of his words by expressing that their speaker is not he himself or not he himself in a serious way. (*ibid.*: 512)[64]

Quoting oneself is, in this regard, an interesting phenomenon. The "splitting off," as Goffman calls it, does not take place between someone else's past thoughts and the speaker's (spoken) repetition in the here-and-now of the ongoing conversation; instead, it is between the past speaker's thought or conduct and themselves in the ongoing conversation. Goffman (1974) explains that tellings require the speaker to take an outside perspective on past events to "replay" them with a narrative structure and certain dramaturgy that works toward a narrative aim so "listeners can empathetically insert themselves into [the replaying], vicariously reexperiencing what took place" (*ibid.*: 504). For direct reported thought, this entails, however, that speakers also separate between their past and present self to be able to achieve a narrating, outer/omniscient perspective. There are several arguments for such a separation:

(1) The speaker of the past is different from the speaker of the here-and-now of the ongoing situation because they may have changed their opinion or thought about the reported matter (Goffman 1974: 517-523). What speakers stage in ongoing talk is the report of their *purported* past thought, thus adapting it to the ongoing activity of a telling, and to their stance.

The mere fact that speakers use direct reported thought to display their stance indicates that, during self-quotations, speakers may be *physically* the same but not more than this. In her article on self-quotation, Golato (2002b) addresses this matter by arguing that "when a speaker designs a quote it is the local context that is relevant to the speaker who is quoting, not the distant context in which the original discourse occurred" (*ibid.*: 64).[65] I will follow Golato's (2002b) proposition that self-quotation is designed to respond to the contingencies of the current speech situation.

(2) Separating the past and present self for self-quotation may be a possibility for a speaker to reduce their responsibility toward former thought (Goffman 1974: 512; see also Macaulay 2001). This may be necessary because direct reported thought may, like direct reported speech,

---

[64]   I would like to note that, in verbatim quotations, I will take over the original gender marking in personal or possessive pronouns even if the feminine form is not explicitly included. For the sake of readability, I decided not to add the feminine form but consider all masculine forms as generic forms.

[65]   McGregor (2019) even goes one step further. Concerning what he calls "phenomena of self-now-quotation" (*ibid.*: 210) of the type *I'm thinking*, he suggests an analysis of the format as a type of reported speech because "there is no other moment than the current speech moment in which the reported utterance occurred" (*ibid.*). Considering that speakers present reported thought as something that has never been pronounced, I will not adopt McGregor's conceptualization.

be associated with tellings that involve gossip. In his book on gossip, Bergmann (1993) states that speakers tend to overstep certain (moral) boundaries (*ibid*.: 117). One of the motivations behind the use of quotation in gossip is that their content "cannot be checked by the recipients for their accuracy and therefore are hard to doubt" (*ibid*.). Bergmann (1993) further argues that there is a "relaxation of the rules of censure [that] clears the way for a *scandalizing reconstruction*" (*ibid*.) through both linguistic and prosodic means. Clift and Holt (2007) explain such transgressions as follows: "The reduction of responsibility for a reported utterance partially accounts for the association between reported speech and gossip" (*ibid*.: 13). The overstepping character of quotations has been more specifically investigated concerning its systematic occurrence with relatively new quotatives like *to be like* (Romaine & Lange 1991) or with French quotatives like *comme* and *genre* (Cheshire & Secova 2018; Dostie 2020; Secova 2015). Buchstaller and Van Alphen (2012) argue that such new quotatives are ways of quoting content while reducing responsibility because the quotatives lexically emphasize approximation instead of faithful reproduction. Since swear words and extreme-case formulations are indeed recurrent when speakers use direct reported thought in my data, reducing responsibility may also be relevant for quoting one's own thought—a further argument that the speaker from the past and the speaker in the ongoing conversation may need to be conceptually separated.

(3) Within the telling, the speaker must be simultaneously the "performer" of the ongoing reenactment (= the animator) in the here-and-now and a character in the telling that the speaker is narrating (= the author; Goffman 1981). These two roles may be physically the same but in different points of time. Goffman (1981) argues that when speakers quote themselves, it is as if there were "embedded animators":

> [T]wo animators can be said to be involved: the one who is physically animating the sounds that are heard, and an embedded animator, a figure in a statement who is present only in a world that is being told about, not in the world in which the current telling takes place. (*ibid*.: 149)

The ambiguity of the role of the speaker may lead to an ambiguity in sequential implicativeness. Recipients (and analysts) must know whether the stance-taking makes a response relevant in the ongoing conversation or only in the story world of the telling. This is exactly what Goffman describes: that a speaker dramatizes events and stylizes themself enhances their perception as a figure in the narration.[66] At the same time, a relevant next by the speaker of the ongoing talk may be awaited even more *because* of such embodied reenactment (see Section 3.1.3).

The problematization of self-quotation regarding the speaker's footing and their roles while quoting has clarified the following: (1) even though speakers quote themselves, they cannot be conceived as the same regarding their roles as reporting and reported speaker; (2) the quotation of thought is shaped by the ongoing conversation, or influenced by the demands and contingencies arising from the current interaction, such as staging and reenacting to dramatize; and (3) the actions that speakers carry out with the pattern [quotative + direct reported thought] can (but need not) have a sequential implicativeness on the interlocutor in ongoing talk.

---

[66]   For a similar observation of stylization, see Brünner (1991). She describes this specific type of characterization as a "means of distancing" (*ibid*.: 19), often achieved through exaggeration. This kind of stylization is tightly linked to stereotypes with the help of which speakers enrich what they report.

To summarize, the analysis of quotatives in first-person singular, or self-quotation (Golato 2002b), and speakers displaying their quoting thought (and not speech) somewhat alter Goffman's concept of footing. I will therefore address the following questions in my analysis: Do the self-quotations of thought remain part of the story, or do they stand as "commenting thoughts" after a change of footing to the here-and-now of the ongoing situation? When in the pattern [quotative + direct reported thought] does a change in footing take place? Would the speaker's embodied conduct be a possibility for displaying parallel footings—meaning, while verbally reenacting past thoughts within the storytelling, the speaker may display an affective stance in the here-and-now of the ongoing conversation through their bodily conduct (e.g., with a facial expression)? The empiric analysis will allow me to analyze, through a detailed sequential and multimodal analysis, the emergence of quotatives introducing direct reported thought along with the changes of footing that speakers perform.

How direct reported thought has been described specifically in interaction will be explained in the next section.

3.2.3. Interactional Use of Direct Reported Thought

The instances of direct reported thought that the present study investigates are distinctive in that they occur in longer narrative sequences that are either storytellings or reasonings formatted as inner monologues (see *supra*). In doing so, they differ from instances of direct reported thought that speakers use as isolated utterances in non-narrative sequences (Haen 2019; Koester & Handford 2018; Park 2018). Studies on these types of direct reported thought nevertheless give interesting insights into when and for which (interactional) purposes speakers integrate direct reported thought into their talk.

Park (2018) investigates advice-giving sequences between writing-instructors and students discussing the student's writing. She demonstrates that the instructor's direct reported thought "provides students access to a reader's real-time reaction to their writing, while allowing instructors to provide a critical assessment and ground their accompanying advice" (*ibid.*: 1). One major point that she makes is that the reported thought is an instrument for writing instructors to deliver critical feedback to a present party, which is a potentially delicate action. The moral implications arising from this setting are specifically present in hierarchical situations, such as the one described by Park. Similar moral issues are also observable in my data, wherein speakers gossip or criticize third parties, an equally delicate matter as research on complaints has shown (Drew 1998; Heinemann & Traverso 2009; Ruusuvuori & Lindfors 2009; Skogmyr Marian 2021, 2023; Traverso 2009). In a similar setting, Haen (2019) shows that tutors and students in writing centers may also use direct reported thought for supportive actions like praising or affiliating. They thereby also resort to bodily resources and non-lexical vocalizations to make academic writing processes maximally accessible to unexperienced students (see Sandlund 2014 for a similar observation). Despite the non-embeddedness of these cases of direct reported thought into larger narrative sequences, speakers seem to use the format to convey their affective stance toward an action or event. What Park's (2018) and Haen's (2019) instances of direct reported thought have in common is that they allow speakers to assess prior actions or events while managing the social impact such an assessment may have on the relationship between themselves and recipients. Potential moral consequences of negative

stance-taking may be alleviated by the format, as the distance between the temporally displaced event (namely, within a narrated scene) created by the reporting format may hedge immediate responsibilities for what is said.

Wessels (2019) is one of the rare studies that investigates the verb *denken* in its function of a quotative of thought in everyday talk. Analyzing the format [PERS+*denk*-Aff+*so*+x] and [*denk*-Aff+PERS+*so*+x] (e.g., *ich denke so* / *denke ich so* in present and past tense)*,* Wessels identifies the form-function pairing of direct reported thought with *denken + so* that allows speakers to position themselves. Exploiting Goffman's concept of footing (see Section 3.2.2), she shows that speakers use this pattern to change footing in dialogic structures order "demonstrate their positioning activity" (*ibid.*: 66, my translation). Her results are very much in line with my findings in the present book. The investigation of a wider range of formats with *denken*, concerning past tense and concerning its combination with additional lexical material, allows me to complement Wessel's (2019) results with additional form-function pairings containing the same verb introducing direct reported thought.

Further studies have also highlighted how speakers report the mental states of others, such as during driving trainings (Broth *et al.* 2019). Usually without quotative, the reporting of other drivers' plans, such as taking over the trainee's car, allows driving trainers to support their instructions while a coordinative event is in process during the training. After such coordinative events, trainers may use the report of others' mental states for "educational ends, yielding the generic inferential practices by which competent drivers make contextual sense of others' actions" (*ibid.*: 7). Even though the authors investigate quotations of a third party's purported mental state or thought, their result clarifies that such reports can have a supporting or accounting function that allows speakers to subsequently make publicly available their personal sense-making during a past situation.

Laury *et al.* (2020) in their article on Finnish *mä ajattelin että* ('I thought that') discuss that the pattern is on its way to routinize into a fixed expression. Investigating two corpora of 50-year-old dialect interviews and present-day conversation allows them to show how the pattern has evolved diachronically. Besides the result that *mä ajattelin että* becomes morpho-phonologically reduced over time as a direct consequence of its frequent use, the authors also identify two main interactional uses: prefacing stance-displays and framing "the speaker's expression of her own plans as well as proposals of joint action" (Laury *et al.* 2020: 133).

Barnes and Moss (2007), combining CA and discursive psychology, investigate mundane telephone calls as well as data from psychotherapy and care consultations. They describe the frequently occurring phenomenon of making publicly available private thoughts as "a speaker's resource for handling everyday rational accountability in reporting and explaining actions and events" (Barnes & Moss 2007: 123).[67] The constructions *I think* and *I thought* are two recurring

---

[67] For German, Imo (2007a: 322f) describes that the verb *überlegen* ('to think about') can be used by speakers as 'condensed structures' (Günthner 2006) to make available a (purported) former thinking process in a telling. As in these cases, the verb occurs without its accusative supplement, it is not entirely comparable to structures with *denken* ('to think'). My analysis in Chapter 6 will show that, in my data, it seems as if actions related to accountability are more often introduced with the quotatives *ich habe (mir) gedacht* ('I have thought (to myself)') in German and *je me suis dit* ('I have said to myself') in French.

structures that they find in their data to report thought.[68] They state that private thoughts, when reported, are often accompanied by non-lexical items, such as *oh* and *yeah* as well as response cries (*ibid*.: 123). This observation is in line with research investigating reported speech and thought in taking an affective stance (Ruusuvuori 2013; Selting 2010, 2017; Wessels 2019). Similarly, Fujii (2006) describes in her study of the Japanese quotative *mitai-na* that direct reported thought is "the speaker's virtual reconstruction of somebody's inner monologue, which represents his attitude, reactions, feelings, emotions, etc. (interpretive thought) but was not actually uttered" (*ibid*.: 61). This result also underlines the close relationship between reporting thought and taking a stance.

Most relevant for the present work is what Barnes and Moss (2007) demonstrate regarding the array of actions that reported thoughts can carry out: "[T]he design and production of the thought trajectory […] from emotional reaction to rational questioning and anger is shaped by the local demands of the interactional context" (*ibid*.: 130). This is one of the rare studies in which rational reflection represents one of the possible actions that speakers may accomplish with direct reported thought (but see Laury *et al.* 2020).[69] One of their conclusions is that communicating private thoughts is a "conversational practice" (*ibid*.: 142) and that "reported thoughts are designed to be heard as if they are the genuine embodiment of thinking in action" (*ibid*.: 133).

The above state of the art evidences that previous research recurrently relates direct reported thought to stance-taking.[70] Different stance-taking practices that range from emotionality to rationality have been detected depending on the interactional environment of their occurrence. Due to the high relevance of mainly two types of stance-taking to the present work, I will briefly address relevant features of affective and epistemic stance-taking in talk-in-interaction in the following section.

Direct Reported Thought as a Stance-Taking Device

Formerly discussed under related terms like *evaluation*, *assessment*, or *positioning*, stance-taking became an established notion primarily with the edited book by Englebretson (2007) and Du Bois' (2007) notion of a *stance triangle*.[71] Du Bois (2007) describes stance as follows:

---

[68]   Indirect reported thought where *I think* is syntactically linked to the quote were excluded in their analysis because of its assumed distinctiveness "in terms of both format and function" (Barnes & Moss 2007: 126).

[69]   Other contexts where reporting formats are used in rather rational contexts are, for instance, cases of hypothetical reported speech for advice-giving (Sandlund 2014). These occurrences differ from the here investigated direct reported thought format in that they are not embedded in narrative sequences or reasonings and cannot be classified as stance-taking devices.

[70]   The phenomenon that direct reported speech or thought allow speakers to evaluate what they report, has already been mentioned by Vološinov (1973), who states that direct speech and thought convey a "manner of thinking and speaking, involving the author's evaluation of that manner as well" (*ibid*.: 133).

[71]   The term *positioning*, which may at first seem as adequate as *stance* regarding the actions carried out with direct reported thought, has been discarded (in favor of *stance*) because of its close relation to the concept of identity (Deppermann 2015: 369) points out, "[p]ositioning theories […] approach facets of identity in the way they are accomplished in and by discourse" (*ibid*.). This dynamic view of identity has been especially put forward by Bamberg (1997) and stands, according to Deppermann (2015), in contrast to more normative sociological (Mead 1934) or psychological theories (Ricœur 1990) of positioning. In the present work, I use the term *stance* to avoid any connotation related to identity.

> Stance is realized, in the usual case, by a linguistic act which is at the same time a social act. The act of taking a stance necessarily invokes an evaluation at one level or another, whether by assertion or inference. […] Moreover, since language is reflexive […], and participants routinely monitor who is responsible […] for any given stance, the very act of taking a stance becomes fair game to serve as a target for the next speaker's stance. (*ibid*.: 141)

With this definition, he evokes three important aspects: (1) stance has a strongly interactional component; (2) by taking a stance, speakers position themselves toward something; and (3) stance-taking bears a sequential implicativeness that makes a subsequent response (affiliation/alignment) relevant. One aspect that can be added to Du Bois' claim that stance is usually displayed through "a linguistic act" is that speakers may also display their affective stance through embodied conduct (see Section 3.1.3).

In the present work, two kinds of stances—displayed through reporting formats—are relevant to my empirical analysis: affective stance-taking, a concept closely related to assessments, and a certain kind of epistemic stance-taking, where speakers present themselves as someone "doing being accountable".[72]

Affective stance-taking can be defined as the display of affect and emotion in interaction.[73] In her sequential study on disappointment in talk-in-interaction, Couper-Kuhlen (2009) points out that "[t]hese displays [of affect and emotion, SF] are realized as *embodied practices*", that "[t]he practices are *situated* at specific sequential positions", and that "[t]he practices are interpreted in a *context-sensitive* fashion." (*ibid*.: 96, original emphasis).

Kärkkäinen (2006) has shown that the most explicit affective stance-taking occurs at the end of the telling, which is in line with Goodwin and Goodwin (1992), who state in their detailed analysis of assessing activities that "[s]uch a shift from **description to assessment of described events** in fact constitutes one of the characteristic ways that speakers begin to exit from a story" (*ibid*.: 162, original emphasis).[74] The pressure on the co-participant to respond increases with the intensity of the speaker's affective stance that closes the telling, often after the climax (see also Selting 2017). Speakers may use direct reported speech or thought to enhance the effect of the storytelling's punch line and "[give] the recipient 'access' to the utterance" (Holt 1996: 241). This may also be one of the reasons why, in my data, direct reported thought is systematically bodily and prosodically reenacted.[75]

Speakers may also display less emotional stances (Barnes & Moss 2007), which allow them to report, for instance, their decision-making or formulate a proposal (Golato 2002b; Stevanovic 2013). These non-affective—or, at least, less affective—actions are kinds of epistemic practices, which must be defined more specifically.

---

[72] For a detailed explanation of accountability, see Chapter 5.
[73] On a problematization of the methodology of CA and the study of emotions and affect, see Couper-Kuhlen (2009: 94-96).
[74] The phenomenon of assessments doing closing work has been most prominently shown by Goodwin (1986) and is an important finding of CA and IL literature (Barnes & Moss 2007; Couper-Kuhlen 2007; Deppermann & Reineke 2017; Ehmer 2011: 62; Goodwin & Goodwin 1987: 18ff on post-positioned assessments in general; Goodwin & Goodwin 1992; Park 2018; Selting 2010, 2017; Wessels 2019).
[75] A characteristic that has already been evidenced for direct reported speech as well (Klewitz & Couper-Kuhlen 1999; Selting 2010: 248).

The most documented epistemic practice for German and English is the display of a discrepant assumption, which has already been mentioned by Sacks (1992: 788, Vol. 1) in one of his lectures and has been discussed more thoroughly by Jefferson (see also Kärkkäinen 2003; Kärkkäinen 2012; for German, see Deppermann & Reineke 2017). This practice will not be treated in the present work because it does not fulfill the formal or functional features of a direct reported thought (see Section 3.1.1 for cross-linguistic formal features). For instance, the practice described by Jefferson (2004) and Deppermann and Reineke (2017) occurs in non-narrative sequences and does not stage affectivity or rationality.

One documented action that speakers can carry out with direct reported thought that displays the speaker's epistemic stance is accounting (Deppermann & Reineke 2017: 344ff).[76] According to Deppermann and Reineke (2017), speakers can use *ich dachte (mir)* / *ich habe (mir) gedacht* ('I (have) thought (to myself)') "to explicate intentions or reasons for actions in narratives or argumentations" (*ibid*.: 344, my translation)[77]. In their collection of *ich dachte / ich habe gedacht*, accounting is the third most common accomplished action. It is part of five epistemic practices that the authors find in their data. The authors carve out the following formal features as the most common (*cf. ibid*.):

- Syntactically, the complement clause is mostly a dependent main clause.
- In 20% of their occurrences, a response cry separates the matrix verb and the complement clause.
- Clearly subordinated complement clauses are rare (only 4% of their cases).
- Deppermann and Reineke (2017) also mention one sequential feature:

> In contrast to the practice of displaying an evaluative/affective stance, which comment on the climax of the telling, intentions and reasons with *ich dachte* [which includes, in their study, *ich habe gedacht*, SF] are uttered in the complication phase (cf. Labov/Waletzky 1967). (*ibid*.: 344)

In their paper, the authors do not indicate explicitly that such accounts with *ich dachte (mir)* / *ich habe (mir) gedacht* correspond to the format of a direct reported thought—a dependent main clause syntax, a prosodic change between quotative and quote, and the social action format of the quotative projecting the quote—even though the analyzed example shows characteristics of a direct quote. For French, the phenomenon of accounting through direct reported thought has not yet been investigated. The format in French seems to share, however, many formal characteristics with Deppermann and Reineke's (2017) findings on German.

In my data, speakers also display their reasoning while carrying out actions other than accounting. This is why I extended the conceptualization of the action to an epistemic stance "displaying an emergent reasoning." In addition to accounting, speakers also seem to make decisions publicly available or simply share their post-hoc rationalization of reported

---

[76]   Of course, I do not claim that accounting is exclusively done with direct reported thought. There are other formats that speakers use to account for prior actions, such as adverbial structures. Broth *et al.* (2019) have shown, for instance, that driving instructors may use the explanation of other drivers' conduct to "explicate, for the benefit of their trainee drivers, their understanding of some other driver's side of a jointly experienced and managed coordination event" (*ibid*.: 18). On accounting in English, see Antaki (1994); for German, see Gohl (2006).

[77]   That accounts occur in narratives has already been described by Antaki (1994).

circumstances, similar to what Broth *et al.* (2019) described for the report of others' mental states after a coordinative event in driver trainings.

Research on rationality and reasoning in conversation is scarce. In an article on concessive repair, Couper-Kuhlen and Thompson (2005) show that self-repair is a practice for speakers to display rationality. Speakers can thus share their "grounds for backing down" and "display that they are being accountable" (*ibid.*: 285). Being accountable is particularly relevant to the present work, as speakers seem to make publicly available a reasoning behind an action to "[further] credibility, rationality and agreement" (*ibid.*); They are *doing being accountable* in a way that is related to how Robinson (2016a) describes accountability:

> [O]ne sense of 'accountability'—which I termed 'the account-ability of conduct's action'—involves interlocutors' abilities to form and ascribe (i.e., recognize and understand) possible actions, which are themselves orchestrations of practices of conduct, which embody relevance rules, or normative structures or reasoning and normative patterns of conduct. (*ibid.*: 11)

What transpires from this description is that speakers and recipients can make best guesses about possible actions in conversation. To do so, speakers possess mutually shared knowledge about the machinery of social interaction, which manifests itself in "practices of conduct" (see *supra*). With direct reported thought, speakers can display their knowledge about the accountability that arises from preceding talk.

Consider the following example, in which *je me dis* is added post-hoc as a reversed pseudo-cleft. Elisa tells a friend that she wanted to appeal her grade on an oral exam she had to pass the previous semester because the grade did not satisfy her.

Ex. 3.5: Faire recours / appeal (Pauscaf 11, 12min34)

```
14        (1.3)

15 ELI: de toute façon c'est dur de faire recours contre un oRAL- h°
        anyway       it's difficult to appeal against an oral exam

16 GER: ouais [NON pis&
        yeah  no then

17 ELI:      [c'est ce que j'me DIS.
              that's what I say to myself
```

The reasoning in this excerpt (line 15) is an account for why Elisa did not appeal the grade even though it did not satisfy her. The pause at line 14 shows that the sequence could be closed at this point. After Germain's lack of uptake, Elisa delivers an account for not appealing. In overlap with Germain's confirmation, Elisa incrementally adds *c'est ce que je me dis* ('that's what I say to myself'). This post-hoc framing allows Elisa to mark her general assertion retrospectively as an opinion by situating it as inner speech. A possible reason for this public display could be Germain's prior insufficient responses to Elisa's telling (not in transcript).

3.3.   Summary

This chapter has reviewed existing literature on formal and functional features of direct reported thought. Formally, direct reported thought can be treated together with direct reported speech,

because, grammatically, there is a tendency to juxtapose two, not-necessarily-clausal parts that are not linked with a complementizer; quotative and quote form a bipartite structure, and deictic and temporal features are parallel in direct reported speech and thought formats. However, as little research as there may be on direct reported thought, my literature review also evidenced that direct reported speech and thought differ in function. The review of the concept of *footing* (Section 3.2.2), the challenges that arise for the analysis of embodied conduct (Section 3.1.3), and the interactional uses of direct reported thought (Section 3.2.3) demonstrated that a distinction between direct reported speech and thought may be fruitful.

Analyzing direct reported thought linguistically and multimodally allows me to extend Goffman's concept of *footing*. As multimodal research has shown, speakers make ample use of bodily resources, in addition to language while quoting. However, what previous research has not investigated is how bodily resources contribute to making shifts in footing recognizable for co-participants. I argue that the speaker's multimodal conduct allows them to mark shifts in footing or perform footings simultaneously, thus accomplishing two actions at the same time (e.g., accounting verbally while displaying a negative stance through a facial expression). To my knowledge, previous literature has not focused on the contribution of embodied conduct in indicating shifts in footing, especially concerning direct reported thought.

4. Why a Cross-Linguistic Analysis?

The basis of the present work is the observation that the German *ich denke (mir)* in present and past tense, when used as a quotative introducing direct reported thought, cannot be translated to the French *je pense*. Instead, in my French data, thought is reported with *je me dis* ('I say to myself') in present and past tense, *je suis là* ('I am there'), and *j'étais là* ('I was there'). It is not through lexical equivalence but through the comparability of actions in everyday talk that I was able to find a suitable translation of German *ich denke (mir)* in present and past tense as a quotative. Cross-linguistic IL aims to find such converging social actions or organizational features between languages to explore the overarching infrastructure of everyday talk (Schegloff 2006; Sidnell 2009). The IL approach explicitly encourages comparison between languages as Couper-Kuhlen and Selting (2018) state in their comprehensive book: "Interactional Linguistics advocates a cross-linguistic approach to the study of language in social interaction" (*ibid.*: 355). They explicitly call for more cross-linguistic work in the field (*ibid.*: 96, 208; see also Couper-Kuhlen & Selting 2001: 3) to detect potential language universals.

The present work sets out to add another piece to the puzzle of cross-linguistic interactional organization by investigating the social action format of quotatives projecting direct reported thought. In what follows, I briefly outline the added value of a cross-linguistic analysis.

The comparison of interactions of two Indo-European languages from two different branches, which are today spoken in several neighboring countries within Europe, may reveal new structural and interactional similarities and differences that have not yet been found through the comparison of written texts. A detailed overview of French and German phonological and grammatical differences and how they emerged over time lies beyond the scope of the present work (for further information, see Dryer & Haspelmath 2013). However, the early separation of the branches of Romance and Germanic languages (Price 1984; Wolff 2004) shows us that comparative work of today's French and German means the comparison of two languages with Indo-European roots but distinct developments ever since. Both languages share certain grammatical features, but there are also differences, such as word order in subordinate clauses, French subjunctive, aspect, etc. What French and German have in common is that they are still used in everyday interaction (compared to Latin) and that they are major languages of their respective branches.

Another advantage of cross-linguistic research is that analyzing one language can reveal parallel phenomena in other languages (see also Haakana *et al.* 2016: 28). Within linguistics, Linguistic Typology is most known to uncover larger structural and formal parallels between languages. Unlike typological studies that often analyze obsolete languages or languages from small speech communities outside the Indo-European language family, the present project compares two Indo-European languages that share a certain number of features. I thereby investigate a phenomenon in languages of two speech communities that may also share certain norms and values in social interaction.

Generally speaking, IL is unconcerned with "the particularity of any [language, SF] system […] and the elements of which it is composed" (Sidnell 2009: 8), but with the overall organization of interaction. Uncovering this "general structure" of a language calls for a micro-

analysis of turns and sequences of few instances. The research focus thereby diminishes, as does the number of languages that can be investigated in cross-linguistic IL studies (see also Section 4.2.2).

Setting actions and the resources that speakers use to accomplish these actions as the point of interest, cross-linguistic IL does not focus on preestablished grammatical categories; instead, "by looking across different languages/communities we can see the way in which the same interactional problem is solved through the mobilization of different resources" (*ibid*.). The special endeavor is to scrutinize resources that speakers use for recurrent actions in interaction. In the case of direct reported thought, I therefore chose to analyze the German *ich denke (mir) / ich dachte (mir) / ich habe (mir) gedacht* ('I think / thought / have thought (to myself)') and the French *je me dis / je me suis dit* ('I say/have said to myself') and *je suis là / j'étais là* ('I am/was there'), as these were the most frequent verbal patterns introducing direct reported thought in my data while accomplishing comparable actions (for a detailed description of the choice of verbs, see Chapter 5).

In the remainder of this chapter, I will first give a short historical overview over the development of cross-linguistic research over the last 70 years, when anthropological and ethnographic studies started to be interested in the everyday sociality of communities. To point out distinctive features of comparative CA and IL, I will first present how they developed out of anthropological and ethnographic approaches (Section 4.1). I will then briefly highlight Schegloff's principles of interaction (Section 4.2) before presenting important results from pragmatic typology (Section 4.2.1) and cross-linguistic CA and IL (Section 4.2.2). I will end the chapter with a brief conclusion (Section 4.3).

4.1. Cross-Linguistic Analysis of Social Interaction

The cross-linguistic and cross-cultural interest of anthropology and ethnography lay ground for the endeavor of CA and IL to find overarching organizational mechanisms in social interaction that are used language and culture independently.

Cross-linguistic CA/IL shares its general interest in understanding mechanisms of communication in various speech communities with mainly three approaches: anthropology (Ochs *et al.* 1996), ethnography of communication (Gumperz 1977; Hymes 1962; Saville-Troike 2003), and cross-cultural pragmatics (Blum-Kulka *et al.* 1989). One major difference, especially between cross-cultural pragmatics and CA/IL, however, is the general understanding of social actions (which are conceptualized as "speech acts" in cross-cultural pragmatics; *cf.* Blum-Kulka *et al.* 1989). Sidnell (2009) criticizes that in cross-cultural pragmatics, actions like requests or apologies "are treated as though they were objective categories of social action providing a template through which one can view interactional data" (*ibid*.: 15). Cross-linguistic CA and IL have shown, however, that these categories are not template-like but vary between languages; even though the action, like an apology, is understood as such by members of the same speech community, the resources, such as excusing oneself, may strongly vary between languages. The approaches have also shown that researchers may categorize an action as, say, a question, when interlocutors actually understand the question to be a request or compliment or as affective stance-taking.

A second difference is that what cross-cultural pragmatics categorizes as actions are often preset categories that are presented as if they were (and would emerge) independently from what precedes in the interaction. CA and IL instead define "practices of speaking" that consider their temporal unfolding in time:

> A focus on practices of speaking, rather than 'action types,' suggests that whatever action a given utterance (or part of an utterance) is understood by the participants to be doing is a contingent outcome. Participants understand some practice of speaking as doing a particular action by considering it in relation to the local sequential context in which it is produced (i.e., after a recognizable news report, an excuse, a suggestion, etc.). In that sense, categories of action are themselves the product of local interpretive work. (Sidnell 2009: 15)

What echoes in Sidnell's comparison between cross-linguistic CA and cross-cultural pragmatics is the methodological specificity of CA not to attribute categories, concepts, or actions as long as they are not made relevant by participants in the data. This methodological process also contrasts anthropological and ethnographic studies, which may also use questionnaires or participatory observation as methodological tools.

Specific to comparative CA and IL is also independence from socio-cultural information of the speaker. Compared to cross-cultural pragmatics, ethnography of communication, or linguistic anthropology,

> CA shifts the burden of evidence by requiring researchers to show that putative 'characteristics' of the participants (such as race, ethnicity, gender) have some 'procedural consequentiality' or demonstrable relevance for the participants themselves in terms of the specific ways in which the interaction is organized (Sidnell 2009: 9)

If the speaker's "characteristics" are not somehow addressed or made relevant by the interlocutors, then the CA/IL approach refrains from drawing conclusions about the speakers' (or interlocutors') sociocultural background.[78]

In the present work, I follow Sidnell (2009) in his argumentation that "particular languages provide specific resources and also establish unique constraints for the organization of interaction" (*ibid*.: 10). At the same time, because all languages used in interaction reflect human sociality, there is also a certain amount of underlying "language- and culture-independent principles" (*ibid*.: 5) that structure everyday talk. The next section presents quantitative and qualitative research dedicated to the comparison of interactions, starting with Schegloff's (2006) generic principles of social interaction.

4.2.   Comparing Interactions: Generic Organizational Principles

A growing body of research in pragmatic typology (Rossi *et al.* 2020) and cross-linguistic CA and IL shows that generic organizational principles exist independent of language and culture. Without investigating several languages specifically, Emanuel A. Schegloff established four "*generic* organizational principles of interaction" that he categorizes as "plausible candidates for universal relevance" (Schegloff 2006: 85). These cross-linguistically valid principles are

---

[78]   Recent CA/IL research has shown that there are indeed instances where socio-cultural issues are addressed and/or made relevant (Robles 2015; Robles & Kurylo 2017; Weerd 2019; Whitehead 2015, 2019; Whitehead & Stokoe 2015), especially in the field of membership categorization analysis (Stokoe 2012).

minimization, *progressivity*, *nextness*, and *recipient design* (for a brief summary, see Couper-Kuhlen & Selting 2018: 552ff).

(1) *Minimization* stands for the principle that, as a default rule, speakers have "initially the right (and responsibility) to produce one TCU to completion" (*ibid*.), or one TCU before the next (earliest possible) turn-taking can take place. Every further TCU from the same speaker must be negotiated if not previously announced (e.g., before a story; on story prefaces, see Goodwin 2015). One sequence consists of at least two turns. If more than one turn is produced, the sequence is considered an "expansion" (*cf.* Schegloff 2006: 85).

(2) *Progressivity* describes the principle that there is recipient orientation to a rapid delivery of a next piece of talk. Schegloff (2006) argues that "recipients orient to each next sound as a next piece in the developing trajectory of what the speaker is saying or doing" (*ibid*.: 86). The principle of progressivity is interrupted, for example, when problems of understanding must be treated before moving on. Whereas Schegloff's brief presentation of this principle suggests that there either is progressivity or not, Raymond (2016) suggests in his PhD thesis a "scalar view of progressivity in interaction" (*ibid*.: 136; for a detailed discussion, see 136ff).

(3) *Nextness* refers to the mechanism behind the basic unit for sequence construction: the adjacency pair. Nextness means that there is always a relation between one turn and its prior turn as long as speakers do not mark a turn as independent from what precedes it. As Couper-Kuhlen and Selting (2018) put it, "[t]alk-in-interaction is processed by reference to nextness: one type of self-initiation of repair, for instance, is located before the next word or next sound of a word" (*ibid*.: 553).

(4) *Recipient design* is the principle that speakers, while talking, "design" their talk according to the recipient's knowledge, role in the conversation, or multimodal conduct (Kaukomaa *et al.* 2015). This design is not a conscious process of thinking about how to assemble words or sounds. According to Schegloff (2006), recipient design means that

> [t]he things one talks about with another are selected and configured for who that other is […]. And how one speaks about them—what words, reference forms, and so forth are to be used—is also shaped by reference to who the recipient relevantly is at that moment for this speaker, at this juncture of this interaction. (*ibid*.: 89)

Speakers build their turns according to the degree of assumed (shared) knowledge, politeness, deontic principles, etc. For instance, when talking to a language learner, a native speaker may talk slower than they would with another native speaker. In the same way, tellings are designed for recipients; speakers may add or remove information so recipients can follow along.

Later studies in pragmatic typology confirmed Schegloff's principles from a quantitative perspective, often working with multiple languages (Section 4.2.1). Recurrently, the focus of research for these quantitative studies was first detected by smaller qualitative studies in cross-linguistic CA/IL, some of which I present in Section 4.2.2.

4.2.1.   Pragmatic Typology

Quantitative analyses, especially in pragmatic typology (Rossi *et al.* 2020: 9ff), have shown that interaction is organized in numerous languages around basic organizational principles, such as turn-taking, repair, and recruitments. Stivers *et al.* (2009) investigate the organization of turn-taking in indigenous communities as well as in more widely spread languages. Starting

from the claim of anthropological research about the striking differences between the turn-taking system (Sacks *et al.* 1974) across cultures, they show that there is a "general avoidance of overlapping talk and a minimization of silence between conversational turns" (Stivers *et al.* 2009: 10587). Their result is that, cross-linguistically, the gap between turns is on average 250 ms. In the ten languages that the authors investigate,[79] there are, of course, variations: In Danish, the average response time is the longest, with 469 ms (*ibid*.: 10588), whereas in Japanese, the average response time is only 7 ms (*ibid*.). Despite this degree of variation, the authors show that the basic organizational principle of turn-taking happens systematically and without larger lapses.[80] Schegloff (2006) already pointed out that a difference in gap length between turns may also be a perception problem[81]: "[W]hat differs between them [two interlocutors of two different cultures, SF] is not that their turn-taking practices are different or differently organized, but the way they 'reckon' the invisible, normative beat between one turn and the next" (*ibid*.: 73).

Dingemanse *et al.* (2015) show in their investigation of twelve languages from eight language families that the conversation-organizational phenomenon of repair shows cross-cultural consistency. Other than repair being frequent (one occurrence per 1.4 min in all the investigated languages), the investigated languages also share the same three types of repair initiation—open requests, (*Huh?*; for a detailed study on *huh*, see Dingemanse *et al.* 2013), restricted requests (*Where?*), or restricted offers—and use repair in similar contexts (*cf.* Dingemanse & Enfield 2015). Additionally, Dingemanse and Enfield (2015), in their special issue on other-repair, bring together authors investigating repair in ten languages, regarding intonation, grammatical resources, preference, and repetition (for a qualitative investigation of other-repetition in five different languages, see also Rossi 2020). In their introduction, Dingemanse and Enfield (2015) state that

> [a] key move […] is the recognition that formats for repair initiation form paradigm-like systems […] are ultimately language-specific, and that comparison is best done at the level of the constitutive properties of these formats. These properties can be functional […] as well as sequential […]. (*ibid.*: 96)

Along these lines, Fox *et al.* (2009) find overarching principles for self-repair depending on word-length during repair initiation. The authors are interested in the place, within a word, where speakers initiate repair. Their results "suggest that there is a great deal of variation with respect to favored sites of initiation but that most of the variation can be accounted for by a few simple interactional factors" (*ibid*.: 59)—namely, creating a temporal delay, or the "late decisions to initiate repair" (*ibid*.: 99).

Cross-linguistic studies on repair thus demonstrate that repair may be an essential part of the organizational principles in everyday talk but that the phenomenon varies. What Dingemanse and Enfield (2015) describe as a "paradigm-like system" can also be applied to the patterns investigated in the present work; the social action format of the quotative projects an open

---

[79]   These languages are Danish, Italian, Dutch, Tzeltal, ‡Ãkhoe Haiǁom, English, Yélî-Dnye, Japanese, Lao, and Korean.

[80]   On lapses and accountability in conversation, see Hoey (2020).

[81]   Stivers *et al.* (2009) also hypothesize that differences in tempo when speaking may lead to a *perception* of longer pauses between two turns, when, in fact, the gaps between turns are no longer related to surrounding talk (*cf. ibid.*: 12587).

slot—the quote—which may be filled with an abundance of formats, assembling themselves to a paradigm of possible "quotables," including sounds, body movements, and verbal utterances.

As these studies show, repair and turn-taking are the very image of human sociality. Speaking alternately or correcting oneself and others are crucial systematics of everyday talk that have developed out of the need to solve mundane problems:[82]

> Everywhere turns-at-talk are constructed and opportunities to speak distributed, courses of action are launched and co-ordinatively managed, troubles of speaking, hearing and understanding are located and their repair attempted. These commonalities suggest that, for all the diversity we see, people everywhere encounter the same sorts of organizational problems and make use of the same basic abilities in their solution to them—a capacity for reading other's intentions, anticipating and projecting actions, calculating inferences and processing information available to them (see Levinson 2006, Schegloff 2006). (Sidnell 2009: 3)

These quantitative studies show that certain mechanisms in everyday talk are common across interactions in languages that have likely never been in direct contact. This evidence indicates that, not language contact, but human sociality forms the origin of interaction-organizational structures: "[H]uman beings across societies may be using language in similar ways to both structure their environment and socialize novices" (Ochs 1996: 409).

In contrast to pragmatic typology, the qualitative approach of cross-linguistic CA and IL focuses on micro-phenomena, such as grammatical patterns, multimodal packages, or convergent actions across a small set of languages. Zooming in on practices in few languages allows these studies to describe how a certain phenomenon works on several levels of granularity (Schegloff 2000), such as the prosodic, syntactic, actional, and multimodal.

4.2.2.   Cross-Linguistic CA and IL

One of the three objectives that Couper-Kuhlen and Selting (2018) list for IL is that

> it aims at *cross-linguistic analysis and comparison* of […] practices in order to determine both how interactional exigencies shape language structure and use in social interaction, and how language and language type impinge on the details of the organization of social interaction. (*ibid*. 16, original emphasis)

Although quantitative studies allow for investigating large data sets, qualitative studies are often necessary to bring an investigated phenomenon to light. Cross-linguistic qualitative studies in IL and CA often compare a small number of languages (usually two but also up to five). They focus primarily on actions or linguistic structures. Concerning actions, there has been research on sequence organization in general (Kendrick *et al.* 2014), question-response sequences (Steensig & Drew 2008), repetition (Rossi 2020), repair (Egbert 1997; Moerman 1977), affectivity (Couper-Kuhlen 2011), and projection (Couper-Kuhlen & Selting 2018: 40ff). Other studies focus on grammatical structures or linguistic formats more generally, such as the grammatical formats of requests (Zinken & Ogiermann 2013), increments (Luke *et al.* 2012; Ono & Couper-Kuhlen 2007), phrases (Couper-Kuhlen & Selting 2018: 416ff), clauses

---

[82]   Evolutionary linguistic studies also argue that gestures, especially deictic gestures, trace very far back in human evolution (Levinson & Holler 2014).

(Laury *et al.* 2019),[83] routinized formats like *I don't know* (Pekarek Doehler *et al.* 2021), *I mean* (Laury & Okamoto 2011a), or turn-initial *yes but* (Steensig & Asmuß 2005), or discourse markers like NU/NÅ (Auer & Maschler 2016) and *okay* (Betz *et al.* 2021)[84].

One of the most studied grammatical phenomena in CA and IL, also cross-linguistically, are increments. In their seminal volume on this format, Ono and Couper-Kuhlen (2007) bring together cross-linguistic studies on turn continuations beyond a transition-relevance place. Investigating various languages from different language families, the collection of papers demonstrates that, because increments have long been investigated (mostly for English and German; Auer 2007: 647; Luke & Zhang 2007; Ono & Couper-Kuhlen 2007: 506), there is a need for a "more dynamic" conceptualization that "does more justice to the emergence of turns and turn units as a multi-modal, multi-layered process" (Auer 2007: 647). For example, in English conversation, increments allow speakers to create a second transition-relevance place after a lack of uptake by the recipient (Ford & Fox 2002). This use does not play a role in Navajo (Field 2007) even though the linguistic format is similar (adverbial phrases or unattached nominal phrases).

In their cross-linguistic paper on the basic linguistic unit of clause, Laury *et al.* (2019) compare Finnish and Japanese talk-in-interaction with traditional grammatical descriptions of clauses (from Finnish and Japanese grammars but also from scientific papers). Until then, most descriptions of clauses had been established for English. With the growing primacy of anglophone linguistics over recent years, some Finnish and Japanese scholars have increasingly oriented themselves toward this (originally anglophone) unit, especially in Japanese (*cf.* Laury *et al.* 2019: 373). Consequently, the grammatical unit of clause[85] has been adopted even though the main characteristic—namely, containing a predicate—"and its accompanying elements" (*ibid.*: 366) is rare in Japanese (*ibid.*: 377). In their Japanese data, participants also do not orient to the clausal unit in everyday talk (*ibid.*: 386) but to "predicates along with occasional overt NPs" (*ibid.*: 393). In Finnish, however, 60% of the utterances of the data are clausal, and speakers and recipients orient to clauses as organizational units. With their sequential analysis of Finnish and Japanese everyday talk, Pekarek Doehler (2011) show that Finnish works similarly to English concerning the clausal unit in interaction whereas Japanese does not, despite prior work suggesting the contrary (*cf. ibid.*: 388). Their work demonstrates that it is essential to examine grammatical categories across languages and compare them using mundane talk.

---

[83]   There are, of course, numerous studies outside the field of CA and IL in which clauses and especially clause-combining have been investigated, mainly in language typology (D'Hertefelt & Verstraete 2014; Gast & Diessel 2012; Louagie & Verstraete 2016; Schmidtke-Bode 2014; Schmidtke-Bode & Diessel 2017; Verstraete 2008). I can name only a few of these studies here, as an extensive literature review lies outside the scope of this work.

[84]   What these studies have in common is that they investigate actions or (grammatical) phenomena that contribute to the organization of everyday talk by starting either from a specific function or from a form. Both methods have advantages and disadvantages. Focusing on an action shows that the same action can be accomplished through different (verbal and embodied) resources that speakers deploy in each language. In contrast, investigating a specific linguistic item demonstrates parallels of form while (eventually) bringing to light differences in function between languages.

[85]   For a discussion of the notion of *clause* across languages, see Couper-Kuhlen and Selting (2018: 355-357).

Zinken and Ogiermann (2013) also examine grammatical formats and their relation to social actions, starting their cross-linguistic investigation by looking at the practice of speakers requesting an object in everyday conversation. What they find is that there are certain contexts in which Polish and British English speakers use equal formats: the imperative format and the polar-question format. Their CA approach allows the authors to identify that these two formats carry out two actions depending on their "home environment," or "situations in interaction in which several cues converge to support the selection of one of the request forms across the two languages" (*ibid*.: 257). While the imperative format, both in Polish and British English, is used when a speaker "*draws on* the other person's cooperation" (*ibid*.: 261, original emphasis), the polar question allows a speaker to "[*enlist*] the other person's assistance" (*ibid*., original emphasis). The authors show that both formats exist in Polish and British English to accomplish specific (and related) social actions and that they are similarly distributed in their specific home environments. Because Zinken and Ogiermann (2013) focus on requests instead of on a specific grammatical format, they are also able to show that there are "mixed environments" (*ibid*.: 257), in which speakers of Polish and British English resort to divergent grammatical formats to accomplish the action of a request. In sum, Polish speakers resort more often to the imperative format than do British English speakers.

Yet another way to investigate several languages is to depart from a specific linguistic (or multimodal) pattern, thus showing its parallel function cross-linguistically. Pekarek Doehler *et al.* (2021) investigate dispreferred responses in second position in Czech, French, Hebrew, Mandarin, and Romanian on a verbal level but also multimodally.[86] The authors show that speakers deploy, turn-initially in a responsive action, an equivalent to the English *I don't know* while averting or maintaining averted their gaze of the interlocutor who has produced the first pair part. This first pair part consists of various actions, such as assessments, informings, or requests. Pekarek Doehler *et al.* (2021) show that, cross-linguistically, speakers use the "bimodal assembly" (*ibid*.: 21) of verbal resources and gaze to "retroactively display resistance to the constraints set by and/or disagreement with the stance conveyed by the immediately preceding action, and prospectively project a dispreferred response" (*ibid*.). In this use, as a token prefacing a dispreferred response, *I don't know* recurs in morpho-phonologically reduced form (*ibid*.: 13), suggesting a high degree of routinization of turn-initial *I don't know* in second position, coupled with gaze aversion.

The level of detail with which this cross-linguistic work has been conducted is exemplar for the present study. Despite the phenomenon not being equally frequent in the five investigated languages (*cf*. Pekarek Doehler *et al.* 2021: 6), the density of characteristics forming this "multimodal assembly" shows that [*I don't know* + averted gaze] in second position is a recurrent practice for a dispreferred response. The present work proceeds similarly to Pekarek Doehler *et al.* (2021); assembling several (verbal and embodied) features that characterize

---

[86]   Cross-linguistic multimodal studies are still rare and have only started to be published in the last few years (see, for instance, Pekarek Doehler *et al.* 2021 for a multimodal analysis of *dunno* in six languages). Studies investigating multimodal packages, or "assemblies," as Pekarek Doehler *et al.* (2021) call them, already form a larger body of research (Auer 2021; De Stefani 2013; Haddington 2006; Hayashi *et al.* 2002; Iwasaki 2011; Kärkkäinen & Thompson 2018; Keevallik 2013; Li 2014; Mondada 2007, 2014b; Pekarek Doehler 2019; Skogmyr Marian 2021; Stoenica & Fiedler 2021; Stoenica & Pekarek Doehler 2020; Stukenbrock 2018, 2014).

quotatives introducing direct reported thought in telling sequences, I will show that, even though the quoting devices are not lexically equivalent in French and German, speakers still carry out similar actions in both languages.

4.3.   Conclusion

In this chapter, I have stressed that a distinctive feature of cross-linguistic CA and IL is, in contrast to anthropological approaches (Section 4.1), their focus on overarching organizational principles in interaction (e.g., turn-taking or repair as well as specific actions like requesting or demonstrating) regardless of the speaker's cultural or ethnic belonging. This has been demonstrated by quantitative and qualitative studies in both fields (Section 4.2). The main argument of cross-linguistic CA and IL is that "universal" interactional principles exist because communities are organized as social groups in which certain actions, such as requests, demonstrations, or repair, occur as expressions of sociality. Independent of culture, language, and ethnicity, these actions are necessary for interactants to solve interactional problems, which is why they resort to generic principles of social interaction (Schegloff 2006).

Starting from actions that speakers carry out, cross-linguistic research in CA and IL enables us to identify parallels in action formation between languages. Especially IL draws on analyses that also rely on the grammar-in-use of each language, enabling an investigation of grammar as an emergent and embodied phenomenon that may form similar patterns for similar actions in many languages.

To my knowledge, quotatives introducing direct reported thought have not yet been investigated cross-linguistically. I will demonstrate that speakers of French and German carry out comparable actions with direct reported thought while resorting to lexically distinct quotatives. A qualitative analysis of prosodic, syntactic, sequential, and multimodal features allows me to describe the social action format on several levels of granularity. Following Pekarek Doehler *et al.* (2021) in their procedure, I use a high density of converging features as a way to point out cross-linguistic similarities despite having limited data. I thereby pick up on the general principle of IL that cross-linguistic parallels are possible because speakers must accomplish certain actions in their everyday lives when living in a community. Tellings have likely always been integral to human sociality. Just as reported speech is evident in multiple languages with structural and functional similarities (see Chapter 3 for a detailed overview), reported thought is expected to exhibit similar cross-linguistic parallels.

Of course, with the present work I am only adding one piece to the puzzle. I hope to show how direct reported thought works in French and German, thus opening new perspectives on quotatives for future research.

5.   Data and Methodology

This chapter introduces the two corpora that I used for the present work (see Section 5.1 for German and Section 5.2 for French). It describes how the data was prepared (see Section 5.3), and how collections were built for German (Section 5.4.1) and French (Section 5.4.2).

The investigation of spontaneously produced direct reported thought necessitates the analysis of naturally occurring talk recorded in mundane situations. My corpus of such mundane situations consists of 27h39min of German video recordings and 11h18min of French video recordings. The difference in the amount of data is for two reasons: (1) In German, the speech situations vary from board game sessions and people driving in a car to cooking events and Tupperware parties. In these interactions, there are sometimes long stretches of silence, as participants think about their next move in the game, focus on a cooking activity, etc. In comparison, during the French recordings, participants only talk, without being involved in a multi-activity. (2) The German *ich denke (mir)* ('I think (to myself)') in present and past tense is multi-functional; it is used both to introduce reported thought and as *epistemic parenthetical* (Kärkkäinen 2012; Thompson & Mulac 1991b) or to indicate a discrepant assumption (Deppermann & Reineke 2017; Jefferson 2004). Consequently, *ich denke (mir)* in its present and past form carries out multiple actions (see Section 5.1 for detailed explanations). To achieve a comparable number of occurrences of *ich denke (mir)* ('I think (to myself)') in present and past tense introducing direct reported thought and *je me dis / je me suis dit* ('I say / have said to myself') and *je suis là / j'étais là* ('I am/was there'), a larger amount of German data was necessary.

In the following two sections, I will briefly describe the two corpora in more detail, first for German (Section 5.1) and then for French (Section 5.2).

5.1.   German Corpus

The German data comprises 23 hr 10 min from the FOLK-corpus (Forschungs- und Lehrkorpus gesprochenes Deutsch), generously provided by the Leibniz-Institut für Deutsche Sprache (Mannheim). I also recorded 3 hr 35 min of conversation between friends and couples. These recordings are labelled with the tag *KAFKU* (*Kaffee und Kuchen* 'coffee and cake') because of the setting; interlocutors have coffee together while mostly eating cake or cookies. Additionally, I have 54 min of a cooking-activity (Kochen01) between two participants,[87] generously provided by Prof. Dr. Anja Stukenbrock from the corpus of her SNF-project named "Deixis and Joint Attention: Vision in Interaction" (DeJA-VI; grant no. 100012_179108). Instances of this recording are labelled with the tag *DeJA-VI*. In the German recordings, the number of participants per recording ranges from two to six. All participants from the FOLK, KAFKU, and DeJA-VI data sets signed an informed consent form in which they agreed to the use of the recordings, including images, in scientific papers and presentations at conferences under the condition of anonymizing all names (of people, towns, etc.) that would identify the speakers.

---

[87]   One of the participants in this recording is me. Since the recording was made in summer 2019, when the topic of my thesis had not yet been decided, I consider the integration of the datum as non-problematic.

5.2.   French Corpus

French video-recordings come from the Pauscaf corpus, recorded in 2013 (Pauscaf 1) and 2019 (Pauscaf 2) by doctoral students of the Centre de Linguistique Appliquée at the University of Neuchâtel (Switzerland). Pauscaf 1 consists of 9 hr 6 min of video recordings, and Pauscaf 2 consists of 2h 12 min. All recordings take place at a university cafeteria in the French-speaking part of Switzerland. Participants are students except for some members of the university staff (administrative personnel or teaching staff). All conversations take place between two and four speakers. Again, all participants signed an informed consent form where they agreed to the use of the anonymized recordings, with or without additional image anonymization, in scientific papers and presentations at conferences as well as for teaching.

5.3.   Transcription and Coding of the Data

All data has been transcribed according to the GAT2 transcription conventions (Selting *et al.* 2009). Although transcriptions force the analyst of spoken language to bring sound into a two-dimensional written format that cannot mirror the temporality, simultaneity, and moment-by-moment emergence of conversation, it remains an important tool for the emic perspective of CA because it enables the determination of recurrent (linguistic) patterns and practices based on a stable source (instead of an ephemeral source, like sound or film). The GAT2 transcription system was originally developed for German.[88] In the present work, for the sake of consistency, I transcribe both German and French data according to this conventions.

The transcription of the interlocutors' body movement (i.e., gaze conduct, gesture, facial expressions, posture, etc.), the so-called multimodal transcription, follows the conventions established by Mondada (2018). The list of symbols that were used for different resources can be found in the appendix. Multimodal transcripts were made in *Transana*, a software that allows to integrate time-codes in the transcript and the video-recording to mark the beginning and the end of verbal or non-verbal conduct. By doing so, I was able to transcribe body movements with high temporal precision concerning their placement within a verbal turn.

For coding purposes of verbal and non-verbal conduct, all instances of the analyzed formats were listed in a spread sheet. In other words, for *denken*, which is also used for other actions than reporting thought, all occurrences, including non-quotative occurrences were initially entered into the table. In addition to basic information on each occurrence—recording number, speaker ID, right and left of the context of the target item—I added multiple columns for coding on a morpho-phonological, syntactic, actional, and multimodal level[89].

In a first step, after a careful qualitative analysis of the sequential context of each occurrence, I added, for German, the categorizations whether *denken* is used as quotative or not (for the detailed process, see Section 4.4.1). All non-quotative uses were then excluded from any further analyses. For French, the qualitative analysis revealed that *se dire* and *être là* are exclusively

---

[88]   For an adaptation and translation in English, see Selting *et al.* (2011).
[89]   I am aware that the following categorizations are partly the result of a qualitative analysis, instead of being its basis. For the sake of clarity, I decided, however, to include all categorization-steps in the description of my coding process.

used as quotatives. Consequently, no instances were discarded due to the non-reporting character of the linguistic unit.

Once the collection of quotative with the respective target constructions was determined, I added the morpho-phonological features of each occurrence and included the exact format of realization, for instance `ha_°i°ch mir gedacht` (lit. 'have I to myself thought') where the auxiliary is reduced and latched to the personal pronoun with its vowel in low volume. By doing so, I enabled a subsequent process of filtration according to criteria of morpho-phonological reduction like *ha* instead of *habe* ('have'). I also marked whether the target construction and the following clause were separated by a pause or not, and the placement of the focal accent if it occurred in the target construction.

Each occurrence was then categorized according to the syntactic features of the introduced quote. In German, the possible categories were subordinate clause, non-introduced subordinate clause, dependent main clause, non-clausal unit (i.e., a verbal unit without predicate), and non-syntactic unit (i.e., bodily conduct only).[90] In French, the syntactic categorization is less complex and restricted to the four categories of subordinate clause, main clause, non-clausal unit, and no syntactic unit.

A last coding step included bodily features that either co-occur with the quote or constitute the quote. I categorized the speakers' gaze concerning direction at the beginning of the quote (e.g., 'gaze up' or 'gaze at recipient') as well as, after a careful sequential analysis, whether the gaze is enacted, in other words, whether it is part of a facial expression that displays an affective stance. In additional columns, I coded the head and face movements (e.g., facial expression, headshake), whether there is gestures or not, and whether the gesture is a gesture of depiction or ception (on the differentiation, see Section 6.3.3), and other bodily conduct (i.e., torso movements).

The categorizations according to the above features allows not only for filtration according to morpho-phonologic, syntactic, and bodily features; it is also a crucial tool to detect recurrent patterns in the grammar-body-interface. Due to the possibility of filtration, I was able to systematically verify first observations that I gathered from my emic approach of the data, for instance, whether specific syntactic formats recurrently co-occur with facial expressions or whether one quotative format is more often used with a specific type of gestures.

After this general description of data preparation, I will lay out, for each language and each quotative-format individually, how collections were built based on linguistic criteria.

5.4.   Collection-Building in German and French

5.4.1.   The German Quotatives with *Denken* ('to think') in Past and Present Tense

After the determination of my phenomenon—direct reported thought with *ich denke (mir)* ('I think (to myself)') in present and past tense—I first searched these forms in my corpus. For this initial lexical search, I did not yet consider that *ich denke (mir)* / *ich dachte (mir)* / *ich habe (mir) gedacht* ('I think (to myself) / I thought (to myself) / I have thought (to myself)') are also

---

[90]   For detailed explanations on the syntactic nuances in German, see Section 6.2.3.

used by speakers for other actions than reporting thought, because the determination of certain actions necessitates a closer analysis. Other formats than direct reported thought were excluded from the collection only after a careful grammatical and/or sequential analysis of all occurrences. To build my collections of *ich denke (mir)*, *ich dachte (mir)*, and *ich habe (mir) gedacht*, I searched the FOLK and KAFKU corpora as well as the video from the DeJA-VI corpus for the forms *denke*, *dachte*, and *gedacht*.

I first investigated the KAFKU-corpus and the Deja-VI video by searching different varieties of the standard-German form of *denke*, *denk*, *dachte*, *dacht*, *denkt*, and *gedacht*, because the transcriptions are not normalized. I then searched the FOLK-corpus, which can be found on the DGD-platform[91] (Datenbank für Gesprochenes Deutsch) of the Leibniz-Institut für Deutsche Sprache. As the online platform normalizes dialectal forms, I was also able to find lemmas that do not match the standard German past participle, such as *denkt*, as well as the common occurrences without the final schwa, *denk* and *dacht*.

Because I treat only first-person singular forms in this work, I first excluded all other conjugational forms. I kept the reflexive forms *ich denke mir*, *ich dachte mir*, and *ich habe mir gedacht* because they were frequent (*n* = 126). I excluded all plusquamperfect forms (*n* = 5) for two reasons: (1) with so few cases, I cannot determine whether they form a coherent practice, and (2) I have only one plusquamperfect form in French (*je m'étais dit* 'I had told myself'), which makes comparison difficult. The double perfect form *ich habe gedacht gehabt* (lit. 'I have thought had') did not occur in my data.

In the following, I first present the different steps for building my collection for the analytic past form *ich habe (mir) gedacht*, for the synthetic past form *ich dachte (mir)*, and for the present form *ich denke (mir)*.

5.4.1.1. The Analytic Past Tense: *Ich habe (mir) gedacht* ('I have thought (to myself)')

The form *ich habe (mir) gedacht* occurs 151 times in my corpus. In a first step, some formats (*n* = 29) were filtered out based on their grammatical features[92]. These forms were excluded because they either have an adverb or noun as object (e.g., *schön gedacht*, 'nicely thought'; *ich hab reh gedacht*, 'I thought deer'), occur with a prepositional object *an etwas gedacht* ('to think about something'), or they refer to a preceding object or turn with a deictic *das/des* ('that') or *es* ('it'). They can also occur with a pronominal adverb like *dran* ('of it'). The collocation *zu NP gedacht* ('made for NP') was excluded because it is a fixed expression with a specific meaning in which the original meaning of *thinking* is semantically bleached. The format *ADV/NP hab ich schon gedacht* is a specific format that occurs only in interactions in which

---

[91]   URL: https://dgd.ids-mannheim.de/dgd/pragdb.dgd_extern.welcome.
[92]   Translations in the following tables are non-idiomatic in English because I translated literally concerning the word order. My goal is to demonstrate how the German form is syntactically built when particles, adverbs, etc. are added. In German, none of the following forms is non-idiomatic. Concerning the translated past tenses, I am aware that the translation of *ich habe gedacht* does not have an equivalent to the present perfect tense in English regarding aspectual issues, mainly because aspect plays a minor role in the distribution of German past tense. However, I preferred a word-by-word translation and an eventual mismatch of temporal meaning to aspectual concerns in the English translation to be able to illustrate the structural features of each form. The same holds for tables with *ich dachte*.

participants play a specific game, and the adverb or noun refers to certain keywords that the participants must find and were wrong about. Another two forms were discarded because of the conjunctive mode with *hätte*. One last form was excluded because *hab ich gedacht* ('I thought') is a parenthesis that functions as a hedge, which is not the focus of this work. Seven further forms were excluded because they contained negated forms (*n* = 2; e.g., *ich habe nicht gedacht*, 'I have not thought') or because they form a standalone turn that functions as a response to a prior utterance (*n* = 5). Another three occurrences were filtered out because they occur with a preceding *so* ('like that'), which functions in this (front field) position as a modal adverb, as in *so habe ich gedacht, dass wir es machen* (lit. 'like this have I thought that we do it'). Another three occurrences were excluded because they introduced a subordinate clause with the complementizer *dass* ('that'). In the end, there were 74 non-reflexive and 35 reflexive forms of *ich habe (mir) gedacht* left. I will first treat the non-reflexive and then the reflexive form in more detail.

A table of all remaining occurrences can be found in the appendix (Table 45). The table shows the various patterns of the non-reflexive form in my data and is divided into several sections. In the following, I will give one example for each section: (1) *ich habe gedacht* with preceding items, such as conjunctions and/or adverbs, such as *und dann habe ich gedacht* (lit. 'and then have I thought'); (2) the plain form *ich habe (mir) gedacht* (lit. 'I have (to myself) thought') without additional lexical items; (3) the syntactical formats of verb-first syntax, coordination ellipsis or a pivot, such as *hab ich so gedacht* (lit. 'have I like thought'); and (4) *ich habe gedacht* in combination with particles or adverbs, which are placed between the auxiliary *haben* ('have') and the past participle *gedacht* ('thought'), like *ich habe noch gedacht* (lit. 'I have PART thought').

Until now, I still included those occurrences where *ich habe (mir) gedacht* does not introduce direct reported thought. Other forms, in which *ich habe (mir) gedacht* is used by speakers, such as to display a discrepant assumption (*n* = 20; Deppermann & Reineke (2017)), are not part of the present work for reasons that I have explained in Chapter 3. Thus, in a second step, after the exclusion of forms due to grammatical features, I analyzed each of the 109 remaining sequences in which one of the forms from Tables 45 and 46 occurred. I determined the actions that speakers carry out with *ich habe (mir) gedacht* and classified them according to the actions determined by Deppermann and Reineke (2017: 341):[93]

- Stating a discrepant assumption (*n* = 18)
- Displaying an evaluative/affective stance (*n* = 59)
- Providing intentions or reasons for a prior action (*n* = 28)
- Claiming prior knowledge (*n* = 1)
- Not claiming secure knowledge (*n* = 2)
- Making an alternative proposal (*n* = 1)

---

[93] Translations were taken from Deppermann and Reineke (2020), an English version of the 2017 paper, which presents the same phenomena. The annotation system did not need to be refined or changed because Deppermann and Reineke (2017) built their categories on the FOLK-corpus, which constitutes the main part of the corpus in the present work.

By doing so, I was able to see that only two of these actions are carried out with direct reported thought: displaying an evaluative/affective stance and providing intentions or reasons for a prior action. The criteria that contributed to the determination of the format of direct reported thought were the formal criteria presented in Chapter 3 but also the sequential environment of tellings or reasoning sequences, where direct reported thought introduced with *ich dachte (mir)* / *ich habe (mir) gedacht* mostly occurs in my data.

I then excluded all occurrences that did not introduce direct reported thought. After their exclusion, 53 occurrences of *ich habe gedacht* and 34 occurrences of *ich habe mir gedacht* that function as quotatives remained, all of which are listed in the following two tables, which are divided into the same blocks as Tables 45 and 46:

Form	Number of occurrences
wo ich gedacht habe (where I thought have)	1
und dann habe ich gedacht (and then have I thought)	3
und dann habe ich so gedacht (and then have I like thought)	1
dann habe ich gedacht (then have I thought)	3
und da habe ich halt gedacht (and there have I PART thought)	1
und ich habe erst gedacht (and I have first thought)	1
da habe ich gedacht (there have I thought)	2
und da habe ich schon gedacht so (and there have I already thought like)	2
da habe ich schon gedacht (there have I already thought)	1
und dann habe ich halt gedacht so (and then have I PART thought like)	1
da habe ich immer gedacht (there have I always thought)	1
da habe ich auch gedacht (there have I also thought)	1
weil ich gedacht habe (because I have thought)	1
adverbial(-phrase) habe ich (nur/schon) gedacht (adverbial(-phrase) have I (PART) thought)	3
zwischendurch habe ich schon gedacht so (sometimes have I PART thought like)	1
ich habe gedacht (I have thought)	5
Verb-first syntax: habe (ich) gedacht	9

(have (I) thought)	
coordination ellipsis	5
ich habe noch gedacht (I have PART thought)	1
ich habe ja gedacht (I have PART thought)	1
ich habe echt gedacht (I have really thought)	4
ich habe nur als gedacht (I have just sometimes thought)	1
ich habe halt gedacht (I have PART thought)	2
ich habe aber gedacht (I have but thought)	1
ich habe schon gedacht (I have PART thought)	1
TOTAL	53

Table 1: Occurrences of *ich habe gedacht* introducing direct reported thought

Form	Number of occurrences
wo ich mir gedacht habe (where I to myself thought have)	3
wo ich selber mir so gedacht habe (where I myself to myself like thought have)	1
wo ich mir halt auch so gedacht habe (where I to myself PART also like thought have)	1
und da habe ich mir nur so gedacht (and there have I to myself only like thought)	1
dann habe ich mir halt gedacht (then have I to myself PART thought)	1
dann habe ich mir aber gedacht (then have I to myself but thought)	1
aber dann habe ich mir aber gedacht (but then have I to myself but thought)	1
da habe ich mir gedacht (there have I to myself thought)	2
da habe ich mir erst mal gedacht (there have I to myself first thought)	1
und dann habe ich mir gedacht (and then have I to myself thought)	2
weil ich mir gedacht habe (because I to myself thought have)	3
aber ich hab(s) mir auch gedacht (but I have (it) to myself also thought	1

ich habe mir gedacht (I have to myself thought)	2
Verb-first syntax: habe (ich) gedacht (have (I) thought)	6
ich habe mir jetzt auch so gedacht so (I have to myself now also like thought like)	1
ich habe mir auch schon gedacht (I have to myself also PART thought)	2
ich habe mir auch gedacht (I have to myself also thought)	1
ich habe mir einfach gedacht (I have to myself just thought)	1
ich habe mir extra gedacht (I have to myself specifically thought)	1
ich habe mir gestern auch gedacht (I have to myself yesterday also thought)	1
ich habe mir schon die ganze zeit gedacht (I have to myself already the whole time thought)	1
TOTAL	34

Table 2: Occurrences of *ich habe mir gedacht* introducing direct reported thought

Comparing Tables 45 and 46 (see appendix) with Tables 1 and 2, or comparing the numbers of all practices established by Deppermann and Reineke (2017) to only those where direct reported thought is introduced, yields the following observations:

(1) The number of occurrences with preceding elements for *ich habe gedacht* (*n* = 23) and for *ich habe mir gedacht* (*n* = 18) remains stable.

(2) Out of 17 plain forms, only five function as quotatives; the plain form (i.e., occurrences without additional lexical items) of the reflexive form remains the same but is generally low (*n* = 2).

(3) Out of 19 verb-first syntax forms for *ich habe (mir) gedacht*, 15 introduce direct reported thought—in other words, the number remains relatively stable. When including the coordination ellipsis, 20 out of 24 occurrences introduce direct reported thought.

This quantitative comparison clarifies that speakers seldomly use *ich habe (mir) gedacht* without additional lexical items (or without verb-first syntax) to introduce direct reported thought. The reflexive form tends to be primarily used to report thought. The almost stable number of occurrences that are preceded by elements like *then*, *there*, or *and* enhances the argument that there is a functional correlation between the form—namely, *ich habe (mir) gedacht*—co-occurring with additional lexical items or having verb-first syntax (including instances of a coordination ellipsis) and its function of introducing direct reported thought. So far, a difference between the reflexive and non-reflexive forms beyond the higher tendency of

the reflexive form to introduce direct reported thought does not seem to exist. This will also be confirmed by the analyses in Chapters 6 and 7.

5.4.1.2.   The Synthetic Past Tense: *Ich dachte (mir)* ('I thought (to myself)')

Ich dachte (mir) occurs 133 times in my corpus. I applied the same steps as for *ich habe (mir) gedacht* when building a collection of *ich dachte (mir)* introducing direct reported thought. Interestingly, *ich dachte (mir)* is, in my corpus, is never used with a pronominal object, like *ich dachte an x* ('I thought of x'), or a pronominal adverb (*daran*, 'of it'), as it is the case with *ich habe (mir) gedacht.* Generally, there is much less variation with *ich dachte (mir)*; I do not find any fixed expressions similar to *so gedacht sein* ('to be meant to be'). Instead, the plain form is much more frequent (*n* = 50, out of which two are clause-final).

There are, however seven forms that I excluded due to grammatical features, like the presence of an NP (*n* = 2; e.g., *reh dachte ich auch*, 'deer thought I also'), because they contained negated forms (*n* = 2; e.g., *ich dachte nicht, dass*, 'I did not think that') or because of their occurrence in a pseudo-cleft construction (*n* = 2; e.g., *was ich kurz dachte ist*, 'what I briefly thought is'; for a complete table, see Table 47 in the appendix).

The remaining forms (*n* = 110 non-reflexive forms, *n* = 17 reflexive forms) were divided, again, into several groups: (1) *ich dachte (mir)* with preceding items such as conjunctions or adverbs, like *und dann dacht ich so* (lit. 'and then thought I like'); (2) the plain form *ich dachte (mir)* ('I thought (to myself') without any additional lexical items; (3) the syntactical formats of verb-first syntax and coordination ellipsis (e.g., *dachte so*, 'thought like'); (4) *ich dachte (mir)* in combination with post-positioned particles or adverbs, like *ich dachte jetzt einfach* ('I thought now just'). The complete tables of all forms can be found in the appendix (Table 48: non reflexive form; Table 49: reflexive form).

I then analyzed the sequences in which these forms occurred and classified them according to Deppermann and Reineke's (2017) categories.[94] The frequencies for *ich dachte (mir)* are distributed as follows:

- Stating a discrepant assumption (*n* = 73)
- Displaying an evaluative/affective stance (*n* = 47)
- Providing intentions or reasons for a prior action (*n* = 4)

Following this step, I excluded, as with *ich habe (mir) gedacht*, occurrences that did not introduce direct reported thought. Similar to *ich habe (mir) gedacht*, this led to an exclusion of forms in which *ich dachte (mir)* was used by speakers to display a discrepant assumption (*n* = 73). It is only in this use that *ich dachte (mir)* introduces a complement clause in conjunctive mode (*n* = 15). The following tables show the remaining occurrences in which *ich dachte (mir)* introduces direct reported thought:

Form	Number of occurrences
und da dacht ich so	1

---

[94]   Three of the practices that allow speakers to display an epistemic stance—claiming prior knowledge, not claiming secure knowledge, and making an alternative proposal—occur together only three times in my corpus.

(and there thought I like)	
und ich dachte irgendwann (and I thought at some point)	1
und dann dacht ich so (and then thought I like)	1
und dann dachte ich (and then thought I)	2
und dann dacht ich auch so (and then thought I also like)	1
dass ich so dachte (that I like thought)	1
und ich dachte so (and I thought like)	4
dann dachte ich (then thought I)	1
ich dachte so (I thought like)	1
aber dann dachte ich so (but then thought I like)	1
dann dachte ich so (then thought I like)	1
und dann dachte ich aber so (and then thought I but like)	1
und ich dachte schon so (and I thought already like)	2
und ich dachte halt so (and I thought PART like)	1
jetzt dachte ich noch (now thought I PART)	1
ich dachte (I thought)	3
Verb-first syntax: dachte (ich) (thought (I))	6
coordination ellipsis	3
ich dachte jetzt einfach (I thought now just)	1
ich dachte halt so (I thought PART like)	1
ich dachte grad so (I thought just like)	1
ich dachte eher (I thought more)	1
ich dachte noch (I thought PART)	1

TOTAL	37

Table 3: Occurrences of *ich dachte* introducing direct reported thought

Form	Number of occurrences
und ich dachte mir PART (and I thought to myself PART)	2
und dann dachte ich mir (and then thought I to myself)	1
und dann dachte ich mir so (and then thought I to myself like)	2
und ich dachte mir (and I thought to myself)	1
aber ich dachte mir auch (but I thought to myself also)	1
und ich dachte mir nur (and I thought to myself only)	1
verb-first syntax: dachte (ich) mir (thought (I) to myelf)	1
coordination ellipsis	3
ich dachte mir nur (I thought to myself only)	1
TOTAL	13

Table 4: Occurrences of *ich dachte mir* introducing direct reported thought

Comparing Tables 48 and 49 (see appendix) with Tables 3 and 4, or comparing all functions with only those in which *ich dachte (mir)* introduces direct reported thought, I observed the following:

(1) In only three out of 50 total occurrences does the plain form (i.e., without additional lexical items) of non-reflexive *ich dachte* introduce direct reported thought (including the clause-final *dachte ich*). This difference is not relevant for the reflexive form, *ich dachte mir*, for which I do not have any plain forms in my corpus.

(2) Only six out of 26 total occurrences of reflexive and non-reflexive *ich dachte (mir)* with *postponed* particles or adverbs introduce direct reported thought.[95]

(3) In contrast, 19 out of 21 total occurrences of reflexive and non-reflexive *ich dachte (mir)* with *preceding* adverbs and/or conjunctions introduce direct reported thought.

---

[95]   If I also include in these numbers those occurrences in which, in addition to the postponed particle/adverb, there is also a preceding adverb or conjunction, the numbers are as follows: For *ich dachte*, 21 out of 37 introduce direct reported thought, and for *ich dachte mir*, 7 out of 8. That the difference in number decreases confirms that preceding lexical items are a recurrent feature of *ich dachte (mir)* introducing direct reported thought.

(4) Thirteen of the 17 occurrences of the reflexive form *ich dachte mir* introduce direct reported thought. This small difference in number is relatively similar to the analytic past form, in which 34 out of 37 occurrences introduce direct reported thought.

The numbers show that, similar to *ich habe (mir) gedacht*, the synthetic past form *ich dachte (mir)* introduces direct reported thought when the complement-taking predicate co-occurs with additional lexical items. The stable number of occurrences in this category evidences that, again, there is a correlation between the form and the introduction of direct reported thought. Verb-first syntax, however, does not seem to play a similarly important role in *ich dachte (mir)* as it does in *ich habe (mir) gedacht*.

5.4.1.3. The Present Tense *Ich denke (mir)* ('I think (to myself)')

In total, *ich denke (mir)* occurs 164 times in my corpus. The non-reflexive form is more frequent (*n* = 139) than the reflexive form (*n* = 25). As with the past forms, I first searched for *denke* in my two German corpora. Of the 164 occurrences, nine were dismissed either because they were interrupted or the following talk was inaudible. I first categorized the remaining 155 occurrences according to grammatical features and according to the additional lexical items they co-occurred with. I first excluded those occurrences in which *ich denke (mir)* occurred with a (prepositional) object pronoun or pronominal adverbs (*n* = 12), like *ich denke da auch immer nicht dran* (lit. 'I think there also always not of it') or *das denke ich* (lit. 'that I think').[96] All these occurrences are listed in Table 57 in the appendix. Two other occurrences contained negated forms (e.g., *ich denke nicht*) and were therefore excluded. The remaining 141 occurrences (119 non-reflexive, 22 reflexive) are listed in Table 51 for the non-reflexive form and Table 52 for the reflexive form (see appendix).

In a next step, I analyzed all occurrences in their larger context. Considering sequential, interactional, and bodily parameters, I determined the actions speakers carry out with *ich denke (mir)*. After doing so, I excluded all occurrences in which *ich denke (mir)* was not followed by a direct reported thought[97] but functioned as a hedge, standalone response, or what has been called epistemic parenthetical (Thompson & Mulac 1991a). The remaining occurrences (n = 53) are listed in the following tables:

Form	Number of occurrences
aber ich denke als immer (but I think PART always)	1
aber irgendwie denke ich halt auch (but somehow think I PART also)	1
und ich denke so (and I think like)	1
und ich denke immer so (and I think always like)	1
und jedes mal denke ich	1

---

[96] There are also two forms that occur with a direct object pronoun but were nevertheless followed by direct reported thought and were therefore included: *immer wenn x denke ich da dran* ('always when x I think of it') and *aber wenn ich daran denke* ('but when I think of it').

[97] I did not have any occurrences of indirect reported thought, which is why I did not have to exclude it.

(and every time think I)	
und letzen denke ich so (and recently think I like)	1
und deshalb denke ich so (and therefore think I like)	1
und wo ich denke (and where I think)	1
wo ich denke (where I think)	3
wo ich so denke so (where I like think like)	2
da denke ich (there think I)	1
da denke ich so (there think I like)	1
dann denke ich (then think I)	1
ADV denke ich so (ADV think I like)	1
wenn x denke ich halt (when x think I PART)	1
verb first-syntax	1
coordination ellipsis	8
ich denke (I think)	3
ich denke noch (I think PART)	1
immer wenn x denk ich da dran (always when x think I there of it)	1
aber wenn ich daran denke (but when I of it think)	1
würde ich denken (would I think)	1
TOTAL	34

Table 5: Occurrences of *ich denke* introducing direct reported thought

Form	Number of occurrences
weil ich mir halt denke (because I to myself PART think)	1
und da denke ich mir (and there think I to myself)	1
und dann denke ich mir so	1

(and then think I to myself like)	
und dann denke ich mir (and then think I to myself)	1
und deswegen denke ich mir halt so manchmal (and therefore think I to myself PART like sometimes)	1
und deswegen denke ich mir (and therefore think I to myself)	1
und ich denke mir so (and I think to myself like)	1
und ich denke mir (and I think to myself)	1
wo ich mir denke (where I to myself think)	2
wo ich mir (echt) denke (where I to myself (really) think)	1
dann denke ich mir (then think I to myself)	1
da denke ich mir auch so (there think I to myself also like)	1
Verb-first syntax	5
ich denke mir so (I think to myself like)	1
TOTAL	19

Table 6: Occurrences of *ich denke mir* introducing direct reported thought

When comparing Tables 51 and 52 (see appendix) and Tables 5 and 6, the following points are of particular interest:

(1) Only three out of 21 total occurrences of non-reflexive *ich denke* in its plain form (i.e., without additional lexical items) introduce direct reported thought.

(2) When occurring with *preceding* lexical items, 30 ($n = 17$ non-reflexive, $n = 13$ reflexive) out of 46 ($n = 31$ non-reflexive, $n = 15$ reflexive) total occurrences of reflexive and non-reflexive *ich denke (mir)* introduce direct reported thought. Interestingly, the number of reflexive forms remains relatively stable whereas the non-reflexive form is almost halved.

(3) The number of reflexive forms generally remains stable (19 out of 22 total occurrences introduce direct reported thought).

(4) *So* ('like') occurs regularly as a post-positioned lexical item ($n = 18$) and exclusively when *ich denke (mir)* introduces direct reported thought. Five of these 18 occurrences occur with verb first-syntax ($n = 1$) or a coordination ellipsis ($n = 4$),[98] which is why these forms do not show individually in the tables above.

---

[98]   For an explanation of the difference between the two syntactic formats, see Section 6.2.3.

(5) The syntactic format of a coordination ellipsis occurs only with the non-reflexive form and only when *ich denke* introduces a direct reported thought (Proske 2019), as shown in the stable number of the format.

These five observations clarify that, as with the past forms, there is a relationship between the occurrence of *ich denke (mir)* with additional lexical items and its function as a quotative of direct reported thought. This does not mean that when *ich denke (mir)* co-occurs with additional items, it must introduce direct reported thought; however, the tendency of doing so is observable in my data. In the case of *ich denke (mir)*, post-positioned *so* ('like') is a strong indicator of following direct reported thought, with all 18 occurrences introducing direct reported thought. These findings demonstrate that the action of quoting thought is tied to regular co-occurring patterns of the type *und da/dann denke ich (so)* (lit. 'and there/then think I (like)').

5.4.2. The French Quotatives *Je me dis* ('I say to myself') and *Je suis là* ('I am there') in Present and Past Tense

Collection-building for French has been less complicated because of less variety in particles and adverbs and because the functions are more clearly distributed compared to German. I will describe briefly how the French collections were built, starting with the past forms of *se dire*, *je me suis dit* ('I have said to myself') and *je me disais* ('I said to myself').

5.4.2.1. *Je me suis dit* ('I have said to myself') and *Je me disais* ('I said to myself')

In my transcripts, I searched for *disais* and *dit*. Being homonym with other conjugation forms (second-person singular imperfective for *disais* and third-person singular present for *dit*), I first had to filter out all forms that were not first-person singular imperfective or present perfect (passé composé). As the present work focuses on self-quotation, I excluded all forms that were not first-person singular forms. This led to eight occurrences of *je me disais* and 47 occurrences of *je me suis dit.* Five of the eight imperfective forms came from the same speaker, which is why I excluded the imperfective form *je me disais* from my analysis.

All occurrences of *je me suis dit* introduced direct or indirect reported thought. In a second step, I excluded those occurrences that introduced indirect reported thought ($n = 4$) because the focus of the present work is on the direct reporting format. This left 43 occurrences of *je me suis dit.* Other functions that Deppermann and Reineke (2017) describe for German *ich dachte (mir)* and *ich habe (mir) gedacht*, such as the display of a discrepant assumptions, which is frequent in German, did not occur. This practice is not accomplished with *je me suis dit* (neither with *je me disais*) but, as preceding examinations of my data revealed, with *je pensais* ('I thought'; $n = 17$) or *j'ai pensé* ('I have thought'; $n = 2$). The collection-building was simplified by *je me suis dit* being, in this sense, "mono-functional," as, in my data, it always introduces reported thought. Co-occurring lexical items do not seem to be as closely intertwined with the quotative-function of *je me suis dit* as it is the case for German: The most frequent pattern is *je me suis dit* without additional items. One of the recurrent co-occurring lexical items is *pis* ('then'), either with *et* ('and') or alone ($n = 8$). Similar to German, this adverb allows speakers to introduce a next step in a telling or reasoning. This function also holds for *après* ('after/then') ($n = 2$). The conjunction *mais* ('but'; $n = 6$) establishes a contrastive relationship between two

units and is expected to be recurrent in argumentative sequences in which pro and contra must be weighed against each other. All other co-occurring items are isolated cases. A larger data set would be needed to uncover regularities in these patterns.

5.4.2.2.   *J'étais là* ('I was there')

French *j'étais là* (*n* = 60) is as consistent as *je me suis dit* regarding its use in everyday talk. The first-person singular form is, in my data, almost always used as a quotative. When searching for *j'étais là* in my corpus, I found only three occurrences in which the deictic *là* refers to an actual location. These three occurrences were excluded from my collection. Most of the cases of *j'étais là* occur without additional lexical items (*n* = 39).[99] Similar to *je me suis dit*, the most frequent additional lexical items were adverbs like *pis/après* ('then'; *n* = 8). All other co-occurring lexical items only occur once or twice, which is why no conclusions can be drawn concerning clear form-function pairings.

The collection-buiding process for French revealed that additional lexical items exist but are neither as frequent nor as consistent as in German. This result indicates that there may be a lower tendency for *j'étais là* to form linguistic patterns.

5.4.2.3.   *Je me dis* ('I say to myself')

For *je me dis* ('I say to myself'), I searched for *dis* in all my transcripts of Pauscaf 1 and Pauscaf 2. With *dis* being homonym with the second-person singular, I first filtered out all forms that were not first-person singular present, thus keeping the focus of the present work on self-quotation. This yielded 35 occurrences of *je me dis.*

Out of 35 total cases of *je me dis*, I discarded one occurrence because of poor audibility and another because it occurred in the negated form. Five further occurrences introduced a quote with the complementizer *que* ('that') and were therefore excluded from my collection. This led to a total of 28 occurrences for *je me dis* introducing direct reported thought. Compared to German *ich denke (mir)* introducing direct reported thought (*n* = 53), the French complement-taking predicate had a higher relative frequency. While the German quotative occurred, on average, every 47 min 24 s, the French *je me dis* occurred every 24 min 12 s.

5.4.2.4.   *Je suis là* ('I am there')

To filter all occurrences of *je suis là*, I searched my corpus for all occurrences of *suis*, the first-person singular present form of *être* ('to be') and excluded all occurrences in which it was not coupled with the deictic *là* ('there'). This led to eight occurrences of *je suis là* in my corpus. One occurrence has literal meaning, in which *là* refers to an actual place. Two of the seven remaining cases occurred in the negated form and were therefore excluded from the collection. Compared to its imperfective form (*n* = 57), the present form *je suis là* is rare in my data (*n* =

---

[99]   I could also find isolated occurrences of non-first-person patterns, like *elle était là* ('she was there'), *ils sont là* ('they are there'), and *mes parents qui étaient là* ('my parents who where there'). A more recurrent pattern is *tu es là* ('you are there') in present tense, usually reduced to *t'es là* (*n* = 14). The most frequent form in my data remains, however, the first person singular imperfective form—a result that corroborates earlier findings from Cheshire and Secova (2018: 27).

5). I include the present form of *être là* despite this low frequency because the social action format *je suis là* is consistent on a sequential, actional, and multimodal level.

5.5.   Conversation Analysis and Interactional Linguistics

The initial goal of the present work is to find overarching rules that organize and structure everyday talk. One main principle of Conversation Analysis (CA), as stated by Sacks in an early lecture, is that interaction manifests "order at all points" (Sacks 1992: 484, Vol. 1). Participants orient to this orderliness by following commonly shared expectations and/or norms.

> If the materials (records of natural conversations) were orderly, they were so because they had been methodically produced by members of the society for one another, and it was a feature of the conversations that we treated as data that they were produced so as to allow the display by the coparticipants to each other of their orderliness, and to allow the participants to display to each other their analysis, appreciation, and use of that orderliness. (Schegloff & Sacks 1973: 290)

The investigation of mundane talk represented a massive change in linguistic methodology and disclosed many new research areas. As an approach heavily infused by sociologic ideas, CA focuses primarily on the "machinery" of talk, thus understanding language as a working system structured into turns, actions, and sequences that enable speakers to accomplish actions. The systematics behind this structure are a main research interest of CA.

CA approaches data from an emic perspective. This means that analysts generally do not have a certain phenomenon in mind and instead try to find it in their data. Conversation analysts practice what they call "unmotivated looking," or repeatedly listening to audio or watching video to discover underlying structures or organizational phenomena. These phenomena do not have to be a specific grammatical format but can be a certain sequence, specific actions, or a multimodal conduct.

To prove the underlying order of social interaction, CA can resort to only the data itself. The conduct of interlocutors alone can show the analyst if actions are recognized as such (on the methodological tools of next turn proof procedure, see Sacks *et al.* (1974: 728f); on accountability, see Section 5.5.2).

Interactional Linguistics (IL), a more recent approach that draws heavily on CA, has a stronger linguistic dimension. The approach concentrates on verbal and multimodal resources that speakers use in interaction. As Couper-Kuhlen and Selting (2018) formulate it in the introduction to their comprehensive book:

> Interactional linguistics […] believe that turns, actions, and sequences are accomplished and made interpretable by the systematic use of linguistic resources; consequently, they make the linguistic forms deployed by ordinary speakers in everyday and institutional encounters the focus of their attention. (*ibid.*: 4).

The stronger focus on linguistic resources does not mean that interactional organizational questions are neglected in IL research; the idea is simply to focus on the interplay of the diverse layers that constitute spoken language, such its social character emerging from the interaction with others.

> [L]inguistic productions—since they take shape in interaction—can no longer be conceptualized as the product of a single speaker. Instead sentence and clause production, indeed speech production in general, must be thought of as an interactional achievement (Goodwin 1981). […] Thus linguistic structures are both emergent in interaction and heavily context-sensitive, in that their use reflects—and may even contribute to creating—conversational structure. (Selting & Couper-Kuhlen 2001: 5)

IL thus studies the interplay of prosody, morphology, semantics, and syntax as well as embodied resources, like gaze and gesture, and contextual information. The present work is more in the tradition of IL than of CA for the following reasons:

(1) The social action format of quotatives introducing direct reported thought is, first and foremost, a linguistic phenomenon, which must be explored on a linguistic (i.e., grammatical) level.

(2) At the same time, because reporting activities in everyday talk often involves staging (Yule & Mathis 1992), or bodily reenactments, speakers use both verbal resources and a multitude of resources in the projected quote. The interplay of these resources and their respective temporal emergence in interaction necessitates closer investigation.

(3) Cross-linguistic work is an explicitly stated goal of the interactional approach. Working with French and German, the present work meets this research aim.

The present work builds nevertheless on organizational principles of talk that have been described early on by CA research. In what follows, I will introduce some of the basic rules according to which mundane talk is structured.

5.5.1.   Organizational Principles in Interaction

In their seminal paper on the organization of turn-taking, Sacks *et al.* (1974) establish "a basic set of rules governing turn construction, providing for the allocation of a next turn to one party, and co-ordinating transfer so as to minimize gap and overlap" (*ibid*.: 704). This set of rules mainly describes what happens at a transition-relevance place (TRP), or a potential moment of turn-taking; at the first TRP of a first turn-constructional unit, the selected next speaker "has the right and is obliged to take next turn to speak" (*ibid*.). If the "current speaker selects next" mechanism does not work, a potential next speaker can self-select or the current speaker can continue. If the current speaker continues, they can do so until the next TRP, where the same mechanisms apply.

To create TRPs, turns and turn-constructional units must be recognizable by co-participants. Couper-Kuhlen and Selting (2018) define turns as "utterances that speakers produce when they occupy the floor" (*ibid*.: 34). The length of turns can vary considerably between single utterances (*single-unit turns*; e.g., one word or sound) and several utterances (*multi-unit turns*). Both kinds of turns are composed of at least one turn-constructional unit (TCU), which can itself be composed of a complete sentence, a clause, a word, a sound, or a body movement (e.g., a nod). Following this logic, one TCU is sufficient to form a turn. As Schegloff (1996) points out,

> By 'turn-constructional unit', it may be recalled, we meant to register that these units *can* constitute possibly complete turns; on their possible completion, transition to a next speaker becomes *relevant* (although not necessarily accomplished). (*ibid*.: 55, original emphasis)

The TCU as an analytical unit has been much debated.[100] Selting (1998), for instance, argues that not every TCU ends with a TRP. Taking so-called "compound-TCUs" and storytellings as examples, she calls for a clearer distinction between TCUs and TRPs (1998: 7f). Syntactic information may overrule a potential TRP at the end of a TCU, as in "if-then clauses" (Lerner 1996), but prosody (Schegloff 1996; Selting 1998: 20ff) or "big-packages" such as storytellings that are projected by pre-sequences (Goodwin 2015; Mandelbaum 2013) may equally do so. Ford *et al.* (1996) refrain, just like Sacks *et al.* (1974), from defining TCU. Instead, the authors stress that TCUs are an interactional achievement that "makes TCUs and turns impossible to precisely define and precisely predict. Rather than a static set of resources to be deployed, TCUs are best understood as ephiphenomena [sic] resulting from practices" (Ford *et al.* 1996: 428f). Because of the incremental nature of turns and TCUs, they are not ready-made units that are uttered, but rather built, in interaction, reflecting the emergent and temporal character of everyday talk (see also Chapter 2).

In the present work, the TCU is the central unit to describe the emergent character of storytellings and, within these tellings, the function of the pattern [quotative + direct reported thought]. If I argue that speakers accomplish a specific action with the reporting format, it is necessary to determine the (actional and grammatical) unit deployed for the action. Even though such an exact determination may not always be possible (see Section 3.1.2), the practices speakers resort to mostly are. According to Ford *et al.* (1996), the recognition of the practice enable understanding TCUs, sometimes retrospectively, and the resources that speakers deploy within them.

5.5.2.   Adjacency Pairs, Turn-Taking and Next Turn Proof Procedure

According to Schegloff and Sacks (1973), "a pervasively relevant issue (for participants) about utterances in conversation is 'why that now', a question whose analysis may […] also be relevant to finding what 'that' is" (*ibid.*: 299). "That" describes actions that speakers carry out with the help of language and/or bodily conduct. Actions become observable for analysts either through the next turn proof procedure (Sacks *et al.* 1974: 728f) or the accountability that arises from a turn (Robinson 2016a).

The next turn proof procedure builds on the idea that talk is organized in *adjacency pairs*— for example, a question (the so-called *first pair part*) and a response (*second pair part*). In third position, there may be a confirmation of the delivered response. Taken together, these two (or three) turns form a sequence—here, a question-answer sequence. The next turn proof procedure as an analytical tool describes the fact that a second pair part of an adjacency pair reveals the co-participant's understanding of the first pair part—for example, a response, demonstrating an understanding of the prior turn as a question. This methodological procedure prevents analysts from hastily ascribing fixed actional categories to certain grammatical formats without considering contextual information. Understanding the recipient's response as analytic

---

[100]   For further discussion on TCUs, see Goodwin (1981) for a general discussion, Chafe (1993) and Selting (1996a) for prosody and TCUs, and Ford *et al.* (1996) for the argument that prosody and gesture may be just as important as syntax for the organization of turns in TCUs and for the determination of TRPs. Couper-Kuhlen and Selting (2018) compare the analysis of TCUs and actions in everyday talk and convincingly argue in favor of the analytical unit of TCUs (*ibid.*: 35ff).

evidence for action ascription allows analysts to evade a potentially subjective interpretation of the data, thus adhering to the primordial "emic perspective" of CA and IL.

In addition to next turn proof procedure, the notion of accountability as another analytical tool has received increased attention at least since Robinson's (2016) edited book on the topic.[101] Originally established in Ethnomethodology, Garfinkel (1967) describes accountability as follows:

> I mean observable-and-reportable, *i.e.* available to members as situated practices of looking-and-telling. I mean, too, that such practices consist of an endless, ongoing, contingent accomplishment; that they are carried on under the auspices of, and are made to happen as events in, the same ordinary affairs that in organizing they describe; that the practices are done by parties to those settings whose skill with, knowledge of, and entitlement to the detailed work of that accomplishment—whose competence—they obstinately depend upon, recognize, use, and take for granted; […]. (*ibid.*: 1)

Following Garfinkel's description, accountability is a more largely applicable "methodological resource" (Sacks *et al.* 1974: 728) than next turn proof procedure. First, it is not restricted to clearly definable adjacency pairs, which are not always as clearly shaped as a question-answer sequence (for instance, incrementally built storytellings). Second, accountability leaves space for several resources to come into play to hold a co-participant accountable.[102] In the present work, both next turn proof procedure and accountability are applied as methodological tools, depending on the recipients' conduct. Because tellings often encompass multi-unit turns, it is not always easy to define first and second pair parts. However, quotatives introducing direct reported thought may be involved in the organization of turn-taking, especially when occurring at the ends of tellings. Depending on the sequence, either next turn proof procedure or accountability may be more appropriate as an analytic tool.

The turn-taking machinery, the major organizational structure of conversation, heavily relies on the way turns are designed by speakers and oriented to by recipients:

> What is at stake in 'turn taking' is not politeness or civility, but the very possibility of coordinated courses of action between the participants (e.g., allowing for initiative and response)—very high stakes indeed. (Schegloff 2006: 72)

For conversation to run smoothly, recipients and speakers offer affordances (Moerman & Sacks 1988); for speakers, this means that their turns must be designed "so that attention to them permits projecting their future" (*ibid.*: 186), or permits projecting a potential end of their turn so the recipient can take over and start the next (potentially responsive) turn. Participants must constantly monitor the current speaker because their prediction of turn-completion is only a "*possible* prospective completion" (*ibid.*: 187, original emphasis), and their upcoming second pair part can always only be a "best guess" (Couper-Kuhlen & Selting 2018). Moerman and Sacks (1988) conclude that "the elaborated understandings which conversational sequencing requires are done locally, immediately, publicly, accessibly, sanctionedly, and continually"

---

[101]   Research has, however, addressed the relationship of turn design, action, and accountability through action formation and action ascription (Enfield & Sidnell 2017; Levinson 2013).

[102]   Recent research has shown that this sometimes is the case, such as when the facial expressions of the recipients lead to the speaker's shift of stance (Kaukomaa *et al.* 2015).

(*ibid*.: 185), thereby anticipating concepts that were later developed in linguistics: emergence, projection, temporality, action formation and ascription, and multimodality.

In the next chapter, I will apply the methodologies of CA and IL by analyzing quotatives in French and German talk-in-interaction.

6. Direct Reported Thought with Quotatives in Past Tense

6.1. Introduction to a Cross-Linguistic Analysis

In the present chapter, I investigate direct reported thought, introduced by quotatives in past tense, in French and German talk-in-interaction. As pointed out earlier, French and German do not use lexical equivalents as quotatives to introduce their purported past thoughts. In German, the most frequent quotatives in past tense that speakers use are *ich dachte (mir) / ich habe (mir) gedacht*[103] ('I (have) thought (to myself)'), whereas in French, speakers use most frequently *je me suis dit* ('I have said to myself') and the pattern *j'étais là* ('I was there').[104]

I will show in this chapter that direct reported thought is a recurrent format for taking a stance with French and German quotatives being functionally distributed according to the kind of stance that speakers display. The comparability of German and French arises from the congruences in action formation instead of the lexical similarity between quotatives.

Concerning the format [quotative + quote], speakers seem to juxtapose in both French and German two parts of a bipartite structure without syntactic link. In line with my explanations in Chapter 2, I consider the pattern [quotative + quote] as emerging in real time and thus as adaptable to the contingencies of an interaction. However, because the pattern is systematically used in telling sequences and in recurrent sequential positions, routinization of the social action format that projects direct reported thought is to be expected in my data. I will address this issue in the conclusion of the analysis of each quotative more specifically.

In the following sections, I will present a grammatical, sequential, and multimodal analysis of quotatives in past tense introducing direct reported thought. I will first analyze the German quotatives before moving on to French. I will close the chapter with a comparison of the direct reported thought format in both languages.

6.2. Reporting Thought with *Ich dachte (mir) / Ich habe (mir) gedacht* ('I (have) thought (to myself)')

6.2.1. Introduction

As I have shown in Chapter 5 on collection building for *ich dachte (mir)* and *ich habe (mir) gedacht*, the use of these two formats as quotatives introducing direct reported thought seems to be only possible when they co-occur with additional lexical items like *da/dann* ('there'/'then'), *so* ('like') or *und* ('and') or when occurring with verb-first syntax.

The tendency of German to form chunks of complement-taking predicates in combination with adverbs or particles has been studied for several verbs by Imo (2007a). With a

---

[103] In what follows, I will refer to the German expressions involving *ich dachte (mir)* or *ich habe (mir) gedacht* generically as *ich dachte (mir)* or *ich habe (mir) gedacht* even though they occur in patterns with additional lexical material like particles, adverbs, or conjunctions (see Section 6.3.4 for a detailed account).

[104] In what follows, I will use the quotatives in italics without translating them repeatedly except for cases where lexical characteristics are crucial for my argument. Additionally, at the beginning of each section, the translation will be repeated for the sake of the meaning's availability.

Construction Grammar approach, he investigates constructions like *sag mal* ('say PART'), *wie soll ich sagen* ('how should I say'), or *sagen wir mal so* ('let's put it this way,' lit. 'say we PART like'). He shows that these constructions are routinized chunks that allow speakers to carry out specific actions in conversation, depending on which particle/adverb is added to the matrix verb. Such routinized forms have also been prominently investigated by Stein (1995) under the notion of *Formelhaftigkeit* ('formulaicity,' my translation; see Section 2.3.1 for a detailed presentation of this notion). One major criteria for a pattern to count as routinized that has been raised especially in the realm of Construction Grammar is the pattern's frequency (for a discussion, see Schoonjans 2017; see also Zima & Bergs 2017). The forms that I investigate are recurrent but would not count as frequent from a strictly quantitative perspective. Filatkina *et al.* (2020), however, argue that "[w]hat seems to be a crucial factor for the emergence of formulaic patterns is not so much just the frequent use of a pattern but its frequent use in a specific situation of communication" (*ibid.*: 7). The patterns that I investigate do occur consistently in a specific context—namely, in the context of tellings, where they introduce direct reported thought. This observation is similar to what Kärkkäinen (2010) describes for English *I thought.* According to her, there is

> strong evidence that *I thought* is part of a social action format or a recurrent linguistic routine or conversational pattern used by speakers for taking a stance: […] it is followed by an explicitly stanced phrase or clause […], or by direct reported speech or thought embodying a clear speaker stance. (*ibid.*: 213)

The next section delivers quantitative (yet not statistic) evidence for the German quotatives occurring in chunks.

6.2.2. Linguistic Pattern with *Ich dachte (mir) / Ich habe (mir) gedacht*

In total, there are 50 cases of *ich dachte (mir)* ('I thought (to myself)') introducing direct reported thought (37 non-reflexive, 13 reflexive forms) and 87 cases of *ich habe (mir) gedacht* ('I have thought (to myself)') doing so (53 non-reflexive, 34 reflexive forms). When speakers use these forms to introduce direct reported thought, they mostly occur in one of the combinations shown in Tables 9 and 10 (for a more detailed table listing all forms and translations, see Sections 5.4.1.1 and 5.4.1.2 and appendix), which shows a condensed overview of the different combinational patterns:

Pattern	Frequency *ich dachte*	Frequency *ich dachte mir*
da/nn dacht ich (so)	5.4% (2/37)	-
und/aber da/nn dachte ich (so)	16.2% (6/37)	23.1% (3/13)
und ich dachte (PRT) so	13.5% (7/37)	-
aber ich dachte auch	-	7.7% (1/13)
und ich dachte (ADV/PRT)	2.7% (1/37)	30.8% (4/13)
verb-first syntax	8.1% (3/37)	7.7% (1/13)
coordination ellipsis	16.2% (6/37)	23.1% (3/13)
ich dachte jetzt einfach	2.7% (1/37)	-
ich dachte (PRT) so	10.8% (4/37)	-

ich dachte PRT	8.1% (3/37)	7.7% (1/13)
jetzt dachte ich noch	2.7% (1/37)	-
no additional item	8.1% (3/37)	-
TOTAL	**100% (37)**	**100% (13)**

Table 7: Combinational pattern with *ich dachte (mir)*

Pattern	Frequency *ich habe gedacht*	Frequency *ich habe mir gedacht*
wo ich gedacht habe	1.9% (1/53)	14.7% (5/34)
da/nn hab ich (PART) gedacht	15.9% (8/53)	14.7% (5/34)
und/aber da/dann hab ich (PRT) gedacht	11.3% (6/53)	11.8% (4/34)
und da/dann hab ich so gedacht	5.7% (3/53)	3.0% (1/34)
und ich habe ADV gedacht	1.9% (1/53)	-
ADV habe ich PRT gedacht	1.9% (1/53)	-
verb-first syntax	17.0% (9/53)	17.6% (6/34)
coordination ellipsis	9.4% (5/53)	-
weil ich gedacht habe	1.9% (1/53)	8.8% (3/34)
ich habe aber gedacht	1.9% (1/53)	-
ich habe auch (so) gedacht		8.8% (3/34)
ich habe PRT gedacht	18.9% (10/53)	11.8% (4/34)
ich habe ADV gedacht	3.8% (2/53)	3.0% (1/34)
no additional item	9.4% (5/53)	5.9% (2/34)
TOTAL	**100% (53)**	**100% (34)**

Table 8: Combinational pattern with *ich habe (mir) gedacht*

The tables show that for *ich dachte (mir)*, patterns with *und* ('and'), *dann/da* ('then/there') and coordination ellipses are recurrent. Also, if all occurrences containing post-positioned *so* ('like') are added up, we arrive at 48.0% (*n* = 19 with the non-reflexive, *n* = 5 with the reflexive form). These percentages are differently distributed for *ich habe (mir) gedacht*: Prepositioned *und* ('and') or *dann*/*da* ('then/there') occur less often with only 33.0% (*n* = 29). The adverbial clause with *wo* ('where') is as frequent as patterns with *dann/da* ('then/there')—namely, 14.7%. Verb-first syntax is even more frequent (17.6%) whereas coordination ellipses do not occur at all. Post-positioned *so* ('like') is considerably less frequent than with *ich dachte (mir)*—namely, 11.5% (*n* = 5 with the non-reflexive, *n* = 5 with the reflexive form). On a purely linguistic level, this quantifying analysis demonstrates that at least [*ich dachte (mir) + so*] can be considered being a recurrent pattern to report thought. This confirms, with a larger data set, Wessel's (2019) results on *ich dachte so* in her exploratory study of *ich dachte* as a quotative of direct reported thought. The fact that there are fewer instances of *ich habe(mir) gedacht* in

combination with *so* may indicate that it is the preterit format, and not the verb *denken* in general, that forms the pattern with *so* when introducing direct reported thought.

In addition to this linguistic evidence, the next section gives an overview on morpho-phonological features that may contribute to pattern-formation.

6.2.2.1. Prosodic and Morpho-Phonological Features of *Ich dachte (mir)* ('I thought (to myself)')

Ich dachte ('I thought (to myself)') often occurs in a morpho-phonologically reduced way and bears certain phonological features: It is pronounced with accelerated tempo compared to the surrounding speech, and does not carry a focus accent.[105] The first table shows an overview of these features, first for the non-reflexive format:

Feature	Frequency
morpho-phonological reduction[106]	86.5% (32/37)
accelerated	51.4% (19/37)
no accentuated syllable	67.6% (25/37)
pause after *ich dachte*	21.6% (8/37)

Table 9: Prosodic and morpho-phonologic features of the quotative *ich dachte*

The table shows that many forms are morpho-phonologically reduced and that in most cases, expressions with *ich dachte* do not carry an accent. These features may be indicators for *ich dachte* being a routinized pattern. As only about half of the occurrences are accelerated, I cannot analyze this feature as a consistent indicator of routinization. It is important to point out that the form itself is already very short, compared to *ich habe (mir) gedacht*. This could be one reason why accelerated forms are less frequent. The pause after *ich dachte* occurs in less than a quarter of all cases, which makes it a less pertinent feature as well. The reflexive form is, in proportion, more routinized on a morpho-phonological level:

Feature	Frequency
morpho-phonological reduction	100.0% (13/13)
accelerated	69.2% (9/13)
no accentuated syllable	100.0% (13/13)
pause after *ich dachte mir*	23.1% (3/13)

Table 10: Prosodic and morpho-phonological features of the quotative *ich dachte mir*

In general, not all features that are usually considered to indicate routinization always co-occur and there is not *one* feature that occurs in all cases. While morpho-phonological reduction and the absence of a (focus) accent are more frequent features, a pause between the quotative and

---

[105] The basis of all following prosodic analyses is an auditory analysis, following the idea of *impressionistic hearing* (Couper-Kuhlen & Selting 1996; Kelly & Local 1989). No software such as Praat has been used to investigate the forms more in detail.

[106] With *ich dachte (mir)*, I included the forms where only the schwa of *dachte* is reduced because this concerns the main verb, compared to *ich habe (mir) gedacht*, where only the auxiliary is concerned. The reduction of the main verb carrying the semantic information may be more meaningful than the reduction of an auxiliary because the latter does not carry the main information. When excluding cases where only the schwa is reduced, only 9 cases count as reduced.

the quote is rather rare[107] and an accelerated pronunciation only concerns about half of all occurrences. Similar to Günthner (2000) describing features of a "reproachful voice," it is impossible to find the one feature making *ich dachte (mir)* a routinized form.

The accumulation and frequent co-occurrence of different prosodic features may lead, however, to a *perceived* routinization. It gives the impression that *ich dachte (mir)* is used by speakers as a pattern that frames the direct reported thought. Similar observations can be made for the analytic past form *ich habe (mir) gedacht.* The following section shows prosodic and morpho-phonological features of the non-reflexive format *ich habe gedacht* (Table 11) and of the reflexive format *ich habe mir gedacht* (Table 12) introducing direct reported thought.

6.2.2.1. Prosodic and Morpho-Phonological Features of *Ich habe (mir) gedacht* ('I have thought (to myself)')

As with the analytic past form, prosodic and morpho-phonological features of *ich habe (mir) gedacht* ('I have thought (to myself)') have been analyzed. The following tables give an overview of the respective frequencies. Table 11 shows the prosodic and morpho-phonological features of the non-reflexive form:

Feature	Frequency
morpho-phonological reduction	69.8% (37/53; 96.2% ($n = 51$) when including cases, where only the schwa on *habe* is missing)
accelerated	52.8% (28/53)
no accentuated syllable	67.9% (36/53)
pause after *ich habe gedacht*	26.4% (14/53)

Table 11: Prosodic and morpho-phonological features of the quotative *ich habe gedacht*

The following table shows the same categories for the reflexive form:

Feature	Frequency
morpho-phonological reduction	52.9% (18/34; 94.1% when including cases, where only the schwa on *habe* is missing)
accelerated	41.2% (14/34)
no accentuated syllable	47.1% (16/34)
pause after *ich habe mir gedacht*	8.8% (3/34)

Table 12: Prosodic and morpho-phonologic features of the quotative *ich habe mir gedacht*

When including the reduction of schwa only, the number of reduced cases in both forms would be high: 83 of 87 forms would count as reduced. As this reduction is very common and cannot necessarily be related to a routinization of the whole pattern, it may be better to exclude forms where only the schwa of *habe* is reduced. When doing so, the numbers decrease: 37/53 for *ich habe gedacht* and 18/34 for the reflexive form only.

---

[107] That there is a tendency of prosodic integration of the quote, or at least of the response cry following the quotative, contradicts Imo's (2007) findings that dependent main clauses—the dominating syntactic format of the quote—tend to be "prosodically non-integrated" (*ibid.*: 48).

Other phonological features are less frequent and cannot be considered as indicators for a possible routinization. It is important to mention, however, that *ich habe (mir) gedacht*, even when it carries an accent, never carries a focus accent when introducing direct reported thought. Instead, the focus accent lies on a syllable within the quote. One recurrent distribution of the focus accent is on the subsequent (and often prosodically attached) response cry. Concerning acceleration as prosodic feature, in my data, only about half of the occurrences are accelerated. The feature of acceleration thus cannot be considered a consistent characteristic. The same holds for the pause after the expression involving *ich habe (mir) gedacht*, which only occurs in about one sixth of all occurrences.

The morpho-phonological reduction and the pronunciation of the expression involving *ich dachte (mir)* / *ich habe (mir) gedacht* under one intonation contour seem to be, in my data, the most important features to at least determine which additional lexical items belong to the quotative and may thus be part of a routinized form.

In the next section, I will address German clause-combining patterns and how they relate to the reported thought format. Positional aspects of the quotative are also addressed.

6.2.3.   Syntactic Features of *Ich dachte (mir)* / *Ich habe (mir) gedacht* ('I (have) thought (to myself)')

6.2.3.1.   The Continuum of Clause-Combining in German

In the following, I will first present an overview of three possible clause-combining patterns[108] in German, all combining, traditionally, two types of clauses. This tripartition is, of course, a simplification of a continuum without clear boundaries:[109] Many forms fill the "in-betweens" of the three patterns that I will present in the next three sections. The two "extremes" of the large scale of forms are prototypically a [main clause + subordinate clause] pattern (maximal indirectness; indirect reported thought[110]) and a juxtaposition of two independent clauses (maximal directness; direct reported thought). One of the formats in between these two extremes has been described as [main clause + dependent main clause / non-introduced subordination] pattern. Some formal features of these three clause-combinational patterns have been described in previous literature.

Maximally Dependent Clause-Combining Patterns

In German, indirect speech and thought have the "prototypical" (Lehmann 1988: 2) syntactic format [matrix clause + complementizer *dass* ('that') + subordinate[111] clause]. In my data, this

---

108   I intentionally use the term *clause combining* in this section because I am focusing here on the actual combination of complement-taking predicate and complement clause.

109   A similar claim has been made by Lehmann (1988), who speaks in his typological study of a continuum reaching from juxtaposed parataxes to embedded structures where the subordinate clause is downgraded to a "particular, well-defined constituent within the main clause" (*ibid.*: 4f). Specifically, regarding reported speech, Plank (1986) describes clause-combining patterns as a continuum with a whole "complex of features" that distinguishes individual shapes (*ibid.*: 302, my translation).

110   Even though I excluded the indirect reported thought format in the present work, I present this format of maximally dependent clause-combining for completeness.

111   In this section, I will stick to the terminology of *main clause* and *subordinate clause* for the sake of clarity. My argumentation is mainly based on Auer (1998) and Imo (2007), who use this terminology to establish

format has been explicitly excluded because it can only occur with indirect reported thought, which is not the focus of the present work. To clarify the idea of a continuum, I will nevertheless briefly outline the features of this format.

In this pattern, the subordinate clause is linked to the matrix clause with the conjunction *dass*, which is why the subordination is called *eingeleiteter Nebensatz* (lit. 'introduced subordination') in German (Auer 1998; see also Imo 2007a: 44f). This type of subordinate clause is maximally embedded in a bi-clausal structure because it has, "in another, superordinate clause, the function of a part of speech" (Hentschel & Weydt 2003: 420, my translation).[112] With the complement-taking predicate *denken* in general, the subordinate clause has the syntactic role of a direct object that fills an empty argument position. This clause combining pattern entails, in German, a change of word order in the subordinate clause with the verb moving from 2nd to final position. Consider the following, invented, example illustrating the verb movement in the subordinate clause:

1) Ich weiß. Marlene macht jeden Sonntag einen Spaziergang.
   Lit.: *I know. Marlene does every Sunday a walk.*
   'I know. Marlene goes for a walk every Sunday'
2) Ich weiß, dass Marlene jeden Sonntag einen Spaziergang macht.
   Lit.: *I know that Marlene every Sunday a walk does.*
   'I know that Marlene goes for a walk every Sunday'

Concerning the degree of dependence, a prototypical subordination is maximally dependent.[113]

On a semantic level, the subordinate clause also fits the main clause verb's valency by "[filling] a valency position" (Gast & Diessel 2012: 5; see also Imo 2007a: 46). This means that the main clause—here, the quotative—would be semantically (and syntactically) incomplete without the quote. Vice versa, Imo (2007a) explains, resorting to Zifonun *et al.* (1997: 2236f) and Verhagen (2001), that the subordinate clause would not be "interpreted in the same way" (Imo 2007a: 43, my translation) neither, were the matrix clause omitted. In the case of a quotation, this would mean that the quote may not be identifiable as a quote without the framing main clause, which does not seem to be necessarily the case unless prosodic or bodily resources mark it as such (Klewitz & Couper-Kuhlen 1999). In addition to its syntactic and semantic features, patterns with a subordinate clause do also have pragmatic features that Auer (1998) described in his article on clause combining patterns in German.

---

a continuum of syntactic dependence. Despite the debates around the syntactic category of subordination (see, among others, Fabricius-Hansen & Ramm 2008; Imo 2007a; Laury & Okamoto 2011a; Matthiessen & Thompson 1988), it allows me to clearly distinguish between formally divergent syntactic patterns which would otherwise be merged in the concept of "complement clause" (for a definition of complement clause see Schmidtke-Bode & Diessel 2017; see also Section 2.3).

[112] Plank (1986) also states regarding syntactic relations of direct reported speech that "[s]ubjects and direct and indirect objects are more subordinated to the structural clause center, the verb, than oblique objects and adverbials, insofar as they are governed as necessary additions (due to the matrix verb) and accordingly also in their case marking and other relational markings by the matrix verb." (*ibid.*: 303, my translation). On the problem of merging objecthood and complementation, see Thompson (2002).

[113] In the specific case of indirect reported speech or thought, the dependence is also marked through additional subjunctive mood in the subordinate clause.

On a pragmatic level, an introduced subordinate clause assigns the matrix clause a high relevance. According to Auer (1998), *abhängige Nebensätze* (lit. 'dependent subordinate clauses,' corresponding to Imo's "introduced subordination") contain information that are more or less known to interlocutors, thus being "presupposing in relation to the recipient" (*ibid.*: 10, my translation). The introduced subordinate clauses are, accordingly, pragmatically backgrounded (*pragmatische Relevanzrückstufung*, lit. 'pragmatic relevance downgrading'). The matrix clause is "asserting in relation to the recipient" (*ibid.*, my translation), thus "shifting the center of relevance to [itself]" (*ibid.*: 11, my translation). Considering Auer's (1998) argument, this means *ex negativo* that syntactically unmarked subordinate clauses (i.e., without the complementizer *dass*) would allow speakers, in the case of direct reported thought, to pragmatically *foreground* the quote, thus leaving the matrix clause—the quotative—in the background. This feature may contribute to the process of foregrounding the quote as the rhematic part of the direct reported thought, thus making it pragmatically more relevant according to Auer's (1998) explanations.

In-Between Clause-Combining Structures

On a continuum from maximal dependence to maximal independence, two types of clause-combinational patterns in German have been determined: so-called *non-introduced subordinate clauses* and so-called *dependent main clauses*. The terminological differentiation between the two patterns still lacks a "commonly accepted use" (Imo 2007a: 46, my translation), despite there being some features that seem to allow a formal differentiation.

Imo (2007a) argues from a Construction Grammar point of view that both syntactic forms belong to the same construction, but constitute "two poles" which are not "clearly separable one from another" (*ibid.*: 46, my translation). Consider the following two examples, which are, for the sake of clarity, invented examples:

1) Sie sagte, *sie träfe* (SUBJ II) euch an der Feier *dort*. (*non-introduced subordinate clause*)
   ('She said she would meet you at the party there')
2) Sie sagte, *ich treffe* (IND) euch an der Feier *hier*. (*dependent main clause*)
   ('She said I meet you at the party here')

The difficulty of separating the two forms lies, as the examples above demonstrate, in their common set of features, which are, according to Imo (2007a: 48):

- no subordinating conjunction (syntactic)
- verb in the subordinate clause in second position (syntactic)
- the subordinate clause fills the open slot of the argument structure of the matrix clause (syntactic)[114]
- the subordinate clause fills the valency position of the matrix clause verb (semantic)

---

[114]   This point is specifically relevant when it comes to the possible (in-)dependence of both syntactic structures. The two clauses in these clause-combining patterns may seem rather independent from the perspective of the complement clause. From the perspective of the matrix clause, however, both formats, the independent subordinate clause and the dependent main clause still fill the empty space of the argument of *denken* and are thus still bound to the main clause.

- the matrix clause is downgraded in relevance, whereas the subordinate clause is upgraded (pragmatic; *cf.* Auer 1998: 10ff)

Imo (2007a) also points out two differences. On a syntactic level, while the unintroduced subordinate clause is in subjunctive mood (*träfe* in so-called *Konjunktiv II*) and with deictic shifts (see Excerpt 1, the shift of the personal deictic *sie* and the local deictic *dort*), the dependent main clause does have neither of them (*treffen*, *ich*, and *hier*). Also, on a prosodic level, there seems to be the tendency of the unintroduced subordinate clause to be prosodically integrated into the matrix clause, whereas this is not the case for the dependent main clause (*ibid.*: 48).[115] When applying the above-mentioned features to my data, I find that only the syntactic complementation pattern of the dependent main clause exists. The following table shows the distribution:

Quotative	Frequency[116]
ich dachte	59.5% (22/37)
ich dachte mir	46.2% (6/13)
ich habe gedacht	66.0% (35/53)
ich habe mir gedacht	61.8% (21/34)

Table 13: Dependent main clause pattern with *ich dachte (mir)* / *ich habe (mir) gedacht* as quotatives

The unintroduced subordinate clause with subjunctive mood and a deictic shift does not exist in my data. The numbers show, however, that dependent main clauses can maximally account for about the half of all my occurrences. Despite the formal consistency of the syntactic pattern in those cases where it is used, there are other formats that *ich dachte (mir)* / *ich habe (mir) gedacht* can introduce in my data, which neither Auer (1998) nor Imo (2007) were able to treat in detail in their overviews, simply because these formats must be classified outside of a clausal structure and were not part of the authors' focus of research.

Syntactically, questions (*n* = 0 for *ich dachte (mir)*; *n* = 7 for *ich habe gedacht*; *n* = 2 for *ich habe mir gedacht*) can also be analyzed as dependent main clauses because they are realized as independent syntactic structures that could occur as such without *ich dachte (mir)* / *ich habe (mir) gedacht* but, in the case of direct reported thought, still fill the slot of their empty argument and valency position. If the questions were dependent structures, they would (prototypically speaking) be realized as *ob*-structures (*if*-structures; without question word), or with the verb in last position (with question word). Imperatives are not verb-second but verb-first structures. If they occurred as a dependent structure, they would have to be reformulated with the German modal verbs *sollen* ('should') or *wollen* ('want'; Radtke 2015: 99), which is not the case in my occurrences.

This overview of in-between clause-combining patterns has demonstrated that an array of syntactic formats can be introduced by *ich denke (mir)* / *ich dachte (mir)* / *ich habe (mir) gedacht* as a quotative. All the previously mentioned formats can be categorized as loosely syntactically linked to the main clause.

---

[115] As pointed out in Sections 6.2.2.1 and 6.2.2.2, my prosodic analysis evidenced that, in contrast to Imo's (2007) description, dependent main clauses reporting thought are, in my data, still prosodically integrated.

[116] These numbers also include cases where the dependent main clause is preceded by a response cry.

The next section presents yet another syntactic configuration where two clauses are merely juxtaposed without syntactical link whatsoever.

Maximally Independent Clause-Combining Patterns / Bipartite Structures

Maximally independent clause-combining patterns are cases where two units are juxtaposed without any syntactic link. In Chapter 2, I explained in detail why I will call such structures *bipartite structures*. In this section, I will briefly describe maximally independent combining patterns and how they may be important to account for some instances of *ich dachte (mir)* / *ich habe (mir) gedacht* as quotatives.

In my data, I have occurrences where *ich dachte (mir)* / *ich habe (mir) gedacht* introduce only response cries, only bodily movements, or verbless structures. There are also a few occurrences where the quotatives are followed by question words or imperatives.[117]

Quotative	Response cries only	Bodily conduct only	Verbless structures	Question words	Imperatives
ich dachte	19.0% (7/37)	-[118]	8.1% (3/37)	-	2.7% (1/37)
ich dachte mir	23.1% (3/13)	-	15.4% (2/13)	7.7% (1/13)	-
ich habe gedacht	9.4% (5/53)	-	13.2% (7/53)	1.9% (1/53)	-
ich habe mir gedacht	8.8% (3/34)	2.9% (1/34)	2.9% (1/34)	-	2.9% (1/34)

Table 14: Formats of direct reported thought introduced by *ich dachte (mir)* and *habe (mir) gedacht*

There are also some instances, where the quotative introduces a longer telling sequence. Similar to what Günthner (2008a) has described for German *weil* ('because'), that can be used as a discourse marker followed by a long telling sequence (*ibid.*: 111), the quotatives *ich dachte (mir)* / *ich habe (mir) gedacht* allow speakers to subsequently produce a multi-unit turn. This format is, however, rare, as the following table shows.

Quotative	Number of occurrences
ich dachte	2.7% (1/37)
ich dachte mir	7.7% (1/13)
ich habe gedacht	5.7% (3/53)
ich habe mir gedacht	5.9% (2/34)

Table 15: Frequency of *ich dachte (mir)* and *habe (mir) gedacht* introducing multi-unit turns

In these cases, the quotative is only pragmatically linked to what follows. In general, clauses that are only pragmatically linked to preceding or following clauses have been described under

---

[117] The numbers of verbless structures, question (words) and imperatives include cases where the quote is preceded by a response cry. Also, 1 imperative and 8 questions that consist of complete clauses are counted as such.

[118] But there is one occurrence where a dependent main clause is added subsequently.

several notions such as routinization (Günthner 2018; Kärkkäinen 2012; Pekarek Doehler 2011, 2016, 2021b; Secova 2015), sedimentation (Couper-Kuhlen & Thompson 2008; Günthner 2011a; Imo *et al.* 2020), or grammaticalization (Lehmann 2015; Romaine & Lange 1991; Thompson & Mulac 1991a; Traugott & Heine 1991b, 1991a). These notions differ concerning their respective criteria that linguistic patterns must meet to count as "routinized." In the following, I will briefly focus on an additional notion that has been established in German research to describe maximally independent clause-combining patterns: *Formelhaftigkeit* ('formulaicity'; Stein 1995). It may allow me to grasp the status of German *ich denke (mir)* in present and past tense regarding its tendency to form chunks with additional lexical items.

Formelhaftigkeit describes routinized forms that contribute to the social or content organization of a "communication event" (*ibid.*: 19, my translation). Stein (1995) calls such routinized forms "pragmatic idioms" ('pragmatische Phraseologismen'; *ibid.*: 18) which range from single words like *danke* ('thank you') to clausal structures. What he describes as clausal structures may have "semantic or pragmatic functions, which go beyond their linking character" (*ibid.*: 19, my translation). In his chapter on conversation-specific formulas ('gesprächsspezifische Formeln'), Stein (1995: 129ff) addresses their high degree of variability, which makes their classification difficult: They often occur with "filled pauses and—even more often—with modal and structuring particles" (*ibid.*: 136). Later in his work, Stein explicitly says that the combination with other lexical items is characteristic for verbal forms occurring in formulas (*cf. ibid.*: 148). When presenting the spectrum of co-occurring forms, he also mentions *und da/und dann* (*ibid.*: 137), which I recurrently find in my data with *ich denke (mir)* in present and past tense (see Section 5.4.1). *Ich denke* (in present tense) is presented as one of the most frequently used forms among the (also most frequent) first-person singular indicative present forms. However, most strikingly, Stein describes that he has only one perfect form in his corpus: *ich hab (halt) gemeint* ('I have (PART) meant'; *ibid.*: 138). This could indicate that *denken* in past form may have been used less at that time—a hypothesis that cannot be verified, of course, without access to his data.

Stein (1995) also indicates that formulaic constructions are often prosodically separated from what precedes or follows (*cf. ibid.*: 143) and that they can be inverted under specific circumstances—two features that also apply to quotatives with *denken*. One major difference between formulaic patterns and *ich dachte (mir)* / *ich habe (mir) gedacht* as quotatives is their position in the clause. Formulaic patterns can be positioned after as well as parenthetically inserted into a clause (*ibid.*: 142). Because of its presentational function and its projective force as social action format recurrently introducing the same action, *ich dachte (mir)* / *ich habe (mir) gedacht* is, in contrast, exclusively prepositioned in my data.[119]

---

[119]   This also holds for the French quotatives that I investigate.

In Stein's (1995) work, *ich denke*[120] functions predominantly as "Gliederungssignal" (roughly translated as 'structuring signal') which organizes conversation:[121] It manages turn-taking and turn-holding and indicates vagueness ("Vagheitsindikator"; *ibid.*: 239). What is important for the present work is that he treats all *verba sentiendi* in their function as epistemic marker (Imo 2007a, 2011a; see also Günthner & Imo 2003). In its use as a quotative, *ich denke (mir)* in present and past tense does *not* function as epistemic marker or hedge, like German *ich glaube* (Imo 2012b). Instead, in this function, the complement-taking predicate seems to routinize into a speech-situating device (Brünner 1991) that allows speakers to frame their reported thought as such. In other functions, however, *ich denke (mir)* in present and past tense seems to be on its way to routinize into an epistemic hedge (Deppermann & Reineke 2017).

This section suggests that *ich denke (mir)* in present or past tense, when fulfilling the criteria established by Stein (1995), has no syntactic relation whatsoever to what it introduces—a finding that my qualitative analysis will confirm. *Ich denke (mir)* in present and past tense—when used as a quotative—seems to fit the criteria of forming patterns with additional lexical items and of being syntactically independent of what it introduces. However, regarding its positional flexibility, *ich denke (mir)* in present and past tense needs to be categorized differently.

6.2.3.2.   Clause-Combining Patterns of Direct Reported Thought

The many names that previous literature has given to quotatives that introduce a direct quote show how difficult their syntactic classification is. Kaufmann (1976), who calls quotation in general *Redeerwähnung* (lit. 'speech mentioning'), calls the quotative *Redeeinleitung* (lit. 'speech introduction'). While first suggesting that direct reported speech is characterized by the reporting speaker repeating the original talk "in exactly the same form as S1 [=the original speaker, SF] has realized it in the speech situation 1 [=the original speech situation, SF]" (*ibid.*: 15, my translation), Kaufmann (1976) elaborates on this claim in a footnote:

> The formulation 'in exactly the same form as S1 has realized it in the speech situation 1' is a grammatical idealization. In practice, the deviation of the quote from the original speech R reaches from shortening to embellishment, from distorting the meaning of the quote to its conscious disfiguration. Often, a completely invented utterance is foisted upon the speaker S1 [=the original speaker, SF]. (*ibid.*: 189, my translation)

His problematization addresses one major problem when analyzing direct reported thought: recipients mostly do not—or, better, cannot—know to which degree a quotation has been changed regarding the original. Speakers can use this opaqueness as a resource in interaction: Quotes may be adapted according to the contingencies arising from the ongoing talk. Kaufmann's label of *Redeeinleitung* ('speech introduction') suggests a certain syntactic dependence between the quotative and the quote. It hints toward the quotative being somehow

---

[120]   Stein's (1995) results concerning *ich denke* remain, unfortunately, non-traceable regarding his analyses because only one out of 76 examples comprises *ich denke.* This is even more surprising as he mentions earlier that *ich denke* has become "recently apparently very popular" (*ibid.*: 145). A systematic and example-based analysis of *ich denke (mir)* as a quotative in its present and past forms remains a research desideratum. The present work sets out to fill this research gap.

[121]   *Ich finde* ('I find'), *ich glaube* ('I believe'), and *ich meine* ('I mean') are classified under the same category but with additional functions.

superordinated to the quote, without specifying how tight the connection between the two syntactic parts is.

Imo's (2007a) labelling of quotatives suggests his considering the quotative not necessarily as superordinate. Regarding *sagen* ('to say') and *meinen* ('mean'), he speaks of *Redeanführung* (lit. 'speech initiation'; *ibid*.: 78 and 128) or *Redesituierung* (lit. 'speech situating'; *ibid*.: 77f), with the latter term being taken from Brünner (1991). Imo (2007) states explicitly that "*Redeanführung* can […] be realized in the most different syntactic variants, be it in an introduced subordinate clause, a non-introduced subordinate clause, or a dependent main clause" (*ibid*.: 48, my translation).

The term *Redeanführung* is already used by Plank (1986). He separates the act of quoting into two parts, the so-called *Redeanführung* (lit. 'speech initiation) and the *wiedergegebene Rede* (lit. 'reported speech'; *ibid*.: 301). He points out that when establishing a continuum of directness of speech, this continuum must run parallel with the "more or less strong syntactic integration of speech initiation and reported speech" (*ibid*.). The continuum that he establishes is based on two factors: the sentence-like structure of the two syntactic parts and the syntactic subordination (*cf. ibid*.: 301ff). He concludes that "[t]he more different two sentential constituents are regarding the respective features [i.e., word order, introducing conjunction, mood, etc., SF], the more pronounced is the subordinate relation between them" (*ibid*.: 303). One aspect that Plank has not integrated in his reflection is that quotatives do not necessarily quote language only. Speakers can also quote body movements, facial expressions, or response cries, or perform these bodily activities parallel to language (König & Oloff 2018b; Zima & Weiß 2020).

Brünner (1991) explicitly states that what is reported does not only include verbal material but also "nonverbal communication, especially gesture and facial expressions" (*ibid*.: 3, my translation), which is why she refuses to speak of reported "utterances" but prefers a description as "reported communicative action" (*ibid*.). She calls the act of reporting metaphorically "window technique" (originally *Fenstertechnik*, *cf. ibid*.: 2), because "[i]n an ongoing communication, a window is opened and another communication is shown" (*ibid*., my translation). The goal of this window is, according to her, to enable the recipient(s) to "take part in the telling" (*ibid*.: 3, my translation) and being a witness of the reported event.[122]

Brünner (1991) establishes the concept of *Redesituierung* ('speech situating') with the following explanation: "The reported communicative action must be made clear as such to the current listener, and it must be situated in a way that the reported speaker, listener and situation become clear" (*ibid*.: 5, my translation). She establishes this term in distinction to *Redeeinleitung* (lit. 'speech introduction'), which is, for her, "too narrow and therefore misleading" (*ibid*.: 6, my translation). The term of *Redesituierung* would also take into account the different positions of the quotative regarding the quote, as well as the different prosodic or syntactic shapes that quotes can take (*cf. ibid*.).[123] Despite her argumentation for quotes being

---

[122]   She calls this function of reported speech "witness function" (Brünner 1991: 7, my translation).

[123]   A last possible conceptualization to describe maximal independence of the quote regarding the quotative and vice versa is when the quotative is a so-called *Floskel* (approximately 'routinized phrase') or a discourse marker. Using the German term *Floskel* for a quotatives places it near the sociolinguistic notion of a contextualization marker, which merely serves to situate speech or thought as such in the telling. What is important with such *Floskeln* is that they have indexical function, not a "presentational" (*cf.* Imo 2007: 89)

multimodally enacted, Brünner (1991) remains vague concerning the multimodal resources speakers can use. One of her, rather cognitive, conclusions is that speakers prefer prosodic resources to characterize others because "internal intonational characterizations offer, in contrast to verbal ones, the advantage that the speaker is less likely to commit and less likely to be attacked. Because intonational stereotypes are less conscious and more difficult to describe" (*ibid*.: 11, my translation). Brünner's assumption needs to be verified by looking at the recipients' responses to such characterizations through quotation. The present work considers not only isolated quotations but also the sequentiality of the reported thought. In my empirical analysis, I will show that multimodal resources may be, in contrast to Brünner's point of view, a very direct way of conveying a characterization, which may be more immediate than verbal resources. That certain quotatives have a higher tendency to introduce such embodied characterizations could be an argument for a need to see language and the body as a closely interwoven communicative system.

The last paragraphs have repeatedly touched upon the position of the quotative regarding the quote. In what follows, I will address this matter in more detail.

The Position of the Quotative

In his chapter on speech introduction, Imo (2007a) describes direct reported speech as a category that might necessitate a classification outside the matrix clause category. He suggests that 'fixed formulas' with or without matrix verb (i.e., *und ich so* 'and I like'; *cf.* Golato 2000) should be conceptualized as a separate category, not least because such quotatives are positionally flexible. Plank (1986) associates different positions of the quotative with different degrees of superordination. He states that "introducing quotatives that are in the foreground [are] more strongly superordinated than postposed or parenthetical quotatives that are inserted into the reported speech" (*ibid*.: 303).

Both syntactical reflections are important to understand the mechanisms of bipartite quoting-structures in general. The argument may necessitate, however, some reconsideration, simply because the data excerpts presented in previous literature suggests that speakers seldomly resort to parenthetically inserted or postposed quotatives.[124] Research working with interaction data suggests that, in everyday talk, quotatives are systematically placed *before* the quote, both in German and in English (Couper-Kuhlen 2007; Golato 2002b; Holt 1996, 2000; Klewitz & Couper-Kuhlen 1999; Selting 2010, 2017; Wessels 2019). As these studies all use examples with the verbs *sagen* ('to say') and *denken* ('to think') in German or English, it is likely that *ich dachte (mir) / ich habe (mir) gedacht* as well as *ich denke (mir)* as quotatives show similar

---

one (i.e., it indexes what comes next and does not project a potentially staged utterance). This is, for example, the case with *ich mein* ('I mean'), which is, according to Imo (2007), "on the way to being a discourse marker" (*ibid*.: 190, my translation, see also Günthner & Imo 2003).

[124]　But see Imo (2007), who discusses occurrences of postposed or parenthetically inserted occurrences of *sagen* ('to say'; *ibid*.: 87f). He explains, however, that these occurrences are "comparatively rare" (*ibid*.). Concerning *meinen* ('to mean'), which can also introduce direct reported speech, he explicitly says that "[i]n everyday talk, the construction with postposed speech initiation [= the quotative, SF] does not seem to play a role" (*ibid*.: 188).

characteristics.[125] This would also match previous syntactic findings by Schmidtke-Bode and Diessel (2017) on a large preference for postverbal complements[126]. In this position within the bipartite structure, the quotative would have the function of "presenting" the quote instead of merely "indexing" it (*cf.* Imo 2007: 87), which matches the grammatical features of a higher pragmatic relevance of the quote and the notion of social action formats projecting an upcoming action. Potential reasons for this non-flexibility of the quotative could be related to (inter-)actional and sequential issues that lie beyond the syntactic or pragmatic arguments presented so far.

First, quotatives allow speakers to frame subsequent talk as something reported from the past. The "early" marking allows speakers to frame their displayed stance "as an immediate emotional and therefore authentic reaction to the events of the time" (Deppermann & Reineke 2017: 343). By indicating a change in footing from the beginning, speakers recipient-design their quote: Recipients have the possibility to understand, from the start, that what is to be reported is not the product of the ongoing conversation but has been—at least purportedly—thought in the past. This is in line with Deppermann and Reineke's (2017) finding that "it seems as if the use of the reenacting, non-declarative format has more the goal to induce empathy and to involve the recipient emotionally […] than first and foremost display the respective stance itself" (*ibid.*: 343; see also Haakana 2007: 153).

Second, reported thought is (also) used to comment on the climax of a telling (Kärkkäinen 2012; Selting 2017) or to assess previous talk (Couper-Kuhlen 2007; Haakana 2007; Wessels 2019). At the same time, assessments have been evidenced to be "apt structures for doing terminal work on an extended sequence" (Goodwin 1986: 215), such as telling sequences. Combining these two findings, the complement-taking predicate that introduces the assessment (i.e., the quote) could be used by speakers to initiate a moving toward closure—either of the sequence or of one part of the telling (Wessels 2019). In this sense, according to Auer's (2005) differentiation of grammatical and interactional projection, the quotative functions as a projecting device on both levels. On the grammatical level, *ich dachte (mir)* / *ich habe (mir) gedacht* project their complement, while on an interactional level, it projects the speaker's moving toward closure.

The last section has aimed to show why, with *ich dachte (mir)* / *ich habe (mir) gedacht* introducing direct reported thought, positional issues may be, interactionally, of less importance than what previous literature working primarily with written or invented examples suggested (but see Imo 2007). In line with previous findings, the quotatives seem to have presentational character, thus tending to being placed before the quote.

---

[125]   The same holds for quotatives without complement-taking predicates. Studies on English *be like*, French *genre/comme* ('like') and German *und ich so* ('and I like') show the same characteristic of the quotatives being placed before the introduced quote (Dostie 2020; Golato 2000; Secova 2015; Streeck 2002).

[126]   It is important to repeat at this point that I excluded those occurrences of *denken* where it functions as an epistemic parenthetical (Thompson & Mulac 1991a). Such occurrences would probably exhibit a high degree of positional flexibility, similar to *ich glaube* ('I believe') because they do not function as quotatives, but hedges (Imo 2007b, 2011a, 2012b).

6.2.4. Combinatory Aspects of *Ich dachte (mir) / Ich habe (mir) gedacht* ('I (have) thought (to myself)')

The majority of occurrences of *ich dachte (mir)* or *ich habe (mir) gedacht* is syntactically bound to the preceding turn. By adding *und* ('and'), *da* ('there/in this moment') or *dann* ('then'), speakers establish a high coherence between two stretches of talk.[127] Golato (2002b), in her article on self-quotation, mentions *da* in a footnote: She argues that its use has the purpose of inducing the reverse word order *verb-subject* (instead of the default subject-verb order). This phenomenon of verb-subject inversion has also been documented for several other languages such as Russian, Modern Greek, or Hungarian (Golato 2002b: 67). Diachronic research has shown that *da* has been used since Old High German, but often *after* a verb in first position. Adolf (1944) argues that this constellation was used "when only a transition to some other phase of the narrative was intended" (*ibid*.: 74). Interestingly, in Old High German, *dô* ('then') had a past meaning only. This temporally more restricted meaning may have developed because of its recurrent use as a connector in narratives.[128] In Middle High German, the preposition of *dô* has replaced the verb-first syntax in tellings (together with the use of *ez* ('it')). In Lexer's dictionary of Middle High German, the lemma *dô* is described as follows: "adv. demonstr. da ('there'), damals ('at that time'), darauf ('then'); introducing a contrast: but still (often only indicating the progressivity of the narration) […]" (Lexer 1992: 32, my translation). The use of *da* in my data is often a mixture of *darauf* ('then') and the introduction of a contrast. Auer (1993) hypothesizes in a footnote that *da* may be on its way to an expletive item that is used as a dummy in the front-field,[129] like German expletive *es* ('it'; *ibid*.: 198). He argues that *da* being able to initiate a topic or a conversation just like *es* ('it') or verb-first syntax could be an argument that *da* ('there') creates cohesion without being anaphoric (which would exclude topic-initiation). This routinization may be due to its recurrent use in tellings.

There are also post-positioned discourse marker-like particles like *noch*, *schon* (focus particle), *halt* (modal particle; approx. 'just'), or *nur* (lit. 'only') by the means of which speakers can also establish coherence (Diewald & Fischer 1998; Thurmair 1989). Occurrences without any additional lexical item are rare: only 5 cases for *ich habe gedacht*, 3 for *ich habe mir gedacht*, 1 for *ich dachte* and none for *ich dachte (mir)*.

A recurrent configuration of the investigated quotatives is verb first-syntax. Adolf (1944) mentions that already in Old High German, instead of a temporal adverb like *da* or *dann* ('then') a verb in first position, "bearing a strong accent, was used likewise, especially when the verbal action was in itself significant" (*ibid*.: 74). This phenomenon of the so-called *Verbspitzenstellung* or verb first-syntax (V1) has also been described by Auer (1993) concerning its interactional function and by Günthner (2006) who discusses the role of V1 in narratives.

---

[127] As my quantitative analysis has shown, 21/37 occurrences of *ich dachte*, 7/13 occurrences of *ich dachte mir*, 20/53 occurrences of *ich habe gedacht*, and 20/34 occurrences of *ich habe mir gedacht* are syntactically attached to the preceding turn with *und* ('and'), *aber* ('but'), *da/dann* ('then'), etc.

[128] It has only been possible to exclude its reference to the future because other adverbs are used for this denotion. Any other use can thus not be excluded (*cf.* Adolf 1944: 74).

[129] The translations of the constituents of the so-called German *sentence brace* ('Satzklammer') are taken from Auer (1996: 62): *front-field* ('Vorfeld'), *left brace* ('linke Satzklammer'), *inner-field* ('Mittelfeld'), *right brace* ('rechte Satzklammer'), and *end-field* ('Nachfeld').

In my data, there are two formats that seem to be verb first-structures at first sight but have to be treated separately: verb first-syntax (which is again split into *eigentliche* and *uneigentliche Verbspitzenstellung*,[130] *cf.* Auer 1993), and the so-called *Koordinationsellipse* ('coordination ellipsis').[131] In both formats, the front-field is either not filled or it is filled with *und* ('and'), and the subject pronoun is either not realized or follows the verb. Consider the following three examples that illustrate the syntactic configuration of the past forms:

(1) coordination ellipsis: aufm rücken / on the back (KAFKU_04, 26min08)

```
01 HAN:  und      dann     lag        ich    da        halt   so
         and.KONJ then.ADV  lie-1SG.PRT PP.1SG there.ADV  PART   like.PART
         and then I lay there PART like

         gezwungenermaßen   auf      _m        RÜcken,=
         forced.ADV         on.PREP  the.DET   back
         under duress on the back

02       =dacht         so              ja TOLL,
         think-1SG.PRT  like.PART        QUOTE
         thought like yeah great
```

front-field	left brace	inner-field	right brace	end-field
Ø	dacht *thought*	so *like*		ja toll *yeah great*

In this pattern, the front field remains empty because the personal pronoun *ich* ('I') in line 01 can be said to fill that position. In paratactical coordination, a word like the personal pronoun can be omitted instead of repeated, which is why the pattern is called coordination ellipsis.

(2) actual verb first-syntax: in der halle / in the hall (KAFKU_04, 06min18)

```
01 JON:  [und er] dann,
          and he then
02       ja ne er will ja im SCHLAG bleiben in der halle,
         yeah no he wants to stay in the shot in the hall

         ((imitates the friend's talk silently, 2.0s))

03       ((laughs, 1.3))

04       °hh [>dacht          mir          so<]      nja GUT,
             think-1SG.PRT    PRO.REFL.1SG  like.PART QUOTE
             thought to myself like well okay
```

---

[130]   I translate these two terms with *actual* and *non-actual verb-first syntax*: "actual" V1 because it is the variety where the subject is not realized; "non-actual" because it is only an inversion of verb and subject.

[131]   V1 in tellings have also been called "dense constructions" (Günthner 2006) or *Vorfeldanalepse* (Zifonun *et al.* 1997: 632ff). The coordination ellipsis has also been described as *fragmentary conversational utterances* (Betten 1985: 274, my translation) or *pronoun zap* (Fries 1988). Both authors emphasize that the condition for such formats is that the non-uttered information can be deducted from context. Proske (2019) calls some cases of coordination ellipsis as pseudo-coordination, arguing that both parts of the coordinative clause "express a single event", with one of them hovering "between a literal and a bleached meaning" (*ibid.*: 116). As I focus more on the formal aspect of the structure—namely, that "we deal with two clauses coordinated by a conjunction meaning 'and' that have the same subject referent that is realized elliptically in the second clause" (Proske 2019: 116), I continue to call the structure coordination ellipsis.

front-field	left brace	inner-field	right brace	end-field
Ø	dacht *thought*	mir so *to myself like*		nja gut *well okay*

In this pattern, the subject, even though denoted by the verb (valency), is not realized. This is why the front field, usually filled with the subject pronoun, remains empty.

(3) unactual verb first-syntax: augen zu / eyes closed (KAFKU_01, 47min07)

```
01 RAB: hat          er        so        die   auge    zu,
        have-3SG.PRS he.PP.3P  like.DEIK DET   eyes.PL close.PREF
        he had like the eyes closed

02      dacht            ich        so,
        think-1SG.PST.PRT PP.1SG    like.PART
        I thought like

03      oh,
        QUOTE
```

front-field	left brace	inner-field	right brace	end-field
Ø	dacht *thought*	ich so *I like*		oh *oh*

In this last pattern, the verb has all arguments decoded by its valency. The empty slot in the front-field can be filled by an expletive form or adverb like *da* ('there') or *dann* ('then').

The syntactic configuration as V1-clauses has caused some reflections on agentivity and word order. Regarding the tempus of the quotative verb, Günthner (2006) names the narrative (or historical) present as one feature of V1-constructions in tellings. Indeed, all her examples are in narrative present, which enhances the function of agentivity, expressivity and liveliness within the telling (*cf. ibid.*: 104). *Ich dachte (mir)* and *ich habe (mir) gedacht*, however, are in past tense and yet they occur in V1-format. Consequently, in the present perfect tense, the auxiliary verb is in first position: *habe ich (mir) gedacht*. The argument of V1 focusing on agentivity, meaning an action-oriented advancing of the telling, would thus become non-valid—a problem that Auer (1993) raises in his article on verb-first syntax.

Auer (1993) states that with verbs in perfective, the usual tense used in tellings, it is not the reported action and thus the semantics of the verb in first position that becomes the focus of the telling but the finite auxiliary verb, *haben* ('to have'), which is only the "carrier of the grammatical information" (*ibid.*: 218, my translation).[132] What I find in my data is a phenomenon that seems to counterbalance this syntactical problem: When the auxiliary *haben* ('have') is in first position, it is mostly morpho-phonologically reduced to *ha* and attached to *ich* ('I'), thus forming one phonological "package" (*ha_ich*). This amalgamation might compensate the auditory separation of auxiliary and subject pronoun and give the past participle

---

[132] In addition to that, Auer (1993) states that *denken*, as a *verbum sentiendi*, does not bear agentivity in itself (*cf. ibid.*: 217) and usually does not advance a telling on the action level. This observation also holds for direct reported thought being introduced by *denken* because it allows speakers to comment on prior reported actions through a change in footing. In that sense, it does not seem to contribute to the progressivity of the telling itself but rather pauses the progressivity of the telling for intersubjectivity reasons.

gedacht ('thought') prominence back that it would not have, were the auxiliary and the personal pronoun pronounced distinctly.

So far, I have only presented prepositioned items co-occurring with *ich dachte (mir)* / *ich habe (mir) gedacht*. There is one post-positioned item recurrently co-occurring with the above-mentioned adverbs and particles: the adverb *so* ('like'). *So* is used as a quotative in spoken German—in combination with a personal pronoun only (Ehlich 1986; Golato 2002a; Streeck 2002), or accompanying a verb (for this observation concerning *ich denke*, see Wessels 2019). There are 21 cases of *ich dachte*, and 5 for *ich dachte mir*, *ich habe gedacht*, and *ich habe mir gedacht* respectively co-occurring with *so*. Often, the item creates coherence with what precedes and what follows by co-occurring with *und* ('and'; as in the routinized quotative *und ich so* ('and I like'))—a pattern that also exists with verbs like *denken*. *So* thereby functions as a deictic that foreshadows an upcoming bodily reenactment by referring to something visible that is about to be depicted by the speaker.[133] In quotatives, *so* seems to have the grammatical function of a deictic device that projects its specification, similar to French *là* in the quotative *j'étais là*.

The pattern *ich dachte / ich denke + so* has been previously described by Wessels (2019) in interactions between adolescents. Working with an interactional and Construction Grammar approach, she shows that speakers use direct reported thought introduced by *ich dachte/denke + so* for their "demonstration of the positioning-activity" (Wessels 2019: 66, my translation). One important point that she makes about the *ich dachte/denke + so* construction is that its sequential format of reacting to a previously described scene allows speakers to evaluate these scenes "still within the telling" (*ibid*.: 69, my translation). This positioning-technique is a possibility to "stage thoughts of the here-and-now of the reporting moment as 'purported' thoughts of the constructed, narrated world" (*ibid*.: 69–70, my translation). The subtleness of this possibility entails that speakers can imply preferred interpretations or recipient responses without the need of making them explicit" (*ibid*.: 70, my translation).

This overview has demonstrated that most of the lexical items and particles that co-occur with *ich dachte (mir)* / *ich habe (mir) gedacht* are (and have been for a long time) recurrent in tellings (e.g., *da*/*dann* 'there/then') or even quotation (e.g., *so* ('like')). I will show in the following analysis that, in line with previous findings on these (lexical items), *ich dachte(mir)* / *ich habe (mir) gedacht* systematically occur in patterns with additional lexical items when introducing direct reported thought.

6.2.5. Interactional Functions of *Ich dachte (mir)* ('I thought (to myself)')

In the following analyses, I will present four excerpts where self-quotation is used in tellings to report thought. An important characteristic of telling sequences is that the speaker holds the turn for a long stretch of talk, while the recipients only respond with verbal continuers (Goodwin 1986; Schegloff 1982), such as *hm*, or nods (Stivers 2008). A more extended response to a telling is typically made relevant after the speaker's closure of the turn. In my

---

[133]   When it is not used in combination with a quotative, *so* can also function as a demonstrative (Stukenbrock 2021), or as adverb, for instance in sentences like *so dachte ich, dass wir es machen* ('like this I thought that we would do it').

data, the affective stance-taking introduced by *ich dachte (mir)*, that is a follow-up to the climax of the story, is systematically followed by an affiliative response of the recipient(s). Following Haakana (2007), I argue that making available their personal affective stance allows the speaker to project the preferred response to the prior telling for the co-participant(s). Especially in tellings in which delicate situations, sometimes involving a present party, are described, it may not be easy for recipients to know which response is the preferred one. The self-quotation with *ich dachte (mir)* allows speakers to share the affective stance toward prior reported events, thereby inviting recipients to affiliate or align.

The first excerpt is a clear instance of this phenomenon. The direct reported thought occurs at the end of a telling during which two of the co-participants do not (yet) orient to the story as being "amusing." Three roommates, Sina, Nadja, and Marius have invited Fiona, a potential new roommate replacing Nadja, to a "roommate casting." The fourth member of the shared apartment participates remotely (computer on the table) but does not take part in the following discussion. Prior to this excerpt, Marius has complimented Nadja for her cooking skills. She confirms that—usually—she succeeds with her dishes but then delivers one counter example where she did not: a traditional dish of rolled, stuffed meat (roulades) that is not easy to prepare because the meat can easily get chewy.

Ex. 6.1: Wie gummi / like rubber (FOLK_E_00253, 20min35)

```
FIO: Fiona, NAD: Nadja, MAR: Marius, SIN: Sina

01 FIO:      [ja ] respe[kt dass du die also rou$LAden ist]
              yes  respect  that you I mean roulades is

02 NAD:                 [wie GUMmi °h und er   $°h       ]
                         like rubber and he
   fio                                        $gz at NAD-->1.15

03    (0.21)

04 FIO: S[U ] per schwe[r oder?   ]
         super   difficult no?

05 NAD: [öh]          [ja un er k]am von der >SCHUle heim<-
          uh              yes and he came back from school

06    !OAH! ich hätt voll bOck auf rouLAde-=
       oah  I   am totally hungry for roulades

07    =ich so ja +KOMM wir machen roulAden;=
       I'm like yeah let's make roulades
   nad           +gz down,then at her hands-->1.14

08    =un dann [s  ][aß er so d ]a-
       and then he sat there

09 FIO:         [ja,]
                 yes

10 SIN:             [((giggles))]

11 NAD: (.)% <<creaky>UH,>
   nad     %imitates MAR sitting in front of his plate,
           torso bent over the plate-->1.16

12    (0.56)

13 NAD: äh hh°
```

```
         uh
14 ???:  [((laughter))]

15 MAR:  [warn halt ¢wi]e$ [%+wie ∫GUM+m]i;¢∫#
          were PART like      like caoutchouc
   mar              ¢gz at FIO----------------->¢
                                 ∫PUOH gesture∫
   fio            -->$gz at MAR-->l.18
   fig                                          #fig1
```


fig1

```
16 NAD:                   [%GEH:     +t,]
                           (it's) okay
   nad               -->%l h to mouth as if to eat-->l.18
                            -->+gz at FIO-->l.18

17 MAR:  [%((laughs, 1.24))

18 NAD:  [%(.) a$(h)h°#aß er so+ DA$ °ge°,
           ((laughter)) he like sa(t) there
   nad  ->%hand down-->l.19
                          -->+gz in front of her-->l.19
   fio        -->$gz at NAD------------->$gz down-->l.22
   fig            #fig2
```


fig2

```
19       %gE+[$H::t,]
          (it's) okay
   nad  ->%slowly l h to mouth as if to eat-->l.21
         -->+gz left-->l.21

20 FIO:      [$gut, ]
              good
             $imitates MAR eating w r h,lifts eyebrows-->l.22

21 NAD:  %+hu[$hum::,+
   nad  ->%l h slowly down-->l.23
         +gz at FIO+gz down-->l.24

22 FIO:      [$(ja$ha,) °h
               yehes
   fio      -->$gz at NAD-->>
               -->$

23 NAD:  %und ich dAcht so    %HM,
```

```
          and I    thought like hm
   nad  ->%l h up,chest height%holds l h-->
24      (.) %<<high pitch>°ZUldigu:ng°>,+#
                             sorry
   nad  -->%covers mouth with l h-->l.28
                        -->+gz briefly l+gz down-->>
   fig                           #fig3
```


fig3

```
25 FIO: [((laughs))      ]
26 SIN: [((laughs))      ]
27 MAR: [((laughs))      ]
28 NAD: [°hab ich %(echt)°] verKACKT. (.) °h
         I've (really) screwed up
   nad       -->%l h down, manipulates pen-->>
```

This is a classic example of a storytelling where the speaker reports a concrete past event while using dramatizing elements like bodily and prosodic animation. The structure of the telling builds up to a punchline (lines 18 and 19) and is designed as amusing story.

At line 02, Nadja already states the result of her cooking roulades: Their texture was *wie gummi* ('like rubber'; i.e., chewy). This overlaps with Fiona, the potential new roommate, who expresses her respect (line 01) and asks whether roulades are not difficult to make, thus downgrading the severity of a failed dish. Nadja responds minimally to this question with *ja* ('yes'; line 05) and initiates her telling. What follows is a multi-unit turn from Nadja (lines 05 to 21), where she depicts Marius' reaction to eating the failed dish (lines 06 to 08). Her telling is only interrupted by Marius' confirming the rubber-like texture of the meat (line 15) and Fiona's co-animation (Cantarutti 2020) at line 20. While Nadja delivers the punchline of her storytelling in overlap (lines 11 and 14), Fiona shifts her attention to Marius. In the following turn, Nadja does not receive the affiliative response that has been made relevant after the climax of her designedly amusing story. Marius starts laughing at line 17, but the main recipient, Fiona, who is potentially the only one not knowing the story, has not yet reacted in a relevant way to Nadja's telling despite the latter's turn-final mobilizing gaze (Stivers & Rossano 2010) at line 16. Nadja repeats the punchline at lines 18 and 19 with laughing voice, reenacting again Marius' eating. This time, instead of gazing at the main recipient, Fiona, Nadja is gazing to her left as if to enact Marius' avoiding gaze while eating something that he does not like (line 19). After Fiona's ironic co-animating response (line 20), she gazes at Nadja who, after gazing quickly at Fiona as if to check her involvement into the telling, produces a follow-up with *und ich dachte so* ('and I thought like'; line 23).

Nadja's self-quotation of her purported thought while watching Marius eat allows her to share her affective stance toward the preceding story, in this case its "moral result" (Selting 2017): She was ashamed of the bad quality of the dish (line 24). What Nadja frames as reported thought is presented as an immediate response to Marius' verbal and bodily display of not liking the food. This format is similar to what Butterworth (2015) describes in her book on techniques of reported speech in interaction, in which she also describes occurrences of reported thought. She shows that there are sequences where direct reported speech and direct reported thought occur in tellings in a dialogic format: While direct reported speech is used to repeat a third party's speech or conduct, the self-quotation of thought of the teller allows them to "metacommunicatively frame the telling, insofar as they [the occurrences of direct reported thought, SF] contextualize the evaluative horizon to which the speaker constantly refers within the whole telling" (*ibid*.: 259). Such a pattern occurs in Excerpt 6.1, where Nadja's direct reported thought responds to Marius' hesitant eating and evaluation of the dish.

The moral result of being ashamed is displayed through bodily means: Nadja covers her mouth with her hand, gazes briefly left, then down and tucks her chin slightly to lower her head (line 24, fig3). Her gesture of covering her mouth develops fluently out of her imitating Marius' eating. After displaying her affective stance verbally and through bodily means with her direct reported thought, Nadja receives laughter from all participants (lines 25–27). Overlapping the common laughter, Nadja closes the sequence with one last assessing comment on her cooked dish while gazing down and manipulating the pen on the table, thus displaying her leaving the floor to her co-participants (line 28).

The use of *ich dachte* in this excerpt is in line with the function that Haakana (2007) describes for Finnish *minä/mä aattelin et* ('I thought that'): Reported thought can be used to "[guide] the recipient in evaluating the story-in-progress" (*ibid*.: 153). As Nadja orients to the story as amusing by laughing herself (lines 18ff), she displays that the recipients could do the same and evaluate the story like she did herself. The change of orientation is observable in Fiona's conduct. She did not orient to Nadja's telling as amusing at the beginning of the excerpt. Instead, she utters an account for why one can easily mess up roulades: they are very difficult to prepare. Fiona's orientation evolves during the telling, along with Nadja's "performance" of the story as amusing, culminating in her co-animation at line 20. The excerpt also shows the change of footing from Nadja first reenacting Marius while eating to reenacting her purported thought in the moment of watching him eat and commenting on it. Even though her responsive stance display is staged as being a spontaneous response within the telling to seeing Marius forcing himself to eat, the response relevance to the (verbal and embodied) affective stance-taking seems to reach into the ongoing conversation, as the recipients' response demonstrates.

Excerpt 6.2 shows a similar case of an emotionally charged storytelling where the speaker uses the direct reported thought to display her affective stance with *dachte ich so*. Compared to Excerpt 6.1, this excerpt illustrates that a speaker may use direct reported thought as climax of a storytelling and extend their turn when the self-quotation is not responded to with the projected relevant next—here, an affiliation (from both participants).

Jana, Lisa und Rabea are talking about sports. Before the beginning of the excerpt, Rabea and Lisa were exchanging about running. After a short lapse, Jana initiates a new, albeit related, topic: A common friend, Jule, regularly went to a gymnastics class of the university sports

(strength and cardio training to music) in summer despite the gymnasium always being packed with students. The specificity of this gymnastics course is that it is for free, which leads to a constant high number of participants running in circles in a small gymnasium with an animator in the middle—which is why everyone at university calls the class "mass fidgeting."

Ex. 6.2: massenzappeln / 'mass fidgeting' (KAFKU_01, 19min33)

```
RAB: Rabea, LIS: Lisa, JAN: Jana

01 JAN: ja also jetz auch diesen $SO¢Mmer-
           yes well PART also this summer
    rab                               $gz at JAN-->>
    lis                                ¢gz at JAN-->>

02      wenn +jUle immer gesagt hat dass sie ins
           when jule always said      that she is going to the
    jan       +gz at RAB-->l.04

        MASsenzappeln geht-=
        'mass fidgeting'

03      =dacht^ich so-=
           I thought like

04      =°+%JU+§le°;$%
            jule
    jan   ->+gz at LIS+gz at RAB-->l.05
             %lifts eyebrows%
    rab          $nods$

05      #sei doch nich <ver+%↑↑RÜ:#CKT,>%=
           don't be PART crazy
    jan                  -->+gz at LIS-->l.06
                          %lifts eyebrows%
    fig  #fig1                    #fig2
```


fig1   fig2

```
06 RAB: =oh ja,+=
           oh yes
    jan       -->+gz at RAB-->l.07

07 JAN: =>da gehn ja immer< %ULtra >vie%+le $hin<-=
            super many people PART always go there
    jan                              -->+gz at LIS-->l.09
                          %lifts eyebrows%
    rab                              $nods-->l.08

08      =in diese kleine $spOrthalle;
            in this small gymnasium
    rab              -->$

09      die ja [völl]ig mit +HOLZ ausgekleidet is,
           which is entirely covered with wood
    jan              -->+gz at RAB-->l.12

10 LIS:       [hm, ]
```

```
11        $(0.3)$%(0.9)
   rab  $1 nod$
   jan           %headshake-->l.12
12 JAN: da kriegsch ja nach ze%+hn minuten keine LUFT mehr;
        PART after ten minutes you can't breathe anymore
   jan                        -->%
                              -->+gz at LIS-->>
```

At lines 01 to 03, Jana describes how Jule told her about going to the gymnastics class regularly during summer. What is framed as responsive thought to that description is introduced by the social action format *dacht ich so* ('I thought like') and allows Jana to display her stance toward her friend's behavior: The quotative *dacht ich so* has verb-first syntax, which accelerates the telling (Günthner 2006) and stages Jana's thought as an immediate response to Jule's informing about going to the gym. The quotative is first followed by a vocative addressing Jule, which functions here as a means to "assign the role of the recipient to the co-participant [here, in the imagined situation, Jule, SF] in relation to the […] expressively marked subsequent action" (Günthner 2016: 432, my translation). Second, Jana utters the imperative *sei doch nicht verrückt* ('don't be crazy'; line 05), still addressed to Jule within the telling. Jana marks the "dialogue" between their friend's informing (line 02) and their reported thought (line 03) with two extreme-case formulations: *Verrückt* ('crazy') negatively assesses their friend, and *immer* ('always') is a "hyperbolic [form]" that characterizes the event as "typical and recurrent" (Günthner 2000: 132, my translation).

Jana's formatting of her negative stance and her subsequent bodily conduct demonstrate how much her turn is recipient-designed to receive an affiliative response. The direct reported thought format allows Jana to downgrade her negative assessment of the friend's actions: She embeds her implicit critique into a telling, which can be easily transformed into a "funny story" or taken back in case the expected affiliative recipient response is not delivered. Jana's subsequent account that allows her to "fade out" her reported thought (lines 07–09) also indicates her careful turn-design toward affiliation.

In contrast to the previous excerpt, the reported thought is not an afterthought of the climax but forms the climax itself which is also reflected in the increased use of bodily resources. Albeit short, the quote is dramatized through an eyebrow lift and a head poke, thus emphasizing Jana's judgement (fig1 and fig2). The assessing adjective is additionally stressed by a high pitch movement and a lengthening on the second syllable of *verrückt* ('crazy').

As the gist of the telling, the direct reported thought already strongly projects the recipients' response. The analysis of the participants' gaze conduct reveals that Jana heavily draws on gaze to orient to a possible next speaker, thus increasing the pressure on her co-participants to respond. Both interlocutors have oriented to Jana by gazing at her throughout the whole sequence while delivering affiliating nods (lines 04, 07, and 11) with the speaker's stance that has been made available through prosodic, linguistic, and embodied means[134] (Stivers 2008). Rabea displays her verbal affiliation with Jana's assessment (line 06) that she had built up toward with her prior nods. Note that, prior to Rabea's response, Jana has been gazing primarily at her (lines 02 to 05), which is a common practice to mobilize the recipient's response after

---

[134]   Rabea's first nod at line 04 starts only after Jana has already begun pronouncing with emphatic prosody the address term *Jule*.

assessments (Stivers & Rossano 2010: 317). At line 05, after the beginning of Rabea's nod "forecasting a likely affiliative stance at story completion" (Stivers 2008: 53), Jana shifts her gaze to Lisa which allows her to monitor Lisa's involvement into the telling and invite a response. After a short gaze back to Rabea, thus orienting to her verbal response (line 06), Jana shifts her gaze again to Lisa, thus addressing her as potential next speaker (Auer 2021), just after the embodied extreme case formulation *ultra viele* ('super many'; line 07).

Jana's gaze conduct shows how carefully she monitors each recipients' response to her telling. Probably because Lisa's response has not yet been obtained, Jana extends her turn and orients to Lisa as the potential next speaker, who finally responds at line 10. The subsequent pause (line 11) demonstrates, however, that Jana may wait for yet another response, which could indicate that the so far delivered responses have been insufficient. To extend her negative assessment, Jana produces a headshake and adds another account: One cannot breathe anymore in this gymnasium. Her efforts remain in vain as this last account is followed by an information seeking question by Lisa about the number of participants in this class (not in transcript). Excerpt 6.2 illustrated that when the expected response is not delivered, speakers use linguistic means (turn-extension) and embodied resources (recipient gaze) to pursue a response.

In the last excerpt, *und dann dacht ich* ('and then I thought') is also responsive to an iterative action that a non-present party performed. What differentiates the following excerpt from the preceding two is the object of the speaker's stance introduced by *dacht ich*. Instead of presenting the affective stance as having been taken toward a prior described action (in Excerpt 6.2 an indirect reported speech), the stance is here responsive to a reported first pair part from a third party. This configuration changes the contingencies of the ongoing conversation.

Jana and her two friends talk about backpains. As an example for backpain being related to sedentary activities, Jana names her former studies that had a very heavy workload:

Ex. 6.3: laber rhabarber / bla bla bla (KAFKU_01, 12min59)

```
JAN: Jana, LIS: Lisa, RAB: Rabea

01 JAN:   $$¢ich hatte (.) +%!KEI!ne% freie minUte mehr,
             I didn't have   a free minute anymore
    jan                      %lifts eyebrows%
                             +gz tw RAB-->l.03
    lis  >>$gz at JAN-->l.05
          >>$nods-->l.02
    rab    >>¢gz down at cookie-->l.05

02        ich hatte lErntage des hab ich$ noch %NIE erlebt,%=
          I had studying days I have never had that before
    jan                                  %headshake--->%
    lis                         -->$

((5 lines omitted where Jana explicates what the heavy workload entailed))

08        und da hat ich SO arge rÜckenprobleme-
          and there I had huge back problems

09        °h$+ und ¢dann bin ich paar mal $zur physiotheraPIE,
              and then  I went a few times to the physiotherapy
    jan  -->+gz at RAB-->l.10
    rab      -->¢gz at JAN-->l.17
    lis  ->$gz at cookie----------------->$gz at JAN-->l.24
```

```
10        +und dann ham die im+mer: (0.3) +bisschen geSCHUMpfen-
            and then they always         scolded me a bit
          ->+gz down----------->+gz at LIS--->+gz down-->l.12

11        <<high pitch>ha[ha,>+]
                        ((laughter))

12 LIS:                 [hihi+]&
                        ((laughter))
   jan                    -->+gz up-->l.14

13        [hihihihi.   ]
           ((laughter))

14 JAN: [>und gesagt<] ja wenn sie sich nich be+wEgen dann
        and said     well if you do not exercise then
   jan                                  -->+gz at LIS-->l.17

        kein [$WUN ]der,    $
        no wonder

15 LIS:      [$mh:;]
   lis       $shrugs r shoulder
               &head right->$nods-->l.17

16 RAB: mhmh-

17 JAN: °h¢ un:d dann dAcht ich $+%naJA: la¢ber rhaBARber,%
          and then I thought     well  bla bla bla
   jan                             %nods------------------>%
                             -->+gz tw RAB-->l.18
   lis                     -->$
   rab   ->¢looks for something under table ¢gz at JAN-->l.20

18      un_+dann: (0.3) >hab ich aber %ge+merkt<% man muss halt
        and then          I realized though    that one has to PART
   jan  -->+gz at LIS------------------->+gz at RAB-->
                                    %shrug w l shoulder%

        echt (.) ∫den $%rÜ+cken irgend%wie [stabi]%liSIERN,=
        really    somehow stabilize the back
   jan                  -->+gz at LIS-->l.20
                        %1 beat,grasping
                                gesture%turns hand %palm up-->l.20
   rab          ∫nods-->l.20
   lis                $nods-->l.22

19 RAB:                                   [mhmh,]

20 JAN: =wei+l @wenn man kei¢ne muskeln hat +%wo @solls denn auch
         because when one doesn't have muscles where should it PART
   jan    -->+gz at RAB-------------------->+,,,,,,,,,,,,,,,,,,
            @lifts shoulders,holds them up-------->@
                                    -->%,,,,,,,,,,,,,,,,,,
   rab                    -->¢

        HER%[komm;] +∫°hh
        come from
   jan  ,->%
        ,,,,,,,,,,-->+
   rab            -->∫

21 LIS:    [hm,  ]

22      (0.8)$
   lis  -->$

23 JAN:      [°ge $nau°.]
```

```
              yeah
24 LIS: ne ich [merk$ des] auch wenn ich halt irgendwie,
       no I    feel it     also when I PART somehow
   lis            -->$
```

Jana starts listing several points to illustrate what the heavy workload meant for her (lines 01 to 07). At line 08, Jana concludes her list by delivering its consequence: She had huge back problems, which she tried to solve by going to the doctor (lines 09 and 10). Subsequently, she reports the physiotherapists'[135] opinion and speech (lines 10 to 14), which she frames with *und dann ham die immer bisschen geschumpfen* ('and then they always scolded me a bit') as a rebuke from the doctors. The framing itself is accompanied by Jana lowering her gaze and head, thus embodiedly displaying shame (line 10), which is, however framed as non-serious through laughter (lines 13 and 14).

The physiotherapists' reported speech is projected by the social action format *gesagt* ('said') and staged through a change in voice quality imitating the doctor's voice (line 14). At this point, both interlocutors have linguistically and embodiedly affiliated or aligned with Jana's back pain: Lisa displays her empathy through a shrug, a serious facial expression, and a head tilt; Rabea delivers the backchannel *mhmh* (line 16).

Jana's reported thought allows her to display her affective stance toward the physiotherapists' opinion by reporting her (purportedly thought) response to the doctors' scolding. The expression *laber rhabarber*, literally translated "babble rhubarb" expressing something like "let them talk," is emphasized through bodily means by Jana's nodding. This formulaic expression and Jana's bodily conduct display her refusing attitude toward the doctor's implicit reproach that she does not exercise enough. Because Jana's negative stance is framed as a direct reported thought addressed, in her telling, to an underspecified group of "the physiotherapists," the pattern [*und dann dacht ich* + direct reported thought] can be analyzed to stage a speaker's negative assessment as part of the telling, without a change in footing. In this sense, "reported" primarily means that Jana presents the quote as if it had been thought at a specific point in time in the past. This technique allows Jana to turn a potentially uncomfortable situation that might put her in a bad light, into an amusing story.

At the same time, the direct reported thought prepares, in Excerpt 6.3, Jana's change of opinion, which constitutes the telling's punchline. In contrast to Excerpts 6.1 and 6.2, Jana's displayed stance does not seem to project the recipients' response. The missing response from both recipients, Lisa and Rabea, but also Jana's immediate continuation at line 18 illustrate this sequential configuration. At lines 18 and 20, Jana clarifies that she knows better now and utters her now revised stance, which is, now, responded to by the recipients with nods (line 18).

In this excerpt, the function of *dachte* comes close to what Jefferson (2004) describes for *at first I thought.* She analyzes *at first I thought* as a means to assert "*the in-principle correctness of the ordinary alternative*" (*ibid.*: 409, original emphasis):

---

[135]   Jana actually uses the plural when referring to the doctors working at the physiotherapy office: She uses the demonstrative *die* ('they') and the verb *haben* ('have') in third-person plural (line 10).

> Their plausibility resides in that they stand as innocuous, ordinary, Anybody's Alternatives to what turned out, on some particular occasion, to be the actuality, i.e., they assert what any one of us would, could, should make of such an event. (*ibid*.: 139)

What Jana purportedly thought "at first" (i.e., *laber rhabarber*) is, in contrast to Jefferson's (2004) examples, formatted as a direct reported thought that allows her to stage her response to the physiotherapists' opinion as a negative affective stance. In Jefferson's numerous examples, no change in voice quality, no response cries, let alone embodied conducts appear and the *at first I thought*-pattern is not used to take a stance toward a prior reported event[136] but merely to report prior thought. In her telling, Jana, however, takes a stance with her direct reported thought not only through her use of the derogatory term *laber rhabarber* and the lengthened particle *naja* preceding it (line 17), which display her resistance, but also through her addressing the doctor in a staged scene of interaction (Yule & Mathis 1992).

The analysis of Excerpt 6.3 has also shown that, when the direct reported thought is embedded into a reported dialogue, a recipient response may be less strongly projected as a relevant next. That the quote is staged as an immediate response to a first pair part from an absent third party seems to change the contingencies of the ongoing conversation, thus making the response from the recipient(s) not immediately relevant.[137] One additional reason for the recipients not to respond after the direct reported thought might be the semi-rising final intonation at line 17, which foreshadows the speaker's continuation of her turn.

Summary

In this chapter, I have shown that *ich dachte (mir)* allows speakers to share their affective stance toward previously described circumstances or reported speech from a non-present party. The pattern [*ich dachte (mir)* + direct reported thought] occurs toward the end of a telling sequence and can constitute a follow-up of a climax or the climax itself. The placement of the direct reported thought seems to make the recipients' subsequent response relevant (Excerpts 6.1), which is, if lacking or insufficient, pursued (Excerpt 6.2). Through bodily and prosodic means, speakers can not only stage their reported thought as part of the telling, they are also able to exploit these resources to mobilize the recipients' response, such as through a gaze at the recipient (Auer 2021).

Formatting their affective stance-taking as direct reported thought (with the social action format *ich dachte (mir)*) allows speakers to stage their stance still as part of the telling. By doing so, the contingencies arising from the affective stance-taking remain somewhat ambiguous. Even though the speaker's footing clearly shifts to the 'past self' for the direct reported thought, which should eliminate any response relevance in the ongoing conversation, a response still seems to be expected in most cases (see specifically Excerpt 6.2). The direct reported thought may therefore be a means to display an affective stance with lowered commitment by framing

---

[136] Jefferson mentions herself that she does not have exclamations in the excerpts (Jefferson 2004: 148), which she supposes to work similarly to immediacy markers such as "immediately" and the like. This lack of exclamations might be because most of her examples are from testimonies from newspapers, where writers may cut out such markers.

[137] This configuration may change when the co-participants are made figures of the reenacted scene, as Pfeiffer and Weiß (2022) have shown.

it as part of a past telling that speakers usually tell "to do something" (Schegloff 1997: 97), for instance to complain, to make laugh, to inform, etc. In that sense, the reported thought becomes part of a larger activity and less exposed as stance-taking.

If speakers simply want to take a stance with lowered commitment, why don't they simply use a hedged assessment? One possible explanation would be that an assessment does not allow for a dramatization and staging as direct reported thought does. As my multimodal analysis showed, speakers make ample use of facial expressions and what I called "reenacted gaze" while quoting their purported thoughts. Compared to a non-staged, simple assessment, a depicting direct reported thought allows speakers to invite their interlocutors to join the telling and to join their perspective. Ultimately, this invitation may also make a preferred response to the speaker's stance more likely.

Syntactically, there is no link between the quotative with *ich dachte (mir)*[138] and the quote. However, my analysis has shown that, on a pragmatic level, the quote becomes foregrounded while the quotative seems to become a presentational device (*cf.* Imo 2007) or a projector-construction that projects the quote while losing the syntactic status of a matrix clause (*cf.* Günthner 2011). Despite the lack of syntactic linkage, *ich dachte (mir)* still seems to have some projective force. Even if it is on its way to being a routinized device introducing direct reported thought, the quotative still grammatically projects "more to come," regulated through the principles of governance and constituency (see Section 2.2 for a detailed explanation). The complement-taking predicate still seems to regulate how many empty argument positions there are (one is filled by the subject pronoun; another can be filled by the quote). However, *ich dachte (mir)* cannot syntactically embed anything when it functions as a quotative of direct reported thought. My analysis has shown that it *must* introduce a complement clause, a non-lexical vocalization, and/or a body movement that allows for depiction and that forms the second part of the bipartite structure. In that sense, Auer's (2002) principle of constituency is still at play.

Some other features contribute to the perception of *ich dachte (mir)* routinizing into a presentational device: Excerpts 6.1 and 6.2 illustrated that there is a tendency of *ich dachte (mir)* co-occurring with other lexical items, especially post-positioned *so* ('like'). When used, *so* seems to contribute to the projective force of the quotative by making its upcoming specification expectable. At the same time, the quotative is often morpho-phonologically reduced, which has been evidenced to be another characteristic implying formulaicity. In that sense, *ich dachte (mir)* may be on its way to routinize into a speech-situating device in Brünner's (1991) sense,[139] where the quotative becomes syntactically maximally independent and serves as a mere framing item. However, because other, less routinized formats co-exist, I analyze this routinization of *ich dachte (mir)* as a quotative only as 'being underway'.

---

[138]   But the reported thought still fills the slot of an argument of the matrix clause (see Section 6.2.3).
[139]   See Section 6.2.3 for a detailed presentation of the concept.

6.2.6. Interactional Functions of *Ich habe (mir) gedacht* ('I have thought (to myself)')

In my data, there are two actions that *ich habe (mir) gedacht* carries out when introducing direct reported thought. Speakers can, as with *ich dachte (mir)*, display their affective stance toward prior actions or events. *Ich habe (mir) gedacht* is, however, also used to display an epistemic stance to display a reasoning. The latter action is recurrent with *ich habe (mir) gedacht* ($n = 28$) and less frequent with *ich dachte (mir)* ($n = 4$).[140] My analysis will show that these two actions come with divergent bodily conduct and are used in different sequential environments. While the affective stance-taking is not only expressed verbally but also prosodically and through bodily means, the epistemic stance-taking is less embodiedly animated and prosodically "flatter." Compared to the preterit, the present perfect format thus seems to have a more marked twofold use in everyday talk that is also distinct on a bodily level.

The possible tendency of an interactional distribution is interesting because past tense use in German is often said not to be distributed systematically, according to, for instance, aspect (Rödel 2007: 50ff), let alone actions. Generally, linguists attribute a variation according to the speaker's region of origin to the distribution of past tense: Speakers from Southern Germany prefer the analytic present perfect form over the synthetic preterit form for most verbs (Zeman 2010). There have also been claims about the preterit disappearing more and more in spoken language (Fischer 2018). These general findings may, however, need further investigation, especially since they have been recently revised regarding the German variety of Alemannic (Leonhard 2022). My analysis shows that in interaction, speakers seem to use primarily the analytic form to display their epistemic stance to display a reasoning. I am aware that the number of occurrences in my corpus is restricted and needs to be verified in a larger corpus. The clarity of my results, however, hints that at least a hypothesis about the distribution of tempus according to action formation with the quotatives *ich dachte (mir)* / *ich habe (mir) gedacht* as social action formats projecting a quote can be made.

Let me begin with the explanation of the grammatical features of the analytic past form, which may be necessary to understand specific changes in word order. In contrast to the synthetic past form, the analytic past form is in need of an auxiliary, *haben* ('have'). The present perfect form can be glossed as follows:

ich	habe	(mir)	gedacht
PP.1SG	have.1SG.AUX	PRO.REFL.1SG	think.PTCP

Because of its analytic character, the form bears some syntactic specificities depending on which element precedes it. When a lexical item, such as a temporal adverbial like *dann* ('then'), is placed before this form, verb and subject pronoun change position to keep the verb in second position: *dann habe ich gedacht* (lit. 'then have I thought'). This is not the case with prepositioned conjunctions such as *und* ('and') after which the word order *ich habe (mir)*

---

[140]   Of course, there are also other practices with *ich habe (mir) gedacht* to display an epistemic stance, which I excluded in the present work because they do not have the format of direct reported thought. For an overview over other practices to display an epistemic stance, see Deppermann and Reineke (2017) and Section 5.4.1.1.

gedacht remains unchanged. When the form is preceded by the combination of *und* and *dann*, the word order also changes: *und dann habe ich gedacht* (lit. 'and then have I thought'). The change in word order does not influence the interactional functioning of the pattern other than backlinking the self-quotation to preceding talk through the additional lexical items (or not). There is, however, a prosodic feature that is related to word orders where the auxiliary *habe* stands before the subject or/and in first position (see Section 6.3.4 for a detailed explanation). This holds for the above-mentioned case with preceding adverbials or conjunctions but also for verb-first syntax and coordination ellipsis. The morpho-phonological reduction of the bi-syllabic *haben* occurs regularly in my data with the auxiliary being reduced to *ha* instead of *hab(e)* ($n = 19$ of 46 total occurrences of inversed subject-auxiliary or verb first order).[141] This reduction leads to a condensation of the pattern which is less possible with the subject-verb word order. The pattern with subject-verb word order ($n = 31$)[142] has seven cases where *ich* ('I') is reduced to *ch*. As the personal pronoun is in first position, this elision does, however, not lead to a condensation but rather to a shortening of the whole pattern. In only three cases, there is a clitization of *ich* ('I') and *hab* ('have') to *ich_ab*.

Excerpt 6.4 illustrates both kinds of stance-taking that speakers can display with direct reported thought introduced by *ich habe (mir) gedacht*. The telling takes place during a Tupperware party with Daria as host. She is the consultant of Tupper presenting different products to her guests. In the following excerpt, she reacts to one of the participant's (Carolin's) affirmation that she never uses a certain Tupper that she has at home, a so-called 'ultra'. An ultra is a Tupper product made of plastic that can be used in the oven or microwave and replaces its heavier equivalents out of cast iron or ceramic. At line 01, Daria initiates a telling as a response to Carolin's utterance. *Eine* ('one') in line 01 refers to the product 'ultra'.

Ex. 6.4: im schrank / in the cupboard + ultra (FOLK_E_00329, 23min30)

```
DAR: Daria, LEN: Lena, CAR: Carolin

01 DAR: [doch, ich hatte AU]CH ma eine a[ls ich]
            yes   I also got      one once when I

02 LEN: [((chuckles))       ]

03 CAR:                                  [ja,   ]
                                          yes

04 DAR: gastgeberin worden war geSCHENKT be↑↑kommen,
        had become a host        offered

05      °h und die stAnd und stand und stAnd bei mir im
           and it  stayed and stayed and stayed here in my

        SCHRANK.
        cupboard

06      $(0.26)
```

---

[141] This is a higher number compared to what Lanwer (2015) found, who had only one occurrence of reduced *haben* to *ha*, probably because he looked at first-person plural, not first-person singular. A reduction from *haben* to *ha* is a bigger step than from *habe* to *hab*. He observed, however, that this highly reduced form co-occurs with higher speed. This observation matches my results. He explicitly states that *ha* is part of these *Allegroformen* ('allegroforms') that exist in dialectal varieties of German (*ibid.*: 183).

[142] The two numbers for the different word order patterns do not add up to the total number of 87 occurrences because there are additional, more complex structures, such as those with *wo* ('where').

```
   car   $1 slow nod-->l.07
07 DAR: °h und§ dann bin ich be↑rAterin geWORden-
              and then I became a consultant
   car        -->§

08      und !UM! mich rum (.) $!AL!le bera%ter;
        and around me           all consultants
   car                      $head turned tw DAR-->>
                                    %....

09      <<glottalized>O%:H>+      %!UL!tra §ultra+%UL§tra;=%
                        oh      (3 times name of the tupperware)
   dar  ..............%elbows opened%2circling gestures%,,,,,,,%
                       to both sides    w hands
                          +gz at car------------>+gz down-->l.10
   car                              $nods----->§

10 =>da hAb ich#@ immer gedach< %ich wEIß% nicht# was
        PART I always thought   I don't know what that's
                                %shrug-->%
                  @facial expression of negative stance-->l.11

   fig
```


fig1

#fig2
fig2

```
        die HA+ben; §(.)
        all about
   dar      -->+gz at car-->l.11
   car              $nods-->l.11

11      %#steht im SCHRANK,+%@§
          stays in my cupboard
   dar                      -->+gz at packaging that she cuts open->l.17
   car  %shrug,2nods--------->%
                          -->@
                          -->§

   fig   #fig3
```


fig3

```
12 ALL: ((chuckle))

13 CAR: [#mhm,   ]

14 DAR: [#°hh #h] >un_dann hab ich mir gedacht< oKAY; (.)
                  and then I thought to myself okay
```

```
fig     #fig4#fig5
```


fig4   fig5

```
15      dann nImmst_se vielleicht DOCH ma;
        then maybe you use it PART once

16      (0.64)

17 DAR: °h und jetz %HAB ich% nix +%anderes% mehr;+
        and now I don't have anything else anymore
   dar                     -->+gz at car------->+gz tw cupboard
                                           behind her-->l.18
               %shrug-->%       %headshake%

18      %jetzt hab ich den: (.) dIcken sch%+rank (.)
        now I have this        huge cupboard
   dar                             -->+gz at car-->>
        %points tw cupboard behind her------->%

        voll mit ULtras.
        filled with (name of the tupperware)
```

From lines 01 to 05, Daria describes a situation in response to Carolin's prior informing that she never uses her ultra (not in transcript). The initial *doch* (line 01)—here, a response particle that signifies adversativity—already projects the upcoming contrast in Daria's telling (lines 07 to 18). She first describes that when she became a consultant, all consultants in her environment were enthusiastic about ultras (lines 07–09). Prosodic mimicking (glottalized voice) and exaggerated arm gestures allow Daria to depict their enthusiasm (line 09).

Subsequently, attached to the preceding intonation unit with latching, Daria displays her past stance toward the consultants' enthusiasm through direct reported thought. Introduced with *da hab ich immer gedacht* ('then I have always thought'), which is delivered with high speed, she quotes her responsive negative stance: she does not understand their enthusiasm, she does not use her ultra (lines 10–11). Daria's negative stance is also displayed through her embodied conduct. Her facial expression manifests incomprehension (fig1, fig2, and fig3) and she shrugs twice (line 10 on *ich weiß* ('I know'), and line 11), thus displaying her disaffiliation. The direct reported thought-format enhances intersubjectivity: It allows Daria to stage her stance for her co-participants to re-experience the situation as Daria herself. By doing so, she may increase the possibility to receive her recipients' affiliative response. Indeed, all co-participants respond with chuckles (line 12), thus orienting to Daria's telling as amusing. Note that Carolin, whose utterance is responded to by Daria's telling, displays more attentiveness throughout the telling than the other co-participants: She nods several times (lines 06 and 07, 09, 10 and 11) and delivers a verbal continuer (line 13) in overlap with Daria's inbreath projecting a continuation.

During Daria's long inbreath, her facial expression displays early on her upcoming change of attitude (fig4 and fig5; Kaukomaa *et al.* 2015). She thereby "[stretches] the temporal boundaries of an action" (Ruusuvuori & Peräkylä 2009: 393)—here, of her epistemic stance. Her direct reported thought introduced by *und dann hab ich mir gedacht* ('and then I have thought to myself') displays her emergent reasoning about finally using the ultra. By doing so,

she gives her recipients the possibility to participate, simultaneously so to say, at her upcoming thinking process in the moment of the described situation. With this second direct reported thought, the speaker shifts from an affectively loaded and depictive telling to an inner monologue that is less staged. Daria first reports what Couper-Kuhlen (2021: 43ff) calls *consequential okay* (line 14), that means an okay "that has some consequences for the speaker's agenda or future behavior" (*ibid*.: 43). It is attached to the quotative without pause and informs the co-participants about her new epistemic state—namely, that she finally started using the Tupperware (line 15), even though reluctantly. After a micro pause, she continues her reported thought (line 15), following the same intonation contour with a semi-falling final intonation. In contrast to the first direct reported thought with *ich habe gedacht,* this second occurrence (lines 14 and 15) is animated neither through bodily nor prosodic means. Except for the prepositioned facial expression during the inhale, Daria produces two prosodically "flat" intonation units without high pitch movements or changes in voice quality.

The direct reported thought allows Daria to publicly display a rationalization of her past decision to finally use the Tupperware despite her initial non-enthusiastic opinion about it. Ultimately, the direct reported thought also contains the final argument in Daria's activity of convincing her co-participants, especially Carolin, to buy an ultra. After a pause during which none of the recipients responds (line 16), Daria continues her turn by adding that she now has many ultras, thus implying that she is now convinced by the product (lines 17–18)—in contrast to Carolin, at whom she has been gazing most of the time. Carolin, however, never explicitly responds to Daria's telling. Instead, one of the other participants asks an information-seeking question (not in transcript).

The two occurrences of *ich habe (mir) gedacht* at lines 10 and 14 demonstrate the two most recurrent uses of direct reported thought in my data. It may be used to take an affective stance, in which case the quote itself is embodiedly reenacted, or it can make publicly available a reasoning, which does not occur with ample bodily conduct nor prosodic animation, probably because of it being part of a less emotional and more argumentative sequence. Despite the different actions that the speaker accomplishes with the direct reported thought format, both occurrences are a possibility for Daria to embed these actions into the dramaturgy of a telling. Making her affective and epistemic stance-taking part of a telling allows her to condense her— probably longer process of changing her opinion—into a short sequence, where the staged direct reported thought lets the recipients co-experience Daria's purported thought "as it emerged." By doing so, the speaker may render the overarching convincing-activity more successful, which is, in this excerpt, not the case.

The next excerpt contains, again, two occurrences of *ich habe gedacht*. The first one is a clear instance of affective stance-taking through direct reported thought. The second one is a less clear-cut case, which can be situated between the two actions of taking an affective stance and epistemic stance to display a reasoning.

In this "roommate casting," Jonas, sitting on the short end of the table, is the invited candidate. He "applies" for Nadja's room, who is sitting to the left of the computer. Prior to this excerpt, Marius has asked Jonas whether he will visit other apartments or rooms on that same day. It is known among all participants that available and especially affordable rooms are rare in the student town where they live. Jonas responds that he has visited only one room,

without knowing that it was in a fraternity,[143] and starts telling the story of visiting the fraternity in the following stretch of talk:

Ex. 6.5: günstig in der Altstadt / cheap in the historical center (FOLK_E_00252, 16min06)

```
JON: Jonas, NAD: Nadja, MAR: Marius, SIN: Sina

01 JON: in einer wohnung war ich die hab ich im bei we ge
        I was in one appartment that I  found in at

        gesucht geFUNden,
        (name of the online platform for shared appartments)

02 MAR: hmHM- $(.)
                $gz at JON-->l.08

03 JON: in der ALTstadt,
        in the historical center

04      °h (0.35) super↑GÜNstig# und so-=
                super cheap and all
   fig                          #fig1
```
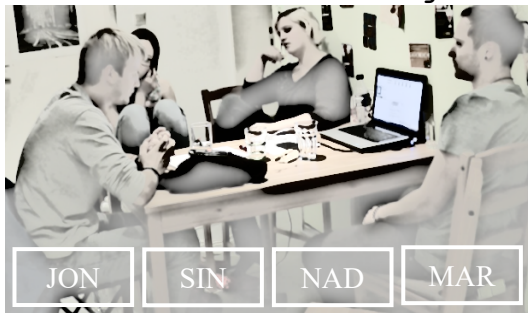
fig1

```
05      =>i°ch° hab gedach< °oh° GEIL #+ey,=
        I thought oh awesome man
   jon                              +gz at MAR-->l.14
   fig                          #fig2
```

fig2

```
06      =weil $da stand nix dabei von wegen verBINdung; (.)
        because it wasn't indicated something like fraternity
   mar       $smiles-->l.08

07      [$norMAL schreibst du_s]
         usually you write it

08 MAR: [$da isch irgendwas    ] FAUL;$
         there is something fishy
```

---

[143]   Fraternities are frequent in German cities with old universities. The rooms that they rent are usually cheap. Living in a fraternity, however, also means to be part of many rituals, often involving a lot of alcohol and dangerous tests of courage. Also, many fraternities are known to be politically situated at the far right and have conservative values, which is why living in a fraternity is generally frowned upon, as the excerpt also illustrates.

```
                                    -->$gz at NAD-->l.11

09        ja[ja;  ]
          yes yes

10 JON:   [normA]l schreiben_s
           usually they always indicate

11        immer da$[ZU,# >na] ha_isch gedach[t< egAL-#      ]
           it              then I thought     doesn't matter
   mar        -->$gz at JON-->l.12
   fig          #fig3                          #fig4
```
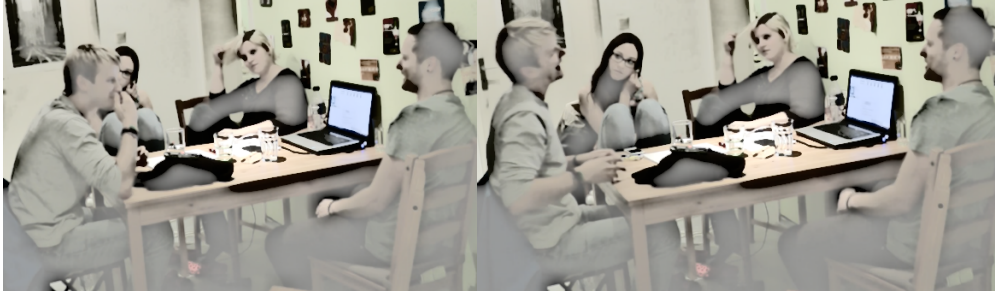


```
          fig3                          fig4
```

```
12 MAR:          [ja.    ]              [ah wenn_s es sch]on
                  yes                    ah when it it PART

          GÜNstig is in der $altstadt is
          cheap in the historical center
                         -->$

          [IRgendwas     fa+ul.        ]
          something is fishy

13 JON: [<<laughing>ja,> +°h           ]
                   yes

14 SIN: [es is ich       +wollt grad S]Agen;
        it is I just wanted to say
   jon                  -->+gz at SIN-->>

15        alles unter dreihundfünfzig euro
          everything under threehundred fifty euros

          [in der altstadt] [is ] irgndwie] KOmisch.
          in the historical center is somehow strange

16 JON: [jaja;           ]
        yes yes

17 MAR:                 [ja-]
                         yes
```

At line 01, Jonas delivers a response to the prior question: He describes that he visited one apartment after seeing the ad for a room on a known website (WG-gesucht, lit. 'shared apartment wanted') where available rooms are posted. After a continuer from Marius, Jonas names two features which led him to the decision of visiting the apartment (lines 03 and 04). At line 05, attached to the prior account with latching, Jonas reports his response to these fortunate circumstances: He evaluates the situation as strongly positive with the assembly of response cries *oh geil ey* ('oh awesome man'). His positive affective stance is formatted as a reenacted self-quotation, framed as such with accelerated *ich hab gedacht* ('I have thought'). *Oh geil ey* is not a response to a present party in the telling, like in Excerpt 6.5, but a response

to the circumstances described at line 04. During his direct reported thought, Jonas leans back with his torso (see change in posture between fig1 and fig2) and strongly stresses the evaluative adjective *geil* ('awesome'). While doing so, he gazes at Marius, who asked the question in the first place, and continues to do so until line 14, when Sina reacts to the thus far described scene.

At lines 06 and 07, Jonas delivers another account for his responding to the (admittedly suspicious looking) ad: *da stand nix dabei von wegen verbindung* ('it wasn't indicated something like fraternity'; line 06) and *normal schreiben_s immer dazu* ('usually they indicate it'; lines 10 and 11). This account displays his anticipating what the others already seem to know: The room was, in fact, in a fraternity. Foreshadowed by his smile in line 07, just after Jonas' initiation of his account with *weil* ('because'), Marius utters the obvious: If these circumstances occur, a cheap room in the historical center, something is wrong (line 08).

With *da hab ich gedacht* ('there have I thought'), Jonas produces a follow-up to his account, which is, again, formatted as a direct reported thought. This time, however, it allows the speaker to share the reasons he had to believe that the apartment was, in fact, not owned by a fraternity. The *egal* ('doesn't matter / never mind') allows Jonas to display, in an economic way, that he seemed to have dismissed upcoming doubts about the ad. During his direct reported thought, Jonas repeats his bodily conduct from line 05: He leans back with his torso (see change in posture between fig3 and fig4) while still gazing at Marius. Jonas' bodily conduct and his turn-final continuative intonation at line 11 could indicate that his telling is not closed at this moment but that he is still involved in the narrative preparation of the punchline—namely, the actual visit at the fraternity. However, at line 12, Marius self-selects, in overlap, and repeats that there is always reason to be skeptical as soon as a room is cheap *and* in the historical center. This is followed by another turn by Sina, who affiliates with Marius (lines 13–15). Jonas' *jaja* ('yesyes') at line 16 is another indication that his telling is not finished yet and that he will move on with his telling after this short side-sequence (not in transcript).

This second occurrence demonstrates that there are fewer clear-cut cases where a taken decision or a reasoning allow speakers to account for prior actions *while* taking an affective stance, as Jonas does with his emphatic *egal* (line 11). The co-participants respond to Jonas' emergent reasoning, formatted as direct reported thought, in a disaffiliative way: they dismiss his accounts and therefore his visit as unreasonable by delivering themselves "evident" reasons for why Jonas should have been skeptical. The recipients' response to Jonas' account (instead of the affective stance) can be understood as proof that the public display of a reasoning or of a decision is foregrounded in this telling.

Excerpt 6.6 again contains two instances of *ich habe (mir) gedacht*. However, this time, the first instance is part of a reasoning sequence—here, an explanation—which then veers into an affectively loaded depicting scene that allows the speaker to close her explanation on an amusing note. This excerpt also clearly demonstrates the difference in bodily conduct between the two uses of [*ich habe (mir) gedacht* + direct reported thought].

The four participants play a game called *Dixit*. Each one has cards with pictures on her hand. One participant chooses a word or sentence that she thinks matches well one of her cards (for example a picture of a cake with candles vs. the word 'birthday'). All the other participants then must choose one of their own cards that they think matches the chosen word or sentence and give one of their cards (face down) to the participant who chose the word. In a second step, all

chosen cards are put on the table (face up) and all participants must guess which card is the "original" based on which the word or sentence has been chosen.

In this excerpt, Gisela has said *im walde steht ein reh* ('a deer stands in the forest'). While the other participants choose their cards, she already announces that she had made a mistake. As Gisela explains, after receiving all the other cards, she thought of *reh* ('deer') playing a role in chess (which is apparently represented on her chosen card), which is not the case. She mixed up chess with skat, a card game. Her confusion leads to the others not choosing her card as the original one and she does not get any points. The excerpt begins with Ramona producing a change of state token, thus displaying her sudden understanding of the confusion after Gisela's explanation.

Ex. 6.6: zu leicht / too easy + scheiße / shit (FOLK_E_00357, 22min34)

```
CAR: Carola, RAM: Ramona, GIS: Gisela, VER: Vera

01 RAM: ach[↑SO::::;              ]
            oh
02 GIS:     [°h und die GANze  ]
              and the whole

        [zEIt wa[r i sch]o auf REH ne,  ]
        time I was already on deer (tag)

03 CAR: [°ah::-°          ]
          oh

04 RAM:          [ob     ]WOHL, (.) jaja;]
                  but            yesyes

05      [ja$ja.         ]
         yesyes
   car     $gz at GIS-->l.14

06 VER: [((chuckles))]

07 GIS: $MA[TT hätt i] [+sOng miassen +ne,     ]
         mate I should have said (tag)
   car  $nods-->l.14

08 CAR:    [(haj,)   ]
            (yes)

09 RAM:                  [+obWOHL des is +n biss]che[n °sch°]
                          although this  is a bit        di
   gis                   +gz at CAR------>+gz at RAM-->l.14

10 CAR:                                          [ja,   ]
                                                  yes

11 GIS:                                          [>aber<]
                                                   but

12 RAM: jaJA;
        yesyes

13      [n bisschen + SCHWER.                     ]
         a bit       difficult

14 GIS: [<<creaky>e:+h> >>hob_i_mer_$$denkt<< des is] zu ↓LEICHT;
                      I have thought to myself this is too easy
   gis          -->+gz at game on the table-->l.26
   car                      -->$
                             -->$gz at game on the table-->l.26
```

```
15        wenn ich sag [MATT,]
          if I say    mate    that

((8 lines omitted where co-participants evaluate 'mate' as too easy))

24 GIS:                    [aber H°e °h hi]
                           but ((laughter))

25     ((chuckles)) %wia_s gl%Eng is dann is mer des %KUMma,       %
                    once I put it ((the card)) then I realized it
   gis              %........%points tw cards,retracts r h
                          lifts l h------------->%both hands
                                           fall on table%

26     °HH >>°na hob_i_[denkt°<<$ %ʃAU=   ]sch (.) sch%+ʃ# sch
          then I thought          oh   sh    sh sh
   gis                            %...................%both h
                                      lifted,,
                                  -->+gz&head
                                     right->l.28
   ver                          ʃ...................ʃgz&head tw
                                     gis-->l.30
   car                  -->$gz at GIS-->l.28
   fig                                              #fig1
```


fig1

```
27 VER:                   [((laughter))        ]

28 GIS: [°SCH%+(h)EIße,°%    ]$ [ʃ((laughs, 4.16s))
          shit
   gis  ,,,,%rebound
           of both h %
        -->+gz&head to game on the table-->>
   car                  -->$gz down,bends over-->l.30

29 VER: [((laughter))     ]

30 CAR:                        [ʃhihi$
   car                          -->$
   ver                     -->ʃ

31 VER: °h okEH,
        okay

32     [wir zähln] TROTZdem die punkte-
        we count the points though

33 RAM: [okeh-     ]
```

From line 02 to line 07, Gisela finishes her explanation by saying what her reasoning has been—she thought of 'deer' (line 02)—and what she should have said: mate (line 07). Ramona's

disaligning response at lines 09 and 12 is not responded to by Gisela, who extends her turn by dismissing the word that she should have said, *matt* ('mate'), as too easy (line 14).

This account, which is at the same time an evaluation of the word mate, is formatted as direct reported thought introduced by *hob i mer denkt* ('have I thought to myself'). The quotative is pronounced with high speed. Its dialectal shape[144] (*hob = habe* 'have', *i = ich* 'I', *denkt = gedacht* 'thought') allows for a morpho-phonological condensation into one chunk where single words are not individually perceivable anymore. Additionally, the verb-first syntax of the quotative, where the subject pronoun is placed after the verb, makes the chunk a "dense structure" (Günthner 2006), which is recurrent in narratives, thus accelerating the narration. All these features make direct reported thought an efficient means for Gisela not only to account for her past actions but also to dismiss them as *zu leicht* ('too easy') from the perspective of the ongoing conversation. By making her past decision publicly available through direct reported thought, the speaker enhances intersubjectivity. Gisela allows her recipients to participate at the rationalization of her ultimate choice as if they were part of its emergence in real-time. Compared to the prior two excerpts, this sequence is not a classical storytelling of a past event with bodily and prosodic animations but rather makes available Gisela's reasoning behind her decision. Instead of simply informing her recipients, Gisela stages her reasoning as an assembly of logical steps from the past that account for its outcome and are not up to changes anymore.

Gisela then extends her turn with the explanation of what would have been too easy: if she had said mate (line 15). After a short side sequence where the participants evaluate whether "mate" would have actually been too easy (lines 16–23, not in transcript), Gisela continues her telling (line 24) with a disruptive *aber* ('but') and a laughter which displays her orientation to the upcoming part of her telling as "amusing." She describes how she came to realize that she mixed two things up: once she had put her card on the table, she realized her confusion (line 25). Note that she uses a dialectal formulaic pattern to express her sudden realization: *dann is ma des kumma* (lit. 'then it came to me'). While doing so, she points toward the card in question and lifts her other hand only to let them both fall on the table during *kumma* ('came')—a body movement that has been described in gestures studies as being pragmatically "[u]sed to dismiss topics of talk" (Bressem & Müller 2014b: 1584)—here, probably Gisela's mistake. This hand movement can also be interpreted as the foreshadowing of her upcoming negative affective stance, which is subsequently verbalized at line 26.

With *na hob i denkt* ('then I have thought'), Gisela introduces her direct reported thought, which is accompanied by ample hand gestures. She starts lifting her hands after the quotative (fig1), holds them up during her hesitation of three *sch* (line 26), and then pushes them down quickly with a small rebound just before the focus accent on the finally pronounced *SCHEIße* ('shit'; line 28). While doing so, Gisela turns her head and gazes right, thus averting the co-participants' gaze. These movements are visible to Carola and Vera, who shift their gaze toward Gisela at line 26. With her direct reported thought, Gisela stages her anger at herself while turning her prior reasoning into an "amusing story."

Whereas the explanation between lines 02 and 20 remains without signs of an amusing story, Gisela's description of her realization (line 25) is preceded by chuckles and produced with

---

[144]   The speaker has a Bavarian accent.

laughing voice. After her direct reported thought (lines 26 and 28), she bursts into laughter and continues laughing for about 4 seconds (line 28). Vera and Carola affiliate with Gisela by laughing (lines 27 and 30), thus displaying their orientation to the story as amusing. Their affiliative responses demonstrate that direct reported thought may be an efficient way—compared to a simple assessment—for speakers to stage their affective stance with their whole body, thereby making available to their recipients which aspect of the telling they should affiliate with. This "technique" may be necessary in Excerpt 6.6 because Gisela made a mistake during the game and topicalizes it—a delicate matter that recipients may find complicated to respond to. The direct reported thought allows Gisela to point out which kind of stance she takes herself (i.e., she orients to her mistake as funny), thus facilitating affiliation. Vera's initiating a new topic, counting the points, indicates that the telling sequence is now closed.

Summary

The three excerpts, each of which contained two occurrences of *ich habe (mir) gedacht*, illustrated the two actions that speakers carry out with the format [*ich habe (mir) gedacht* + direct reported thought]. I have shown that speakers use it to (1) take an affective stance in storytellings, in which case the quote is, in my data, mostly coupled with an embodied or prosodic reenactment or (2) to display an epistemic stance in reasoning sequences to display a post-hoc rationalization. All three excerpts contrasted the two actions that speakers accomplish with direct reported thought on a linguistic and embodied level. My analysis has shown that bodily resources tend to be distributed according to the actions that speakers carry out. While a display of the speaker's affective stance is accompanied by ample bodily conduct, speakers deploy less bodily resources during the display of an epistemic stance.

These qualitative observations are corroborated by quantitative results. Facial expressions and reenacted gaze are less frequent with the analytic past form: Whereas facial expressions occur in 56.0% of quotes with *ich dachte (mir)*, only 25.3% of *ich habe (mir) gedacht* do. Reenacted gaze occurs in 60.0% of all cases of *ich dachte (mir)* as a quotative and only in 41.4% of *ich habe (mir) gedacht*. This result is one further argument in favor of the multi-functionality of *ich habe (mir) gedacht* as a quotative, which is also reflected in the speakers' multimodal conduct accompanying the affective or epistemic stance-taking.

Despite a lack of syntactic linkage, there is still a pragmatic link between the quotative and the quote. As with *ich dachte (mir)*, grammatical projection still plays a role despite the tendency of *ich habe (mir) gedacht* routinizing into a social action format: Not anything can follow *ich habe (mir) gedacht* when functioning as a quotative. It seems to primarily introduce the second part of the bipartite structure in the format of a complement clause, a verbless structure, non-lexical vocalizations, or body movements. Prepositional complements do not seem to be introduced by *ich habe (mir) gedacht* functioning as a quotative.

The following section will be dedicated explicitly to the differences between *ich dachte (mir)* and *ich habe (mir) gedacht* regarding tense: I suggest that a distribution according to action formation seems to be a possible tendency to consider.

6.2.7. Reflections on the Distribution of Past Tense: Which Relation between Tense, Action Formation, and Aspect?

After my analysis of *ich dachte (mir)* and *ich habe (mir) gedacht* as quotatives, I would like to analyze one excerpt where both forms are used within the same sequence. As for now, I have mostly addressed the similarities in use between the two past forms. I have also shown that when introducing direct reported thought with *ich habe (mir) gedacht*, speakers can carry out two distinct actions, which is less often the case for *ich dachte (mir)*. What I have not addressed so far is the question whether the two tenses, preterit and present perfect, may tend to be distributed according to the actions that they allow speakers to accomplish. Excerpt 6.7 shall illustrate this hypothesis.

Fiona is at a roommate casting. Prior to this excerpt, the three students living in the shared apartment asked Fiona about her studies and it turns out that she has studied art history for two semesters before studying law, which she accounts for starting at line 02.

Ex. 6.7: jurastudium / law studies (FOLK_E_00253, 23min01)

```
FIO: Fiona, SIN: Sina, NAD: Nadja, MAR: Marius

01 MAR: uh $o[+kay;]
        uh okay
   mar    $gz at FIO-->l.20

02 FIO:     [+joa ha]t [mich irgendwie int]ress
            yeah (it) interested me somehow
   fio    >>+gz at mar-->l.07

03 NAD:                 [uh:              ]

04 FIO: also (.) fand ich irgendwie immer
        well    (I) found it somehow always

05      intress[ANT <<dim>kuns]tgeschichte,>=
        interesting      art history

06 MAR:        [hmhm-          ]

07 FIO: =un_dann dAcht ich mir;+
         and then I thought to myself
   fio                  -->+closes eyes-->l.08

08     °h +>WAR ich hier no_nich+ so< £ähm+ (0.53)£ +berei:t
           I wasn't here like ready yet uhm
   fio -->+gz right----------->+gz up--->+gz down r+gz in front
                                              -->l.09

09     jetz so+ voller e+l!AN! in so_n +großes JUrastudi$um=-
       now like start full of verve such long law studies
   fio    -->+gz front +gz down left->+gz at MAR-->l.13
   mar                                        $nods
                                              ->l.11

10     =oder was weiß I[CH ir$gendne,           ]
        or whatever       some

11 MAR:                [ja:  $ isch halt schon H]EFtig ne,=
                        yes    it is PART PART a big deal (tag)
   mar                   -->$

12 FIO: =ʃja:-
   sin   ʃnods-->l.13
```

```
13      und dann >hab_ich gedach°t ich°<+%ʃ# (0.58)% mach
        and then I thought          I           first make
   fio                           -->+closes eyes-->l.14
                                      %head l&r,purses
                                      lips---->%shrug->l.14
   sin   --------------------------------->ʃ

   fig                                      #fig1
```

fig1

```
14      jetz% mal zwEI +semester %was ich% §EINfach-
        now  PART two semester    what I    just
   fio    -->%                    %shrug  %
                  -->+gz at MAR-->l.16
   mar                                    §nods-->l.18

15      was mich irgendwie so (.) °ja;°
        what I'm kind of like     yes

16 SIN: int[res ]SIER+T-
        interested in
   fio            -->+gz at SIN-->>

17 FIO:    [was ]
             what

18 SIN: °h§ ja.
            yes
   mar  ->§
```

Within the same turn, the quotative is first used in preterit tense *und dann dacht ich mir* (line 07) and then in present perfect tense *und dann hab ich gedacht* (line 13). The *dachte* quotative is used in response to a temporally extended (and potentially still valid) emotional state, marked through the temporal adverb *immer* ('always'): Fiona has always found art history interesting (line 04). With the imperfective form *dachte*, Fiona would, according to the pattern that I found in my data, introduce an action that is staged as spontaneous and emotionally charged (see also Excerpts 6.1 or 6.2). The continuation of her turn shows, however, that some information still seemed to be missing in order for her recipients to understand her point. With a self-repair of the actional trajectory from quoting herself—as projected by the social action format—to the assertion *war ich hier noch nich so bereit* ('I wasn't here like ready yet'; line 08), Fiona parenthetically inserts further information (lines 08–10), describing additional (emotional) reasons for first studying art history. The verb-first syntax in line 08 and the preterit tense demonstrate that this is probably the continuation of her ongoing turn. Lines 02 to 09 allow Fiona to describe a state of mind without clear temporal boundaries. The use of imperfective

forms throughout this stretch of talk fits the actions that Fiona accomplishes: She frames her informings as potentially necessary background information for what comes next.

After an affiliative response from Marius (line 11) and Fiona's confirming *hm* (line 12), the latter continues her turn with direct reported thought, introduced with *und dann hab ich gedacht* (line 13), now using a quotative in present perfect (and thus potentially with perfective aspect). Instead of delivering background information without temporal boundaries, Fiona here reports—and depicts[145]—an individual event, which is marked as such using *jetzt mal* (lit. 'now once') in the reported clause (line 14): Fiona's decision of first studying two semesters something that she is interested in (line 16) as a hobby (not in transcript).

This excerpt raises the question why the speaker uses the same quotative in two different tenses. I suggest that, in this specific case, the respective use of *ich dachte mir* and *ich habe gedacht* may be related to action formation.[146] Whereas a thought reporting a decision or conclusion from a past situation is mainly introduced by a quotative in present perfect in my data (*ich habe (mir) gedacht*), thoughts that report emotional states are often introduced by a quotative in preterit (*ich dachte (mir)*). Excerpts 6.2 and 6.3 in Section 6.2.5, where affective stances are displayed through direct reported thought, support this hypothesis. As my corpus is restricted in number, the hypothesis of a verb-specific actional distribution needs to be checked with a larger corpus, of course.

The distribution of German past formats according to action formation has not yet been addressed by prior literature (but see Fiedler forthc.). Instead, other distributional mechanisms have been debated, such as aspect (Eisenberg 1994; Heinold 2015: 102ff; Henriksson 2006; Rödel 2007: 38 and 42ff) even though German tempus is considered not being ruled by aspect (Henriksson 2006; Rödel 2007: 50ff) unlike, for instance, Russian. Research on the German tempus system has also focused on a mainly semantic description following the rules of formal logic which operates with truth-values (following Reichenbach's (1947) tempus system). More generally, a closer analysis of tense formats and their use in everyday talk has not been carried out yet (but see, on an analysis according to semantic action categories, Leonhard 2022).

Geographical variation, too, has been the focus of tempus research, indicating that in Southern Germany, the so-called *oberdeutscher Sprachraum*, the analytic past is preferred to the synthetic past for most verbs (Zeman 2010). Recent research on the German Alemannic region, however, suggests that the preterit may have found its way back from standard German into southern German dialects (Leonhard 2022). Another common finding is that the preterit is generally used in written language while the present perfect is the tempus of spoken language (Fischer 2018: 213; Zeman 2010).

In my data, independent from the geographical origin of the speaker, the preterit form *ich dachte (mir)* is amply used. Zeman's claim that the present perfect has replaced the preterit cannot be confirmed in the present work and the overall development called *Präteritumschwund* ('preterit decrease'; see also Fischer 2018) may need to be revised

---

[145]   Fiona's bodily conduct contributes to this depiction: she shrugs, purses her lips, and closes her eyes (fig1).

[146]   It remains an open question so far whether this potential distribution is not limited to speakers who actually have both tenses available in their "inventory" of spoken language. Speakers who do not use preterit at all and never have may not have the linguistic resources to accomplish different actions through tense distribution.

regarding certain verbs. A systematic verb-specific study of the distribution of preterit and present perfect in spoken language (and verb-specifically) could clarify whether action formation is important—a claim that Fiedler (forthc.) makes in her study on German *finden* ('find'), *meinen* ('mean'), and *glauben* ('to believe') in their past tense formats.

It is not unusual for routinized forms to be an exception in processes of language change (see also Section 6.3 on French). When occurring in fixed patterns, specific forms, especially with frequent verbs like *haben* ('to have') or *sein* ('to be'), may persist while less frequent verbs disappear. A standard example are the preterit forms of *sein* ('to be'), occurring in various routinized patterns like *war*+ADJ (e.g., *Der Film war gut*, 'The film was good'). A similar (certainly ongoing) process of routinization of quotatives with *denken* may have the impact that *ich dachte (mir)* is, regardless of its tense marking as preterit, recurrently used to introduce direct reported thought, thus allowing speakers to display their affective stance. In contrast, when a decision-making is reported, speakers may resort recurrently to the quotative *ich habe (mir) gedacht* in present perfect tense.

A detailed analysis of a potential distribution according to action formation of *ich dachte (mir)* and *ich habe (mir) gedacht* lies beyond the scope of the present work. The initial observation that both forms are used but do not necessarily carry out the same actions hint toward a distribution according to action formation. If we look at the numbers, the present perfect ($n = 87$) is actually more frequent than the preterit ($n = 50$). The reason for this difference in number may be that direct reported thought with *ich habe (mir) gedacht* allows speakers to take an affective *and* an epistemic stance (i.e., reasonings), whereas *ich dachte (mir)* is primarily restricted to affective stance-taking. An argument in favor of this hypothesis is that, when including all functions (i.e., the change of an epistemic state, etc.), the frequencies become more balanced ($n = 133$ for *ich dachte (mir)* and $n = 151$ for *ich habe (mir) gedacht*).

These result showsthat one *cannot* say that synthetic past forms are generally not used anymore in German or that they are only used in certain geographical areas. What my data shows is that, when *denken* functions as a quotative, the synthetic past form continues to be used in spoken language. This also refutes the second, recurrent claim that the preterit is more and more used in written language only (Fischer 2018: 213). Leonhard (2022) recently argued that standard German might, in fact, influence southern German varieties so that the latter re-integrates preterit forms into spoken language. I observe that even speakers from the German Alemannic area, the variation that Leonhard (2022) investigates, use both past forms. There does not seem to be a preference for either of them. With the results of the present work, I would like to suggest a hypothesis that could add to Leonhard's (2022) findings. It could be possible that the frequent verb *denken* is in the process of routinizing, similar to the English simple past form *I thought* (Kärkkäinen 2012), into patterns that are recurrently used for the same interactional purposes: While the perfect form, *ich habe (mir) gedacht*, allows speakers to recurrently introduce a post-hoc rationalization as a kind of epistemic stance, the preterit form *ich dachte (mir)* shows a clear tendency toward allowing speakers to take an affective stance. Both past forms may have reached routinization in their role as social action formats recurrently projecting those two different actions independently from aspect or tense.

6.2.8.   Conclusion

The preceding two analytic sections have demonstrated that there is a tendency for *ich dachte (mir)* and *ich habe (mir) gedacht*, when introducing direct reported thought, to carry out specific actions with distinct frequencies: *Ich habe (mir) gedacht* allows speakers, in my data, to take both an affective or an epistemic stance to display their emergent reasoning, whereas *ich dachte (mir)* is almost exclusively used for affective stance taking. The ratio of the distribution regarding the two actions is about 2:1 with the analytical past form and 8:1 for the synthetic past form, as the following table illustrates.

Form	Frequency affective stance	Frequency display of reasoning	Frequency other functions
ich dachte (mir)	78.0% (39/50)	10.0% (5/50)	10.0% (5/50)
ich habe (mir) gedacht	64.4% (56/87)	36.8% (28/87)	3.4% (3/87)

Table 16: Distribution of tense according to action formation

This distribution according to action formation can be demonstrated with two characteristics where the two patterns [*ich dachte (mir)* + direct reported thought] and [*ich habe (mir) gedacht* + direct reported thought] differ: The format of the introduced quotes and the speakers' bodily conduct. Let me begin with the bodily conduct of the speaker.

What my analysis illustrated is that the two different stance displays come with a difference in multimodal conduct. Speakers tend to embodiedly animate their affective stance-taking, which is less the case when they take an epistemic stance to display a reasoning. This difference in embodied conduct may be due to the different activities in which the reported thought occurs. While the epistemic stance to display a reasoning is used in argumentative or explanatory environments, the affective stance-takings occur in more emotional sequences, where speakers stage their story as, for instance, amusing through bodily means.

In addition to the opposition depicting *vs.* non-depicting, two gaze patterns co-occurring with direct reported thought have been determined:

(1) Speakers gaze away during the reenactment and gaze back at the recipient—at the latest—at the end of the quote (n = 62/137[147] for *ich dachte (mir)* / *ich habe (mir) gedacht*). This pattern has already been described by previous literature as a means to check the recipient's attention (Thompson & Suzuki 2014). My analysis has shown that this gaze pattern might also indicate shifts in footing. Gazing back at recipient(s) at the end of the quote seems to be a means to mark a shift back from the story world to the here-and-now of the ongoing conversation. Gazing at the recipient(s) may thus also be a means to yield a turn and allocate it

---

[147]   The low number for the German data may be due to the nature of this data. In some videos, interlocutors are involved in activities such as cooking, driving, or renovating a room. This involvement directs the interlocutors' gaze mostly to the involved objects or activities, such as painting or stirring. If I exclude all occurrences in which the speaker is involved in an activity that occupies their gaze, this reduces the number of occurrences by n = 22. This heightens the ratio of the pattern from 62/136 to 62/114. If I do not count either in the total number those cases where gaze is not visible because of the recording angle (n = 12), the ratio increases again to 62/102.

to a potential next speaker, especially in triadic settings (Auer 2021). The second pattern, however, has, to my knowledge, not yet been addressed.

(2) Speakers gaze away during the quotative and the reenactment and do *not* gaze back at the recipient at the end of their quote (*n* = 19/137 for *ich dachte (mir)* / *ich habe (mir) gedacht*; e.g., Excerpt 6.1, Section 6.2.5). This gaze pattern seems to be related to the activity speakers are involved in. I systematically observe this pattern in my data when the direct reported thought occurs in delicate situations (14/15 occurrences for *ich habe (mir) gedacht*, all four occurrences of *ich dachte (mir)*). As the recipient's gaze at the speaker is recurrent for *ich dachte (mir)* / *ich habe (mir) gedacht* (*n* = 74/121),[148] mutual gaze would be highly probable (for a detailed discussion of this pattern, see Section 7.4).

A quantitative analysis of other multimodal conduct apart from gaze revealed the following:[149]

Multimodal conduct	Frequency *ich dachte*	Frequency *ich dachte mir*
gaze at the interlocutor(s) at the end of the quote	51.3% (19/37)	69.2% (9/13)
reenacted gaze[150]	54.1% (20/37)	76.9% (10/13)
eyebrow raise	10.8% (4/37)	-
facial expression	54.1% (20/37)	61.5% (8/13)
dropping hand gesture	8.1% (3/37)	7.7% (1/13)
palm-up open-hand gesture	5.4% (2/37)	15.4% (2/13)
pointing gesture	2.7% (1/37)	7.7% (1/13)
other gesture	13.5% (5/37)	15.4% (2/13)
nod	5.4% (2/37)	-
shrug	16.2% (6/37)	-
change in posture	10.8% (4/37)	15.4% (2/13)
freezing	-	7.7% (1/13)
no bodily conduct	8.1% (3/37)	-
not visible	18.9% (7/37)	-

Table 17: Multimodal features of [*ich dachte (mir)* + direct reported thought]

Multimodal conduct	Frequency *ich habe gedacht*	Frequency *ich habe mir gedacht*
gaze at the interlocutor(s) at the end of the quote	54.7% (29/53)	26.5% (9/34)
reenacted gaze	47.2% (25/53)	32.4% (11/34)
facial expression	28.3% (15/53)	20.6% (7/34)

---

[148] Again, the low number for the German data may be due to the nature of the data. When recipients are involved in another activity, their gaze is directed toward this activity (*n* = 36). Also, as mentioned before, in some recordings, gaze is just not visible because of the positioning of the cameras (*n* = 12). If I exclude both of these numbers from the overall number of occurrences, the ratio becomes significantly higher: recipients then gaze at the speaker in 74 of 89 occurrences.

[149] Because some resources can occur simultaneously, such as speakers making a gesture while lifting their eyebrows, the numbers do not add up to the total number of occurrences of both forms.

[150] This category comprises instances where speakers reenact gazing at an imaginary interlocutor on their left or right when they depict an affective stance through gaze (looking up, squinting their eyes), or when they reenact a gaze away while reporting a dispreferred action during the quote.

head tilt	7.5% (4/53)	8.8% (3/34)
dropping hand gesture	18.9% (10/53)	8.8% (3/34)
palm-up open-hand gesture	1.9% (1/53)	-
pointing gesture	-	5.9% (2/34)
iconic gesture	3.8% (2/53)	8.8% (3/34)
other gesture	5.7% (3/53)	2.9% (1/34)
nod	1.9% (1/53)	2.9% (1/34)
headshake	17.0% (9/53)	11.8% (4/34)
shrug	18.9% (10/53)	5.9% (2/34)
change in posture	9.4% (5/53)	5.9% (2/34)
freezing	3.8% (2/53)	-
no bodily conduct	11.3% (6/53)	17.6% (6/34)
not visible	7.5% (4/53)	11.8% (4/34)

Table 18: Multimodal features of [*ich habe (mir) gedacht* + direct reported thought]

The one body movement that is recurrent are facial expressions during the quote.[151] However, they are more frequent with the synthetic past form (54.1% for *ich dachte*, 61.5% for *ich dachte mir*) than with the analytic past form (28.3% for *ich habe gedacht*, and 20.0% for *ich habe mir gedacht*). One reason for this result may lie in the actions that speakers accomplish with each of these past forms. Facial expressions are systematically involved in affective stance-taking activities (Ruusuvuori & Peräkylä 2009; Peräkylä & Ruusuvuori 2006, 2012; Kaukomaa *et al.* 2015), which is why they may occur more with *ich dachte (mir)*, which primarily allows speakers to display their affective stance with direct reported thought. That eyebrow raises only occur with *ich dachte* may support this hypothesis, as they have also been demonstrated to contribute to affective stance-taking, specifically while displaying surprise (Debras & Cienki 2012; Ekman *et al.* 2002).[152,153]

One main result of this quantifying analysis is that only some body movements are somewhat recurrent. Given former research on the co-occurrence of multimodal conduct and linguistic practices, this result is, however, not surprising (see, for instance, Schoonjans 2018). Because of the low number of cases, no conclusions can be drawn from most percentages. Also, the different settings of my German data may influence especially gaze conduct, as a side-by-side

---

[151] Due to the diversity of my German data, this quantification needs to be treated with caution. Because speakers sometimes sit side-by-side, are involved in activities that are not stationary, and because the participant numbers range from two to five participants, bodily conduct is suspected to be more diverse than in the French data. The multitude of participant frameworks lead to the decision to primarily focus on the production side of multimodal conduct and less on the recipient side. It would be interesting for future research to investigate the same action within different participant frameworks, thus determining potential differences in multimodal conduct.

[152] Note, however, that head tilts have been described to similarly do so (Debras & Cienki 2012) and only occur with *ich habe (mir) gedacht*.

[153] Two gestures occurring more often with the analytic past form are headshakes (15.0%), which do not occur at all in my data with *ich dachte (mir)*, and the so-called 'dropping hand gesture' (Bressem & Müller 2014b: 1584)i.e., gestures where speakers let their hand fall down suddenly and quickly), with 15.0% for *ich habe (mir) gedacht* compared to 10.0% for *ich dachte mir*. Shrugs are relatively equal for *ich dachte (mir)* (13.8%) and *ich habe (mir) gedacht* (15.0%). Unfortunately, because of the low number of cases, these body movements do not allow for any conclusions to be drawn from them.

seating configuration may lower the speaker's recipient gaze drastically. These data-related difficulties are to be born in mind when interpreting the quantifications of multimodal conduct.

What is interesting is the category "no bodily conduct." Of course, speakers cannot "not have a bodily conduct." However, there are cases where speakers simply do not change their bodily conduct during direct reported thought. When they simply hold the same gaze, posture, and gesture—also by *not* performing them—I counted them as "no bodily conduct." This is more often the case with the analytic past form (13.8%) than with the synthetic one (7.5%). One possible reason for this difference in frequency may be, as with facial expressions, the action that speakers accomplish with direct reported thought.

In addition to an analysis of body movements, I also investigated prosodic features. Four morpho-phonological and prosodic features were considered for each occurrence of *ich dachte (mir) / ich habe (mir) gedacht*, including acceleration, accentuation, a pause between the quotative and the quote, and whether the quotative is morpho-phonologically reduced. There is no clear correlation with the action of displaying a reasoning with *ich habe (mir) gedacht* and the elision of the schwa ($n = 6/14$ for the non-reflexive, $n = 9/14$ for the reflexive form). On the other side, when speakers take an affective stance, the introducing pattern *ich dachte (mir)* is, in general, more reduced than only with an omitted schwa ($n = 37/39$ for the non-reflexive; $n = 18/20$ for the reflexive form). This means that, when displaying an affective stance, quotatives introducing direct reported thought tend to be more reduced. More generally, the reflexive forms—*ich dachte mir*, and those occurrences of *ich habe mir gedacht* that speakers use to display their affective stance,[154]—show, on a morpho-phonological level, a higher degree of routinization than the non-reflexive forms (accomplishing the same action).

Concerning the different syntactic patterns of the bipartite structure [quotative + quote], the quotes that *ich habe (mir) gedacht* introduces slightly diverge from those introduced by *ich dachte (mir)*. In 64.4% of all occurrences, *ich habe (mir) gedacht* is followed by a dependent main clause. This percentage is lower for the synthetic past form (56.0%). The reason for this divergence may, like the difference in bodily conduct, be related to action formation. Because speakers use the pattern [*ich habe (mir) gedacht* + direct reported thought] more often to make publicly available their emergent reasoning, complete clauses as quotes may be generally more frequent. Syntactically, in these cases, *ich habe (mir) gedacht* is not completely independent from what follows: Instead, there is still a loose syntactic bind that goes beyond a mere pragmatic relation between quote and quotative (see Section 6.2.3.1 on dependent main clauses and their syntactic status). The second part of the bipartite structure seems to be, however, maximally independent when [*ich habe (mir) gedacht* + direct reported thought] is used for affective stance-taking, as with *ich dachte (mir)* (see Section 6.2.5), when the quote consists of response cries, non-clausal units, or body movements.

The partial distribution of the two past tenses according to two actions raises the question whether these actions are not also somewhat related to the past tense formats of *denken*. Section 6.2.7 illustrated that simply stipulating the decline of preterit tense in German may be too restrictive and may need some revision regarding certain verbs. In my data, the preterit format *ich dachte (mir)* and the present perfect format *ich habe (mir) gedacht* show a tendency of being

---

[154]   All but one occurrence of *ich habe mir gedacht* is used by speakers to take an affective stance.

distributed according to the actions that they accomplish in conversation,[155] regardless of the speakers' region of origin. While the amount of data in this work remains too small to make any general claims about the distribution according to action formation, my observations open a possible new line of investigation concerning the German tempus system. Further research is needed to investigate this hypothesis with a larger data set.

6.3. Reporting Thought with the French *Je me suis dit* ('I have said to myself') and *J'étais là* ('I was there')

When speakers report their own thought in my French data, they do so with the *direct* reporting format, which meets the following three syntactic features:

- no complementizer between the quotative and the quote,
- no change of the personal deixis (Holt 1996),
- no temporal shifts of the verb in the quote.[156]

In French linguistics, Blanche-Benveniste (2010) has called this direct format *l'effet deux points* ('the colon effect'), thus picking up on punctuation rules indicating directness in written language. Research on prosody in direct speech formats has shown, however, that in spoken language, such parallels are not systematically distributed (Klewitz & Couper-Kuhlen 1999). In what follows, I will treat the investigated formats, *je me suis dit* ('I say to myself') and *j'étais là* ('I was there'), in two separate chapters. The formats differ on three levels: syntactically, concerning their interactional use (including sequential features), but also multimodally.

Before presenting both constructions in more detail, let me briefly introduce their respective morphological structure and some morpho-phonological features. *Je me suis dit* is composed the following way:

je	*me*	*suis*	*dit*
PP.1SG	myself.REFL	be.AUX.1SG	say.PTCP

'I said to myself'

It is the first-person analytic past of the reflexive form of *dire* ('to say'). *Je me suis dit* can occur in a morpho-phonologically reduced form and with certain prosodic features, as the following table shows:

Feature	Frequency
omission of one or both of the final 'e's in *je* ('I') and *me* ('myself')	19.5% (8/41)
omission of the *m* in *me* ('myself')[157]	22.0% (9/41)

---

[155] These actions are not to be confused with Vendler's (1967) *Aktionsarten* which are independent of the speaker and refer to semantic characteristics of a verb such as *state*, *achievement*, or *accomplishment*.

[156] In French, the conjugated verb of the complement clause has to follow the so-called rule of the *concordance des temps*—the temporal shift of the verb in the quote when the introducing verb is in past tense (Le Goffic 2019: 81). A recurrent phenomenon is, however, a shift to the narrating present tense in the quote (Léon 1988: 121).

[157] Indeed, speakers do not pronounce the *m*. Speakers do not separate *je* ('I') and the *e* of *me* ('myself') by a glottal stop but pronounce something close to an elongated *e*. Out of these 10 cases, nine are also

omission of the personal pronoun *je* ('I')	2.4% (1/41)
accelerated	58.5% (24/41)
no accentuated syllable	80.5% (33/41)
pause after *je me suis dit*	14.6% (6/41)

Table 19: Prosodic and morpho-phonological features of *je me suis dit*

The reduction of final 'e's is usually called *élision* in French. However, this reduction process concerns cases where the following word starts with a vowel (*je imagine > j'imagine*) and is usually also graphically displayed in written language (Abeillé & Godard 2022: 2120). With some exceptions, the *élision* is obligatory in spoken and written French. When the following word starts with a consonant and the drop of the *e* is not also a written convention but a feature of spoken language only, the phenomenon is called *la chute du* e *caduc* (*ibid.*: 2121). Dropping an *e* in the middle or at the end of a word is a common practice in oral French and cannot be necessarily interpreted as a sign of routinization (Riegel *et al.* 2003: 57 and 79).[158] In about a quarter of all cases, another reduction process occurs: the omission of the 'm' in the reflexive pronoun *me* ('myself'). The omission of the consonant 'm' and of the final 'e' both contribute to the perception of *je me suis dit* as one chunk rather than an assembly of several words—a phenomenon that Pekarek Doehler (2016) also evidenced for the French *je ne sais pas* becoming *chais pas*. Compared to German, the personal pronoun is only dropped in one case out of 41. Reducing or omitting the personal pronoun does not seem to be the site of morpho-phonological reduction in French, as it is in German where *ich* is recurrently reduced to *ch* or *i* (see Section 6.2.2).

In my data, the quote that is introduced by *je me suis dit* remains mostly prosodically flat. That means that prosodic features of emotional involvement or affectivity like high pitch movements or strong accentuation are rare (*n* = 6). Also, the quotative and the quote are primarily delivered under one intonation contour, with the focus accent being placed in the quote, which creates an effect of cohesion between the quotative and the quote.

J'étais là is composed the following way:

j'	*étais*	*là*
PP.1SG	be-IPFV.1SG	DEIK

'I was there'

The pattern consists of the first-person imperfective form of *être* ('to be') and a deictic particle that usually refers to a farther location or moment. When *j'étais là* is used as a quotative, which

---

accelerated, which may indicate a tendency to morpho-phonological reduction, compared to the omission of E's, where three of the seven cases are pronounced with normal speed. Of course, these results are the outcome of impressionistic hearing and have not been checked with a software like praat, which is why they should be checked with a bigger data set and a more technical approach.

[158] Riegel *et al.* (2003) note that "most of the final *e* [are] maintained in written language: *un(e) autr(e) histoir(e) est arrivée.*" (*ibid.*: 79). This is also what I observe in the literature on French prosody, for instance Avanzi and Glikman (2009), where the final *e* of *je* ('I') in *je pense* ('I think') or *je crois* ('I believe') is orthographically realized in the examples even though the Analor-pictures in the annex clearly show a morpho-phonological reduction (*ibid.*: 141).

it is almost exclusively in my data (57/60 occurrences), the deictic *là* ('there') seems to project an upcoming depiction of the speaker's (physical or emotional) state in the here-and-now of the ongoing conversation. The data shows that speakers reenact their state at a specific point (in time or space) that is deictically referred to with *là* for the co-participant to literally 'see' what happened in the past, but in the ongoing conversation. The construction thus invites co-participants to experience the reported situation with the speaker by monitoring the speaker's bodily conduct as it emerges in real time.

Concerning its morpho-phonologic features, see the following table:

Feature	Frequency
morpho-phonological reduction: omission of the initial 'e' in *étais* ('was')	73.7% (42/57)
accelerated	66.7% (38/57)
pause after *j'étais là*	29.8% (17/57)
no accentuated syllable[159]	96.5% (55/57)

Table 20: Prosodic and morpho-phonological features of *j'étais là*

The table shows that *j'étais là* is recurrently morpho-phonologically reduced. When the initial "e" of *étais* is not pronounced, speakers either use *j'tais là* with voiced [ʒ] or *chtais là* with unvoiced [ʃ]. In addition to that, *j'étais là* is almost never prosodically stressed. These two features may be signs of frequent use. Compared to *je me suis dit*, it is striking that the pause after the quotative is more frequent with *j'étais là*. As *j'étais là* has been evidenced to being systematically followed by (bodily) depictions (Cheshire & Secova 2018), these pauses may be an indication for the speaker either making a dramatic pause to "prepare" the depiction, or for already performing a body movement. Indeed, in 14 of the 17 cases, speakers prepare or perform already a gesture (*n* = 3), a posture (*n* = 2), or a facial expression (*n* = 7). In two other cases, a facial expression is coupled with gestures.

In the following, I will present *je me suis dit* and *j'étais là* as social actions formats systematically projecting a quote in everyday talk.

6.3.1. The French *Je me suis dit*: Direct Reported Thought as a Stance-Taking Device to Display a Reasoning

Whereas *être là* does not occur in dictionaries (yet), *se dire* can be found in the dictionary as "dire à soi-même, penser" (Robert 2013: 745) and has been documented in literature since the 19th century as a synonym for *to think*, for example in Stendhal's Lucien Leuwen: "Il se disait vaguement qu'il ne jouirait pas longtemps encore de toutes ces belles choses, et à cause de cela elles ne lui donnaient pas d'humeur."[160] (Stendhal 1835: 368, cited in TFLi, *dire*). Whereas examples from literature always comprise a main and complement clause which are connected with the complementizer *que* ('that'), the Petit Robert from 2013 also gives an example without complementizer: "*Je me disais: il faut partir.*" ('I told myself: I have to leave'; (Robert 2013: 745, original emphasis). The Petit Robert also provides a synonymous clause "je me faisais

---

159   As with *je me suis dit* and with the German quotatives, there is, however, a recurrent stress on the item standing between the quotative and the quote (*n* = 17).

160   Roughly translatable as 'He vaguely told himself that he would not enjoy all these beautiful things for a long time yet, and because of that they did not give him any joy.'

cette réflexion" (*ibid*.: 745) with the literal meaning 'I have made myself this reflection'. In line with my data, this synonym suggests that the form *se dire* is used to report a thinking-process or reasoning.

Whereas the verb *dire* and constructions with *dire* have been the object of several studies (Anscombre 2010; Gómez-Jordana Ferary & Anscombre 2015; Moreno 2016), none on these articles treats the reflexive form of *dire* explicitly. Prior studies on direct reported speech in French primarily focus on syntactic and semantic features of direct reported speech (Authier-Revuz 2019; Blanche-Benveniste 1988; De Cornulier 1978; Gachet 2012), often based on the investigation of written language (but see Bangerter *et al.* 2011; Berger & Pekarek Doehler 2015; on Canadian French, see Vincent & Dubois 1997). Building on the (mainly formal) features that these studies have investigated, the present work focuses on *je me suis dit* as a social action format that allows speakers to project direct reported thought.

6.3.1.1.   Combinatory Patterns with *Je me suis dit* ('I have told myself')

From a total of 43 occurrences, two cases had to be excluded because of an interruption directly after the construction. In contrast to the German past forms *ich dachte (mir)* / *ich habe (mir) gedacht* ('I (have) thought (to myself)'), *je me suis dit* ('I have said to myself') is used exclusively to introduce reported thought. Consequently, my collection includes all 41 instances (after exclusion of the two interrupted cases) of *je me suis dit* because they are all used to introduce a quote.

The following table illustrates that *je me suis dit* can occur in patterns. These patterns are not as frequent as they are in German (see Section 6.2.4); however, a certain tendency can be observed despite the small data set. Lexical items preceding *je me suis dit* mark the high topical and syntactical cohesion between the turn with *je me suis dit* and (a) prior turn(s). In 27 of the 41 occurrences, *je me suis dit* is preceded by lexical items. The following table gives an overview:

Form	Number of occurrences
(et) pis je me suis dit (then I have said to myself)	9
après je me suis dit (then I have said to myself)	2
mais je me suis dit (but I have said to myself)	5
(mais) moi je me suis dit (but me I have said to myself)	2
et je me suis dit (and I have said to myself)	1
du coup je me suis dit (so I have said to myself)	1
à un moment je me suis dit (at some point I have said to myself)	1
alors je me suis dit (so I have said to myself)	1
où je me suis dit	1

(where I have said to myself)	
tu vois je me suis dit (you see I have said to myself)	1
en plus je me suis dit (in addition I have said to myself)	1
parce que je me suis dit (because I have said to myself)	1
[pause][161] je me suis dit (I have said to myself)	2
je me suis dit (I have said to myself)	13
TOTAL	41

Table 21: Combinatory patterns with *je me suis dit*

Similar to German, the most frequent pattern contains the adverbials *puis* or *après* ('then') and/or the coordinative conjunction *et* ('and'; $n = 12$). The relative frequency of pattern with adversative *mais* ('but'; $n = 6/41$) is higher with the French *je me suis dit* than with German *ich dachte (mir) / ich habe (mir) gedacht* ($n = 4/137$), even though the absolute frequency remains low. Other adverbials (*alors*, 'so'; *en plus*, 'in addition') or particles (*enfin*, 'well') are not frequent but still allow speakers to link their quotation to preceding talk and/or to create intersubjectivity with an interlocutor by using particles in second person (*tu vois* 'you see').

Among the two occurrences with a preceding pause, one comes in after a lack of uptake by the interlocutor. The other occurrence follows a 1.9 s silence where none of the two interlocutors takes the turn after topic closure.

6.3.1.2. Syntactic Features of *Je me suis dit* ('I have said to myself')

On a syntactic level, *je me suis dit* ('I have said to myself') calls for an obligatory direct object, which is, in the specific case of direct reported thought, filled by the quote. With the obligatory argument position filled, the complement-taking predicate *se dire* is syntactically complete. Additionally, *je me suis dit* can, but does not need to have an indirect object. In the *je me suis dit* construction, the reflexive pronoun *me* fills the position of the indirect object.

Je me suis dit can introduce various formats of quotes. Quotes with main clause syntax are the most frequent format ($n = 21$, out of which two are preceded by response cries). There is no syntactic relation between *je me suis dit* and the clause that fills its empty argument position because the complementizer *que* ('that') is missing. I would like to remind the reader that in French, the complementizer is, other than in German, the only possibility to syntactically link a complement-taking predicate and its complement. However, because the quote still fulfills the standard format of a complement (i.e., a clause), and because in French, the word order in the complement clause does not change without complementizer, the juxtaposition of

---

[161]   I considered as pauses gaps that are longer than 0.25ms. Everything under 0.25ms has been shown to be the usual, non-perceived gap between two turns (Stivers *et al.* 2009).

the two clauses comes close to the syntactically linked form. As I pointed out in Chapter 2, to grasp all possible combining patterns in my data, I describe the format [quotative + quote] as a bipartite structure. One of the reasons for this decision is that there are cases where the quotative is not followed by a clausal format. The following table shows these configurations:

Quotative	Response cries only	Bodily conduct only	Verbless structures	Question (words)	Imperatives
je me suis dit	4.9% (2/41)	2.4% (1/41)	7.3% (3/41)	14.6% (6/41)	2.4% (1/41)

Table 22: Formats of direct reported thought introduced by *je me suis dit*

Because questions and imperatives keep their direct format and are not reformulated into dependent structures (e.g., *je me suis dit* qu'est-ce qu'elle fait ('I have said to myself what does she do') and not *je me suis dit* ce qu'elle faisait (lit. 'I have said to myself that what she did')), the syntactic independence may become more visible[162] than with "simple" main clauses. Another five instances are followed by either a complex clause like a left dislocation or another CTP-complement clause structure, like in the following example:

Ex. 6.8: fouiller partout / snoop around everywhere (Pauscaf 19, 18min40)

```
10 ORE: j'me suis dit j'pense pas qu'c'est possible
        I told myself I don't think that it's possible

        qu'ils aient une caméra quand même,=
        that they have a camera though
```

In another two occurrences, the speakers initiate a self-repair just after *je me suis dit*. Similar to German, there is thus a tendency for speakers to use the direct reported thought format more often than the indirect format even though an omission of the complementizer has been argued to not be common in spoken Modern French[163] (Blanche-Benveniste & Willems 2007: 226), compared to spoken German, where (non-negated) *verba sentiendi* and *verba dicendi* have a stronger tendency to 'attach' complement clauses without complementizer (Helbig & Buscha 2001: 568ff; for a justified correction of the verbs listed by Helbig & Buscha 2001, see Auer 1998: 4).

---

[162]   When questions have a question word, the differences between a dependent and a non-dependent syntax are more subtle because dependency is then only marked through a tense shift in the dependent clause. For instance, *Je me suis demandé pourquoi elle **a** fait ça* (independent complement clause) vs. *Je me suis demandé pourquoi elle **avait** fait ça* (dependent complement clause).

[163]   Blanche-Benveniste and Willems (2007) mention in a footnote, however, that even though there does not seem to be a similarly systematic omission of *que* in French as in English, there seems to be the possibility of zero complementizer in Parisian French and in Canadian French as well as the tendency of *que* becoming phonologically reduced (*ibid*.: 226, especially footnote 10). Glikman (2008, 2013) has shown that in Ancient French, it was very common to not use *que* ('that') after certain verbs, among others, *verba sentiendi* or *dicendi* like *croire* ('believe') or *penser* ('to think'). Comparing the epos *Chanson de Roland* to later works, she convincingly argues that the use of *que* is a more recent phenomenon which might have its origin in oral language for disambiguation purposes: "The tendency to repeat *que*, to avoid a possible ambiguity, goes hand in hand with the more and more frequent habit of expressing *que* at the head of subordinate clauses, where the oldest French preferred the juxtaposed form" (Glikman 2013: 11).

Due to the absence of syntactic marking, the syntactic status of direct quotes has been much discussed by researchers working on reported formats in French (Authier & Meunier 1977; Authier-Revuz 1978, 1979; Blanche-Benveniste 1988; Mosegaard Hansen 2000). Different criteria have been established by prior research to show that direct quotes do not have the status of the direct object of a matrix verb. Blanche-Benveniste (1988), for instance, shows that direct quotes cannot be pronominalized neither can they be subject to right dislocation (*ibid*.: 55ff). This rule applies, according to her, to verbs like *murmurer* ('to whisper') or *soupirer* ('to sigh') because they also indicate *how* something has been said, sometimes even without linguistic "material," let alone complete clauses. However, the rule does not apply to *dire*.[164] The author explicitly states that *dire* ('say') can be pronominalized and it can also be subject to right dislocation because it syntactically demands a direct object. Her grammatical explanation for this exception is that direct quotes introduced by complement-taking predicates such as *dire* correspond to a syntagm "which is taken as a regular complement, equivalent to […] *elle me le disait, ce truc-là* ('she said it to me, this thing') […]" (*ibid*.: 56, my emphasis and translation). Such instances with *dire* are not part of her analysis, which influences, of course, her claims about the syntactic status of the quote.

At least in my data, *se dire* shares many features with Blanche-Benveniste's (1988) examples with *soupirer* or *murmurer*: *Se dire* can, for instance, introduce prosodically animated and/or non-clausal elements. Consequently, Blanche-Benveniste's (1988) conclusion for the grammatical status of quotes introduced by *soupirer* or *murmurer* seems equally valid for *se dire*: "In absence of a relation with a pronoun, direct discourse does not represent any grammatical category; it is, from a grammatical point of view, 'without form'" (*ibid*.: 56, my translation). As I mentioned earlier, Blanche-Benveniste draws this conclusion also on the observation that what is reported does not have to be verbal material but can comprise anything, including sounds and gestures (*ibid*.: 57). With her analysis of the quote lacking a grammatical form, Blanche-Benveniste (1988) contradicts De Cornulier (1978), who classifies direct quotes generally as *mimes*. When such *mimes* are marked as quotes by a complement-taking predicate, he argues that "they can clearly constitute a complement, usually a direct object like in (71): Il dit : – Zut! [('He says: – Damn!'), SF]" (*ibid*.: 79, my translation). Mosegaard Hansen (2000) draws similar conclusions for the syntactic status of quotative and quote. She states that, despite several arguments against a grammatical status of the direct quote as complement, that

> on the whole, the syntactic properties of the French direct reporting construction indicate that […] the relationship is realized in the canonical way, i.e. a direct quote **can** be considered as the direct object of the introductory reporting clause, the latter being the matrix clause of the sentence as a whole. (*ibid*.: 290f, original emphasis).

This opposition raises one main question, which is equally valid for German and French: Do we deny quotes the status of complement of the CTP just because they are animated through bodily and prosodic means? Or do we extend our grammar by body movements and prosodic animations *in the specific case of direct reporting formats* because of linguistic material and

---

164 Also, I do have cases where the quotation is pronominalized through a cleft-construction (see Section 3.2.3, Excerpt 3.5).

bodily conduct being easily interchangeable in the grammatical role as complement of a CTP?[165]

Even though there may be several arguments in favor of De Cornulier's (1978) and Mosegaard Hansen's (2000) point of view that the direct quote constitutes the complement of the introducing quotative *se dire*, I will, as mentioned before, simply describe the pattern [quotative + quote] as a bipartite structure.

One feature that can complicate the delimitation of quotative and quote (see Section 3.1.2) are elements that are inserted between them. This is a common practice documented for direct reported speech: Such items mark the onset of the reported segment and "can be used to anchor the utterance to the 'original' interaction or situation" while allowing a speaker to "begin to display [their] stance" (Holt 1996: 238). Speakers use the following *particules d'amorces*[166] (Guérin & Moreno 2015) between *je me suis dit* and the quote:

Particule d'amorces	Frequency
mais ('but')	17.1% (7/41)
bon/ben ('well')	9.8% (4/41)
euh ('uh')	7.3% (3/41)
ouais ('yeah')	7.3% (3/41)
response cries	7.3% (3/41)
TOTAL	48.8% (20/41)

Table 23: *Particules d'amorces* between *je me suis dit* and direct reported thought

Indeed, the lexis of the *particules d'amorce* shown in this table suggests that speakers do take some kind of stance toward something prior to the reported thought: for example, *mais* ('but'), the most frequent initial item, expresses a contrast, whereas response cries like curses allow speakers to evaluate prior talk or events. Because the *particules d'amorce* are used at the beginning of the speaker's reported thought, they may "convey some relation between what the current speaker is about to say and what the previous speaker has just said" (Heritage 2002: 197), like Heritage described it for *oh*-prefaced responses to assessments.

The following analyses will show the interactional functions of *je me suis dit* when it introduces direct reported thought.

6.3.1.3.   Interactional Functions of *Je me suis dit* ('I have said to myself')

Reported thought introduced by *je me suis dit* allows speakers to publicly share a reasoning from the past. In my data, this emergent reasoning occurs in two different positions in a telling sequence. It is either used toward the end of the telling (*n* = 20) or in the middle of it (*n* = 21). Depending on the position of the reported thought, speakers either close the telling with a conclusive thought or they introduce a next part of the telling.

---

[165]   This is, of course, not valid for *j'étais là*, where the syntactic relation between quotative and quote is a different one (see Section 6.3.2.2).

[166]   Guérin and Moreno (2015) define *particules d'amorces* ('initiating particles') as follows: "Thus, the particles appear as elements of the speaker's involvement in the representation of the quoted discourse. Their function could be to guide the interpretation of the quoted speech. The initiating particles thus make it possible to announce that the following utterance will be framed as staged" (*ibid*.: 76, my translation).

In the following, I will present three excerpts comprising the social action format *je me suis dit*. The first excerpt illustrates an occurrence of *je me suis dit* at the end of a telling. It allows the speaker to make publicly available a post-hoc rationalization that led to a decision in the described situation. Cédric tells Camille how, while being at the supermarket with his mother, he spotted a package of cookies, so-called *mendiants*. Knowing that Camille likes these cookies, Cédric bought them for her. Just prior to the excerpt, Camille says that it is a pity not to have brought the cookies, which leads Cédric to initiate his story.

Ex. 6.9: mendiants[167] (Pauscaf 07, 00min59s)

```
CED: Cédric, CAM: Camille

01 CED: j'ai fait les courses seul avec ma MÈre,=
        I was shopping alone with my mother

02      =et pis du coup elle elle voulait acheter un TRUC;
        and then PART she she wanted to buy something

03      pis [j'ai vu qu'c/]
        and then I saw th-

04 CAM:     [hmhm-        ]

05 CED: pis je me souve↑nais ↓bIEn que tu m'avais parlé de
        and I remembered well that you had talked to me about

        mendiANTS; (.)
        mendiants

06      j'ai donc BIEN pensé;=
        so I thought well

07      =parce que j'ai vu les mendiANTS-=
         because I saw the mendiants

08       =je me suis DIT, %(0.4)
         I told myself
    ced                   %extends r hand->l.10
09      +mendiANTS-
         mendiants
    cam +gz at CED-->l.14

10      caM%ILLE;%$
        camille
    ced         $turns head&gz right-->l.11
          %points with index finger tw front%

11      (0.2)%(0.5)%$
    ced          -->$gz at CAM—>l.18
          %nods,purses lips%

12 CAM: ouais,=
        yeah okay

13 CED: =ça va [°BIEN $ensem$(ble).°]
        that goes well together

14 CAM:       [t'as  $fait $      ] +le BON lien.$
               you made the right connection
    cam                              -->+
    ced           -->$gz left$gz at CAM------------>>
```

---

[167]   *Mendiants* are a certain kind of cookies with nuts and chocolate.

```
15 CED: °hm.°
```

Je me suis dit allows Cédric to make publicly available an emergent reasoning that is staged as suddenly occurring in the reported situation. In the storytelling at lines 01 to 03, Cédric provides the circumstances from which the realization emerged—which are already framed as thoughts with *je me souvenais bien* ('I remembered well'; line 05) and *j'ai donc bien pensé* ('so I thought well'; line 06). At line 06, he then positively assesses his thoughts and adds an account for this assessment, which is introduced with *parce que* ('because'; line 07). The account comprises a description of spotting the cookies and then relating them (line 09) to his friend (line 10). Just after *je me suis dit*, Cédric extends his right hand toward Camille. While saying her name he points in her direction (line 10), thus emphasizing the rhematic part of the quote: Camille. Seeing the cookies is presented as the triggering event for his emergent reasoning and the resulting purchase. By nodding and pursing his lips during the pause (line 11), Cédric first displays through bodily means that Camille and the cookies "go well together," thus confirming his, already positively assessed, conclusion. This post-hoc rationalization is formatted like a logical combination of thoughts (and remembrances) which were presented as triggered by the circumstances of the past situation.

The direct reported thought is embedded in a sequence of reasoning that does not share all characteristics with a classical storytelling (see Section 3.2.1 for a differentiation). While the sequence does have a climax, the narration does not build up toward it with the typical features of an increasing affective load, bodily animations, or reported speech (Holt 1996, 2007). Instead, because Cédric narrates almost exclusively his inner monologue (except for line 01), he seems to focus on reporting the steps of his thinking process that led to the conclusion that *mendiants* and Camille are a good match. As the whole sequence is an explanation for buying the cookies, Cédric seems to stage the narration as a logical succession of events and thoughts, thus providing the interlocutor with the possibility to experience the emergence of his thoughts as if it was happening at the very moment of the conversation.

Camille shifts her gaze at her interlocutor after the short pause following *je me suis dit*, which allows her to see her friend's bodily conduct. Cédric turned his head and gazes to the right at line 09 but gazes back at Camille after the pause. He thereby projects the end of his turn (Kendon 1973; Streeck 2014). Simultaneously, his gaze toward his friend allows him to mobilize Camille's response (Stivers & Rossano 2010). She confirms Cédric's conclusion by aligning with his prior turn (line 12). Cédric then extends his turn confirming verbally his preceding embodied positive assessment.

Je me suis dit may grammatically project an empty argument position. However, it introduces—at least in the beginning—a non-clausal unit: *mendiants*, *camille* (lines 09 and 10) and a pointing gesture during the subsequent pause (line 11). Only then follows a complete clause, *ça va bien ensemble* ('that goes well together'; line 13), which is produced in low volume and in overlap with Camille's response (line 14). The format of direct reported thought enables Cédric to share with Camille his purported inner monologue. The public display of his reasoning leading to a concrete resulting action—that may also affect the co-participant— makes the recipient's response conditionally relevant (Sacks *et al.* 1974). With this specific turn design, the speaker invites the recipient to re-think with him the individual steps in the reflection

process. The lack of bodily enactment can be explained by the depicted situation not being an emotional but a rational one, which is not in need of a dramatization.

On an interactional level, the social action format *je me suis dit* is used to project the gist of the telling and thus the upcoming turn-closure. It allows the speaker to share efficiently his (sudden) reasoning behind a described situation after building up toward a climax. In Excerpt 6.9, during the pause after *je me suis dit*, Camille does not show any sign of preparing to take the turn, which demonstrates that the projection of "more to come" is successful.

The excerpt also demonstrates that recipients seem to be expected to respond to the telling—and thus to the sequence-closing direct reported thought—with an elaborate, preferably affiliative response. Indeed, in 15 of the 21 occurrences of [*je me suis dit* + direct reported thought] being used in the closing environment of tellings, recipients respond in an affiliative way. In four of the six remaining instances, speakers pursue an affiliative response by providing additional explanations. Note, however, that, compared to usual affectively loaded storytellings, recipients do not seem to affiliate with a high degree of affectivity of the direct reported thought (or the telling). Instead, they affiliate with the sequencing of thoughts that speakers stage as logical and rational step-by-step process.

Excerpt 6.10 shows a case where the recipient responds with disaffiliation. Here, the pattern [*je me suis dit* + direct reported thought] allows the speaker to make publicly available a reasoning that could potentially affect her future and is thus still "open to discussion." In this regard, it differs from the preceding excerpt, in which *je me suis dit* introduces an emergent reasoning from the past with its consequential actions lying in the past as well.

Marie complains about the exams she still must write before finishing her bachelor's degree. In this excerpt, Josiane, the recipient, has epistemic access (if not primacy over) the topic that the interlocutors discuss because she has already written the exams in question.

Ex. 6.10: trop chaud / so tricky (Pauscaf_02_05, 29min27)

```
JOS: Josiane, MAR: Marie

01 JOS: on est d'accord que LÀ- (1.0)
        we agree that now

02      euh: si tu réu$ssis tous tes exa@mens-=
        uh  if you pass all your exams
   jos                  $gz at MAR-->>
   mar                                @smiles-->l.03

03      =maint'ant il t'reste ↑quOI: en sep@↑TEMbre?
         now what do you have left in september?
   mar                                -->@

04 MAR: hehe[((laughs))                    ]

05 JOS:     [bé cé di pri comme tout l'MONde?]
               ((name of course)) like everyone?

06 MAR: <<suffering voice>£et conTRAT-£>
                          and contract

07 JOS: <<:-)<oh putain;> °hh
           oh shit

08 MAR: [<<:-)<oh puTAIN ouais.>]
         oh shit yeah
```

```
09 JOS: [ah ouais c'est CH        ]AUD:-=
            oh yeah that's tricky

10 MAR: =ouais c'est TROP [chaud.  ]
            yeah it's so tricky

11 JOS:                   [faut pas] qu'tu LOUpes;
                           you cannot screw that up

12 MAR: non.
            no

13      (0.6)

14 MAR: du coup: j'me suis dit (.)
            so      I told myself

15      [>dans  # l'<(.)+%PI:re     ] des cas% je (.) déplace UN;
            in the        worst        case   I     postpone one

16 JOS: [mais ça# ↑VA,  +%°°au fait.°°]
            but that's actually okay
    mar                      +gz at jos-->l.20
                             %3 beats l hand------>%

    fig         #fig1
```

fig1

```
17 MAR: >ouais mais le<
            yeah but the

18      (2.0)

19 MAR: %°mh.°% ((click))
   mar  %nods %

20 JOS: %((laughter, 0.4))
   mar  %extends r index, nods once-->l.20

21 MAR:   +mh.%
   mar  ->+closes eyes-->.21
            -->%

22       °h+ c'est qu'j travaille à ((name of workplace)) tout l'éTÉ,
            it's that I work at ((name of workplace)) the whole summer
    mar ->+
```

At line 03, Josiane asks an information question and subsequently delivers a candidate answer (line 05),[168] thus suggesting a repair of her initial open question to a more restrictive one (Svennevig 2013). After a laughter, Marie adds, with 'suffering prosody' a second exam, *contrat* ('contract'), thereby confirming Josiane's candidate while displaying her negative stance toward it. At line 07, Josiane affiliates with the curse *ah putain* ('oh shit'). Both

---

[168]   On the routinization of such question formats, see Pekarek Doehler (2021b).

participants are slightly laughing at the same time, which may be, in this specific case, a sign of discomfort because Marie's bachelor's degree is at stake. This is also confirmed by Josiane's admonition that Marie cannot screw this exam up (line 11). After Marie's minimal aligning response (line 12) and a 0.6 second pause, she extends her turn with initial *du coup* ('so'), implicating consequentiality (line 14; Rossari & Jayez 2000) . The sequentiality of this excerpt shows even more clearly than Excerpt 6.9 that direct reported thought with *je me suis dit* is mostly not embedded in a classic storytelling: Before Marie's direct reported thought, both speakers establish common ground without being involved in a narration.

Je me suis dit is followed by a micro pause and a reported thought that responds to Josiane's admonition: Marie may postpone one of the two exams that lie ahead of her (line 15). The direct reported thought is coupled with a so-called *ceiving* gesture (Streeck 2009: 9; see *infra* for a detailed description). Marie extends her left hand up with her palm facing her coffee cup just before *pire* ('worst'; fig1). With her extended hand, she performs, while gazing up, three beats on the subsequent three words *pire des cas* (lit. 'worst of cases'), thus emphasizing the unlikeliness to postpone one of the exams. The direct reported thought introduced by *je me suis dit* allows Marie to make publicly available her emergent reasoning concerning her upcoming exams. The temporal configuration of Excerpt 6.10 is different from the preceding excerpt; it places the reasoning that led to Marie's decision in the past while leaving open the option of not realizing the consequences. By doing so, the speaker distances herself from the decision itself while giving her interlocutor the option to discuss the open outcome—an opportunity that Josiane seizes.

Marie's direct reported thought overlaps with Josiane's disaffiliative response (line 16) which downgrades the seriousness of the situation: two exams are in fact feasible. This sudden change of stance is a dispreferred response to the interlocutors' prior common evaluation of Marie's situation as *chaud* ('tricky'; lit. 'hot'). Josiane's sudden positive assessment of the circumstances is only possible because she has epistemic access to the situation: She already wrote the exams herself. Marie's subsequent response demonstrates the disaffiliate character of Josiane's stance. Starting at line 17, Marie delivers an additional reason for the eventual non-feasibility of both exams. She works the whole summer at a summer job (line 22), which implies that she may not have enough time to study.

In contrast to the preceding excerpt, because the recipient has epistemic access to the matter, Josiane has the possibility to disaffiliate with Marie. Marie's subsequent account demonstrates that she works toward affiliation by delivering further reasons that support the, initially common, negative assessment from lines 07 to 10.

This excerpt has shown that *je me suis dit* can also be used to make available an action that may or may not be carried out in the future (*n* = 8). These cases are, however, less frequent than those making publicly available conclusive thoughts from the past (*n* = 33).

The next excerpt shows a case where *je me suis dit* occurs in the middle of a telling and allows the speaker to introduce a next part of their multi-unit turn, which comprises, most probably, new information that is necessary for the ongoing turn. The direct reported thought comes in after the speaker's affectively charged telling, to which the recipient responds in an affiliative way before the reasoning starts. Because the direct reported thought is, in this sequential position, not part of a closing activity, recipient responses are designed differently

from the previous two excerpts. In Excerpt 6.11, Elinda reports to her friend Ekti that she almost fainted during the last gym class.

Ex. 6.11: je m'asseye / I sit down (Pauscaf 20, 32min10)

```
ELI: Elinda, EKT: Ekti

01 ELI: et$ je suis <relevée> j'ai vu ↑TOUT noir-
        and I came up          and I saw all black
    eli    $both hands flat on the table-->l.06

02      j'étais là; #(1.1)#$
        I was there
    eli                     $gz at EKT-->l.06
    fig             #fig1 #
```

fig1

```
03 EKT: >ah MERde-<=
        oh shit

04 ELI: =>je me suis dit< (.) je m'asΩSEYE ouΩ j'fais quoi,=
        I told myself   do I sit down or what do I do
    eli                              Ωlifts
                                     eyebrowsΩ

05      >>mais t'sais<< tout le monde était deBOUT-
        but y'know     everyone was standing

06      >alors $$je me suis# dit<$ je vais$ pas $m'asseOI:R,
        so       I told myself    I won't sit down
    eli         $gz up------------------------>$blinks-->l.07
            -->$points with r h $........$
    fig                #fig2
```

fig2

```
07      je vais attendre °qu'° $ça reVIENNE-#
        I will wait for it to come back
    eli                     -->$gz at EKT-->l.09
    fig                             #fig3
```

fig3

```
08       si ça revient pas je m'as%(h)#SEYE-
         if it does not come back I sit down
    eli                      %smiles-->l.09
    fig                      #fig4
```


fig4

```
09 EKT: %(0.3)  % $((laughter,(0.6))
   ekt  %2 rapid
           nods%
   eli          -->$
```

```
((2 lines omitted where she elaborates how she would sit down))
```

```
12       enfin BREF,
         however

13       pi c'est reve↑NU?
         then it came back
```

Elinda describes that during the stretching exercises she saw all black (line 01). What Elinda reports with *j'étais là* (line 02) depicts her bodily and emotional state in the described situation: she freezes for 1.1 s (line 02), with both hands imitating her wide stance, accompanied by a facial expression of shock (fig1).

After Ekti's affiliative response, Elinda continues her telling with a reported thought introduced by *je me suis dit* (line 04). With her quote, she makes publicly available her emergent thoughts while trying not to faint. By doing so, she shifts from an emotionally charged depicting activity to a less depictive way of reporting that is staged as "rational" because she lays out the various steps according to which logic she has decided to act in this specific situation. The reported rhetorical question of whether she should sit down is accompanied by a lift of her eyebrows (line 04), but no further bodily enactment. After a brief change of footing to the ongoing conversation for a parenthetically inserted informing (line 05), Elinda shifts footing again and continues with a second reported thought introduced with *je me suis dit* (line 06).

This second occurrence of direct reported thought allows Elinda to make publicly available the different steps of her decision-making process from the past as they purportedly emerged:

She won't sit down (line 06), she will wait until she can see her environment again (line 07), and only if this is not the case she will sit down (line 08). That the social action format is followed by a multi-unit turn may indicate the routinization of *je me suis dit* as a device projecting "more to come"—namely, reporting a (potentially extended) reasoning. During *je me suis dit* (line 06), Elinda stages her thinking process by gazing up. The simultaneous pointing gesture (fig2) may indexically "point" at the moment of her decision. Note that the following steps of her shared reasoning are also accompanied by gestures.[169] These gestures allow the speaker to emphasize or illustrate what is verbally uttered while structuring her talk.[170]

In this sequential position, the pattern [*je me suis dit* + direct reported thought] does not introduce a consequence or final decision of a reasoning. Instead, it allows Elinda to stage parts of her narration, that are important for the understanding of her larger telling, as emergent thoughts. This is confirmed by the recipient's conduct. Compared to the recipient responses in the first two excerpts, Ekti does not react any further after her affiliative response cry at line 03 in response to the pattern [*j'étais là* + quote]. Ekti neither nods nor does she deliver signals of recipiency. Only after line 08, when Elinda displays to having finished her self-directed thought, Ekti produces two rapid nods and laughs, thus affiliating with Elinda's smile at the end of her direct reported thought (line 08). The nods display, at the same time, Ekti's access to Elinda's telling. Elinda's continuation of the turn (not in transcript) is closed with *enfin bref* ('well PART'; line 12), which clarifies that Elinda's reasoning has only been one part of her telling and that she now returns to moving toward the climax: Seeing all black did come back (line 13). Elinda's turn-continuation demonstrates retrospectively that Ekti's nods and laughter are not treated as insufficient, because Elinda continues her story with new information instead of pursuing an affiliative response.

This excerpt has shown that the sequential position of the direct reported thought seems to influence which recipient response is made conditionally relevant. While noddings in response to a turn-final direct reported thought may be treated as insufficient, Excerpt 6.11 illustrated that they are not treated as such in the middle of a turn. This is in line with previous research on storytellings: In the middle of a telling, nods are sufficient and function as continuers (Stivers 2008: 51).

Summary

The analyses have shown that direct reported thought introduced by *je me suis dit* allows speakers to make publicly available a (sudden) reasoning or rationalization in sequences that do not necessarily share many features with affective storytellings other than the triggering

---

[169]   At line 07, her extended hands with both palms down support iconically her waiting (fig3); at line 08, her lifted forearms and hands with both palms facing Ekti, emphasize the protasis of her conditional clause by conveying her 'giving up'—and consequentially sitting down—if she continues to see all black.

[170]   Without wanting to anticipate results from the next section on *j'étais là*, I would like to point out, however, that the bodily conduct following *j'étais là* (line 05) does something different: It depicts a scene where bodily conduct is the main (or even only) resource to carry out an action, which is emotionally charged. The gestures accompanying direct reported thought with *je me suis dit* do not carry the main message of the utterance.

event[171] (Labov & Waletzky 1967). The triggering event can be a description in a telling (Excerpts 6.9 and 6.11) or it can emerge from the ongoing conversation, such as in response to a co-participant's response (Excerpt 6.10). The reasoning itself can be rather short (Excerpt 6.9) or more elaborate (Excerpts 6.11). While in most cases the reasoning and the resulting consequential actions are part of the telling and do not affect the present of the ongoing conversation, in some cases, consequential actions have, however, not yet been undertaken (Excerpt 6.10). When the outcome of a rationalization is still unknown, it seems as if interlocutors can negotiate the outcome with the speaker. These cases remain, however, an exception in my data and should be further investigated with a larger data set.

I have shown in the preceding three excerpts that direct reported thought occurs mainly in two sequential positions in the telling: in the closing environment (Excerpts 6.9 and 6.10) or in the middle of a telling (Excerpt 6.11). Depending on the position in the sequence, the direct reported thought makes different responses from the co-participant relevant: affiliation (Excerpt 6.9) at the end of a telling and alignment or the display of a forthcoming affiliation in the middle of a telling (Excerpt 6.11). In this latter position, speakers usually continue their multi-unit turn after the recipient's display of alignment only.

My analysis has demonstrated that *je me suis dit* is not syntactically linked to the quote that it projects. The exact format of the quote seems to be secondary even though there is a tendency for *je me suis dit* introducing complete clauses (see Section 6.3.1.2 for quantitative evidence). Despite an absence of syntactic linkage, the quotative situates following talk as direct reported thought—a format that has been described for German by Brünner (1991). What is specific to the social action format *je me suis dit* is that speakers often stage their thinking process, through hesitations and pauses, but also through bodily means like a thinking posture.

Typically, speakers use little resources (compared to direct reported thought introduced by *j'étais là*, see following section). In all three excerpts, gestures seem to be mostly used to structure talk, thus emphasizing the rational aspect of their reasoning. These gestures belong to a certain category which Streeck (2009) describes as gestures of "*ceiving* or *caption*" (*ibid.*: 9, original emphasis). They essentially help speakers "without attending to the process and without wishing to depict anything, [to] use their hands to give form to—i.e. construe—content" (*ibid.*). 43.1% of all occurrences are accompanied by a hand gesture such as PUOH (according to Streeck (2009) a "giving" or "inviting" gesture), a pointing gesture, or an extended hand with palm down or to the side. The stroke of these gestures (Kendon 2004) is concerted with the information-relevancy within the turn: either the stroke occurs on *suis* within the *je me suis dit*-construction so that the gesture is retracted by the end of *dit*, or the stroke occurs on the new information of the following clause.

The characteristics of the social action format [*je me suis dit* + direct reported thought] seem not only tailored to the actions that speakers accomplish with it but also to the kind of sequence in which the direct reported thought occurs: Their degree of affectivity is rather low, and they do not contain reported speech of third parties or ample bodily/prosodic depictions. Instead, speakers are involved in an inner monologue that allows them to make an auto-generated

---

[171] As in two excerpts, speakers present seeing an object as a triggering event, I use this rather untypical term, thus following Jefferson's (1978) description. I am aware that "trigger" does not necessarily fit regular adjacency pairs.

thinking process available that is displayed as emerging independently from others. My analyses have shown that these kind of narrations do not necessarily correspond to the typical emotionally charged storytellings where speakers dramatize and exaggerate events to allow their recipients to (emotionally) experience past events that were—so far—outside their epistemic realm (Berger & Pekarek Doehler 2015; Günthner 1997, 2009a; Imo 2009). Instead, it seems to be the display of a rational assembly of thoughts that speakers foreground in their tellings. To my knowledge, only Bangerter *et al.* (2011) have evidenced for nursing shift handover meetings that narrations can also be used for post-hoc rationalizations in, as they call them, "institutional storytellings" (*ibid.*: 183).

The next section demonstrates that, in French, speakers seem to functionally distribute quotatives according to the degree of affectivity. In contrast to *je me suis dit*, French *j'étais là* is used by speakers to display their affective stance.

6.3.2.   French *J'étais là* ('I was there'): Displaying an Affective Stance through DirectReported Thought

The pattern *j'étais là* ('I was there') roughly compares to English *I was like*. Studies on *I was like* describe features converging with *j'étais là* like a high degree of (bodily) reenactment of what is reported (Sidnell 2006; Streeck 2002), a colloquial use (Fox Tree & Tomlinson 2008), the use during storytellings, and the introduction of *direct* reported speech or thought only (Buchstaller 2001).[172] Compared to the other investigated constructions in this book, *j'étais là* does not allow speakers to clearly display purported past *thoughts*. Instead, it allows them to leave open whether what they report has been said or thought, which can be a forceful interactional tool, especially when strong affective stance-taking is involved. The ambiguity allows speakers to first monitor the recipient's response to the strong stance-taking before—if necessary—addressing whether something has been said or thought.

Following Haakana (2007), I assume that speakers use this ambiguity interactionally. Co-participants can, just as I do as an analyst, only guess whether the quote has been uttered.[173] The ambiguity of the quotative is, however, not important for my main argument that the reported speech or thought following *j'étais là* is used as a vehicle to carry out a specific action, mostly taking an affective stance. It is thereby crucial that the social action format *j'étais là* has much in common with the other quotatives that I investigate in the present work despite its capacity to project both speech and thought. That stance-taking can be displayed through direct reported speech has already been described by Holt (1996). She analyzes reported speech as a means to "[convey] the reported speaker's state of mind" and, at the same time, to give evidence for it "more economically" (Holt 1996: 241). Put like this, *j'étais là* and the reported event are, formally speaking, a self-quotation that consists of an underspecified projector-device (*j'étais là*) and its specification (reported speech or thought). The pattern [*j'étais là* + quote] altogether carries out the action of affective stance-taking.

---

[172]   That *j'étais là* cannot—grammatically—introduce an indirect quote will be discussed in Section 6.3.2.3.

[173]   In most cases, co-participants do not topicalize this ambiguity. Excerpt 6.15 is one of the rare examples where the recipient shows through her subsequent talk that she understood the quote as direct reported thought, not talk.

J'étais là does not, in contrast to other relatively new quotatives in German or English, semantically indicate approximation.[174] Instead, it seems to refer to an imaginary space, which is about to be staged in the ongoing conversation with the deictic particle *là* ('there'). In general, patterns that introduce reported speech or thought have been much investigated in English (Buchstaller 2001; Fox Tree & Tomlinson 2008; Meehan 1991; Romaine & Lange 1991; Schourup 1982; Streeck 2002). In French, quotative patterns that occur primarily in spoken language have not been much investigated (but see Blondeau & Moreno 2018; Secova 2015; Cheshire & Secova 2018; on Canadian French, see Levey *et al.* 2013; on a general picture of reported speech in French, see Moreno 2016). In their seminal paper on Canadian French reported speech, Vincent and Dubois (1997) do not (yet) investigate quotatives like *j'étais là*, which have only been studied after 2010 (namely Cheshire & Secova 2018; Dostie 2020; Levey *et al.* 2013; Secova 2015; Secova *et al.* 2020). Recent sociolinguistic research demonstrated that *j'étais là* is a quotative that does not occur among speakers born between 1923 and 1948 while being present as a quotative among speakers born between 1995 and 2001 (Secova *et al.* 2020: 125; see also Dostie 2020: 93 for Canadian French).[175] Research on it remains scarce.[176] There have, however, been some studies, which I briefly present in the following.

Dostie (2020) has investigated *j'étais là* in Canadian French together with a few other, rather recent quotatives occurring in spoken language. Resorting to Bühler's concept of *Deixis am Phantasma* ('deixis in the imagination'; Bühler 1978 [1934]), she shows that *j'étais là* refers to a "reality broader than the one circumscribed by [it]" (Dostie 2020: 83, my translation) because it does not only introduce speech but also gestures or non-lexical vocalizations. Dostie argues that the (body) quotes are cataphorically referred to by the deictic *là* ('there'). According to her, quotatives like *j'étais là* are

> natural 'auxiliaries' for the deixis in the imagination. Thanks to them, the speaker signals that he is about to render audible and/or visible the saliant elements of a scene that he apprehends in thought. (*ibid.*: 86, my translation)

---

[174] If English *like* and German *so* really indicate approximation and do not rather function as demonstratives has been object to discussion (Buchstaller & Van Alphen 2012; Romaine & Lange 1991; Wessels 2019).

[175] This result does, of course, not necessarily mean that the quotative *j'étais là* has not existed before. Future diachronic research may reveal that *j'étais là* has only disappeared for a certain period of time before being reused again more recently.

[176] There are, however, formal and interactional features of *j'étais là* that converge with English *(to) be like* or *(to) go* as well as with the recent quotative *this is + speaker* (Cheshire *et al.* 2011; Secova 2015). The research body concerning these three English quotatives has been growing since the 1980s. In one of the earliest papers on the topic, Butters (1980) describes *(to) go* in narratives. He states that, compared to *(to) say*, *(to) go* does not have to report elaborated speech but also non-lexical vocalizations (such as animal sounds) or "the mimicking of bodily actions, gestures, and postures" and concludes that it "has a somewhat parader range than mere *say*" (*ibid.*: 305). Fox Tree and Tomlinson (2008) show similar characteristics for *like* as a quotative. The quantitative results of their corpus study show that between 1980 and 2000, *like* has exceeded *go* by far (4% vs. 11% in 1980 compared to 92% vs. 2% in 2000). This development goes hand in hand with an extreme drop of the use of *say* as an oral quotative, which declines from 85% to only 7% (*ibid.*: 92). One important point Fox Tree and Tomlinson (2008) make is that *like* does not just replace *say*. They argue, instead, that the different enquoting devices code different meanings in talk (*ibid.*: 93). This hypothesis seems to be also valid for my French data, where *je me suis dit* is used for other interactional purposes than *j'étais là*. The authors also show that there has been a shift of tasks speakers carry out with *like* over time. While in 1980 other's speech or thought is reported with *like*, in 2000 speakers' have a higher tendency to quote themselves with *like* (*ibid.*). These results indicate that the distribution of quotatives may change over time and that such a change may be interactionally motivated.

She shows that more "traditional" verbs like *dire* ('to say') and verbs of movement like *sortir* ('leave') quote speech whereas quotatives like *j'étais là* refer to body movements or sounds. This observation is specifically relevant for the recipients' orientation to the speaker's quote. If *j'étais là* does introduce bodily conduct recurrently, this may be observable through a systematic recipient gaze at the speaker.

Cheshire and Secova (2018) found out for Parisian French that, in contrast to *être comme ça* ('to be like that'), *être là* ('to be there') is the preferred quotative when the quote consists of "non-lexicalised sounds or internal thought, and when the quote is uttered with mimesis" (*ibid*.: 27)—a result that has been later confirmed by Dostie (2020) for *j'étais là* in Canadian French. What Cheshire and Secova (2018) name *mimesis* is similar to Clark and Gerrig's (1990) concept of *depiction*.[177] Comparing *j'étais là* to *this is*, Cheshire (2018) shows that both constructions also allow speakers to "'[perform]' the appearance, location, action or activities of a protagonist in a narrative" (*ibid*.: 29). While stating that *j'étais là* is systematically followed by a bodily reenactment, none of the three articles uses detailed transcripts to evidence this observation. Drawing on the methodology of multimodal Interactional Linguistics, I will trace through sequential analyses which resources speakers use before, while and after reporting thought or speech with *j'étais là*.

None of the above-mentioned articles uses the methodology of CA or IL. Consequently, they do not investigate how *j'étais là* is used sequentially and which actions speakers carry out when using it. My interactional approach allows me to address, in addition to prior findings, the following questions: What leads to the use of the (self-)quotation with *j'étais là*? Why do speakers use the social action format *j'étais là* at a specific point in their telling? What differentiates *j'étais là* from *je me suis dit*? Approaching *j'étais là* with the conversation analytic question of *why that now?* (Schegloff & Sacks 1973: 299) allows me to complete previous pragmatic and sociolinguistic findings. An additional multimodal analysis enables me to investigate further the "mimetic nature" (see *supra*) of the quotes introduced by *j'étais là*. I will also demonstrate that the recipients' orientation to these quotes is crucial for the understanding of the pattern [*j'étais là* + quote],[178] which builds on a deictic particle that mostly refers to a reenacted depiction through bodily means.

6.3.2.1. A General Picture: A Quantifying Analysis

J'étais là ('I was there') occurs 60 times in my 11 h 18 min video corpus. Three occurrences had to be singled out because of *là* referring to an actual locus, thus excluding the possibility of *j'étais là* projecting a quote.

Out of 57 occurrences, only eight are turn-initial. The notion of 'turn-initial' includes turns where *j'étais là* is preceded by temporal adverbials like *puis* ('then') or *l'autre jour* ('the other day'), by the conjunction *parce que* ('because') or by a click (*n* = 4). Forty-nine occurrences are turn-internal. Within the turn, however, 37 tokens constitute the beginning of a new

---

[177]   For a detailed explanation of the concept, see Section 3.1.3.
[178]   In contrast to the other patterns treated in this work, I refer to this pattern as [*j'étais là* + quote] and not as [*j'étais là* + direct reported thought] because the speaker does not have the possibility to clearly display direct reported *thought* with this quotative.

intonation unit and a new TCU,[179] meaning that they introduce a new action within a multi-unit turn from the same speaker (see also Section 5.3; for a description and discussion of the segmentation unit 'TCU', see Clayman 2013; Ford & Fox 2002; specifically for German, see Selting 2000; for French, see Persson 2017: 34ff). Compared to *je me suis dit*, *j'étais là* is, in this TCU-initial position, less often preceded by lexical items (*n* = 13); None of these items occurs more often than 3 times.

The mid-turn position of quotations with *j'étais là* is not surprising, given that direct reported thought and speech usually occur during storytellings. A feature of storytellings is that the same speaker keeps their turn for a long time while receiving primarily backchannels from the recipients (for a detailed description of storytellings, see Section 3.2.1). Telling a story allows the speaker to extend (or not extend) their telling further in time (Auer 1996), thus reacting to the emergent contingencies of the ongoing conversation (e.g., because of the recipient's change of facial expression or a comparable action). 10 further occurrences of self-quotations with *j'étais là* are not only turn-internal but also intonation unit-internal.

Before analyzing sequentially individual excerpts, I will first present some general syntactic features of the pattern [*j'étais là* + quote].

6.3.2.2. Syntactic Features of *J'étais là* ('I was there')

Mosegaard Hansen (2000) states regarding direct reported speech in general that "the direct quote can be considered as the direct object of the introductory reporting clause, the latter being the matrix clause of the sentence as a whole" (*ibid*.: 291). This explanation, however, does not hold for *j'étais là*, which does not call for a direct object. *J'étais* ('I was') is syntactically complete with the addition of the deictic *là* ('there'). The deictic remains, however, underspecified; with *là* ('there') referring to a location or moment that has not yet been specified, the clause projects a specification. In this sense, the pattern [*j'étais là* + quote] is a bipartite structure *par excellence* as grammatical projection is not at work. Despite this difference to the other social action formats projecting a quote that I have investigated so far, the pattern functions similarly in interaction: It is systematically followed by a quote, the bipartite structure emerges in time according to the contingencies of the ongoing interaction, and *j'étais là* seems to be routinely used for the same action. This raises the question how *j'étais là* and the quote are actually linked.

Spronck and Nikitina (2019), establishing three main components for quotatives, with one of them being that there is a "deictic relation between […] the alleged original situation of discourse production […] and the current speech moment" (Spronck & Nikitina 2019: 143).[180] Similarly, De Cornulier (1978) already points out that a reenacted direct quote "can be explicitly announced and designated" (*ibid*.: 83, my translation), for example through a deictic; However, the direct quote "does not enter, just because of this one reason, into the grammatical syntax of

---

[179]   Because I transcribed according to the GAT2 transcription system, my transcripts are segmented in intonation units, not in TCUs (for a detailed explanation, see Section 5.3). However, because direct reported thought carries out an action on its own, it does happen, in my data, that an intonation unit describes the same unit as a TCU.

[180]   Picking up Bühler's concept of *origo* (Bühler 1978 [1934]), one could also argue that *là* is a deictic to shift the origo into what he calls *Phantasma* (translated by Stukenbrock (2014) as 'imagination'). A detailed consideration of Bühler's model is unfortunately beyond the scope of the present work.

the discourse" (*ibid*., my translation). It is also for this reason that I conceptualize [quotative + quote] as a bipartite structure. Authier-Revuz (2019) calls formats like *j'étais là annonceur* (*ibid*.: 205ff; 'announcing device'). An *annonceur* distinguishes itself from an *introducteur* ('introducing device') in that it cannot be pronominalized and that it cannot be reformulated into a passive construction (*ibid*.: 206). According to her, this explicitly holds for quotatives including the speaker's multimodal conduct (*ibid*.: 206f). These observations apply to reported thought with *j'étais là.* With *là* referring to the speaker's (visually perceivable) state at a specific moment in time, the quotative invites the recipient not only to relive the demonstrated scene but also to orient to the multimodal character of the introduced speech or thought.

In my data, speakers depict their state or action from a specific time or place that is deictically referred to with *là* ('there'). By using several bodily resources other than speech, speakers do not only invite the recipient's auditory attention but also their visual one. In this, *j'étais là* is a quotative that does not only introduce "body quotes" (Streeck 2002), but also invites a bodily orientation from the recipient.

In general, in addition to the most frequent format of a complete clause ($n = 25$), the quote following *j'étais là* can have various forms, as the following table shows:

Quotative	Response cries only	Bodily conduct only	Verbless structures	Question (words)	Imperatives
j'étais là	14.3% (8/57)	5.2% (3/57)	17.5% (10/57)[181]	10.5% (6/57)	7.0% (4/57)

Table 24: Syntactic features of direct reported thought introduced by *j'étais là*

The remaining instance is a quote that has the shape of a multi-unit turn—a format that is more frequent with *je me suis dit* ($n = 5$). These numbers suggest that *je me suis dit* has a higher tendency to introduce longer quotes than *j'étais là*, which seems to deliver briefer, emotional utterances. Of course, a larger data set is needed to confirm this assumption.

The next section presents the functioning of use of *j'étais là* in everyday talk. My analyses will show that *j'étais là* recurrently allows speakers to display their affective stance in an economic way, not only linguistically, but also through embodied means.

6.3.2.3.   Interactional Functions of *J'étais là* ('I was there')

In the following sequential analysis, I will show that the pattern [*j'étais là* + quote] allows speakers to display their affective stance that is conveyed through a multitude of resources. The analysis also demonstrates that the displayed affective stance, which occurs toward the end of tellings, makes a response from the co-participants relevant. If this response is not delivered, it is pursued by the speaker. I will first analyze two excerpts with an affiliative response from the co-participants and one excerpt in which the interlocutor disaffiliates with the speaker's strong stance. An additional excerpt illustrates how a recipient can demonstrate their understanding of the quote introduced by *j'étais là* as reported thought, not speech.

---

[181]   3 of these 10 verbless structures comprises only the response particle *non* ('no').

In Excerpt 6.12, Elinda is involved in a longer telling with a constructed dialogue. At the end of the telling, she comments on the described circumstances with [*j'étais là* + quote]. Elinda talks about a weekend where she planned to study but participated in the production of a video clip. She expected this event to take place on another weekend because she hadn't been informed about anything until Friday, when she received a message from the organizer, who asks all girls not to wash their hair. The following excerpt is the beginning of Elinda's depictive telling on how she found out that the shooting took place that weekend:

Ex. 6.12: coiffure / hair style (Pauscaf 20, 23min05)

```
ELI: Elinda, EKT: Ekti

01 ELI:      [genre] euh::: le vendreDI, (0.2)
                  like   uh      on friday

02      maTIN? (0.3)
        morning

03      je reçois un message de la FILle- (0.3)
        I get a message from the girl

04      ↑↑oh cou↑COU ↓euh si jamais les ↑FILles_↓euh (0.5) euhm::::-
          oh hi         uh  just in case girls     uh        uhm

05      $°attends qu'est-ce qu'elle m'a-
         wait     what did she
   eli  $gz&head down-->l.07

06      comment est-ce qu'elle m'a sorti ÇA,°
        how did she tell me that?

07      si jamais $les +FILles euh: juste une$ un petit consEIl$ euh-
        just in case girls     uh   just a   a little advice  uh
   eli           -->$gz at EKT---------------->$gz left-------->$gz at EKT
                                                                  -->l.08
   ekt              +gz at ELI-->>

08      pour euh: ce$ week-END-% (0.4)%
        for this weekend
   eli           -->$gz left-->l.10
   ekt                         %tucks chin, blinks twice%

09      'fin pour deMAIN, (0.4)
        well for tomorrow

10       $si vous pouvez ne pas vous laver les cheveux
          if you  could  not       wash your hair
   eli  ->$gz at EKT-->l.12

        le jour# MÊme%# euh-
        the same day
   ekt                %opens eyes wide,eyebrows up-->l.12
   fig         #fig1a #fig1b
```
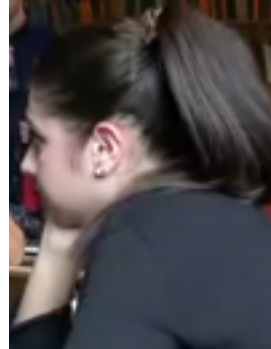
fig1a   fig1b

```
11      qu'on puisse faire les coiffures beaucoup plus facileMENT,
        so that we can make the hair styling more easily

12      en$fin bref,%
        well anyway
   eli  ->$gz left-->l.14
   ekt          -->%

13 EKT: ils t'ont [fait QUOI?      ]
            what did they do? ((with the hair))

14 ELI:          [donc je- $>donc] $du coup j'tais LÀ;< $(.)
                   so I-   so     PART I was there
   eli                    $lifts hands tw ears-->l.15
                             -->$gz at EKT---------->$gz up
                                                      ->l.16

15 ELI: °h$ #ça me Ωconcerne Ω# ↑PA:S,=
        that's not relevant for me
   eli  ->$lifted hands, palms up-->l.18
                Ωeyebrows upΩ
   fig       #fig2a          #fig2b
```
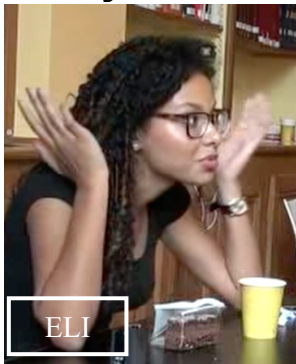

fig2a   fig2b

```
16         =$avec ma Ωtouffe de [cheVEUXΩ d't'façon]&
             with my tuft of hair
   eli  ->$gz at EKT-->l.18

17 EKT:                     [<↓mais Ω ↑ouAIS:,>]
                              of course
   eli          Ωeyebrows up       Ω

18 ELI: &tu peux faire$ ce$ que tu veux# propres ou SALES,=
        you can do whatever    you want clean or dirty
   eli          -->$gz left-->l.21
                    -->$moves hands l&r in the air-->l.21
   fig                          #fig3
```

fig3

```
19 ELI: =hh° huhu [hahahahaha °HH     ]
20 ELI:          [<<:-)>enfin voilà,>]
                  PART PART
21       °hh$§ <<:-)>et pis euh: attends-> [hh°]
         and then uh wait
   eli   -->$
          -->§
```

At lines 01 to 12, Elinda repeats the message that the organizer sent to her and the other girls. The content of the message is introduced as a written message (line 03). However, Elinda reenacts the organizer's "speech" with a high pitch by mimicking a female voice. During the quote, Ekti reacts only through bodily means: At line 08, she tucks her chin and blinks twice. This display of surprise is probably due to her realizing how short notice the organizer's message has been. Ekti's bodily response follows immediately *ce weekend* ('this weekend'), which is in relation to Elinda's introductory informing *le vendredi matin je reçois un message* ('on friday morning I receive a message'), obviously being short notice for an event on the weekend. A second minimal response follows at line 10, where Ekti opens her eyes wide and lifts her eyebrows (fig 1a and fig1b), thus displaying—again—surprise, this time toward the organizer's request of not washing the hair the same day.

Elinda has just closed her first part of the telling with *enfin bref* ('well anyway') at line 12, thus cutting the third party's reported speech short and proceeding with her response to this reported first pair part. The response is displayed as a quote, which allows her to take an affective stance toward the organizer's request. That the quote is a case of direct reported thought becomes clear through the circumstance that the "first pair part" to which the quote responds within the telling is a written message. At line 14, she introduces her response with *donc du coup* (lit. 'so therefore'), thus marking it as consequential in the order of events in the telling. *J'étais là* forms one chunk with the two preceding items: *donc du coup j'étais là* is pronounced with accelerated speech and under one intonation contour. Additionally, *j'étais là* is morpho-phonologically reduced to *j'tais là* (line 14). These features, accelerating the telling, come to a sudden halt with the micro-pause (line 14) and the subsequent inhale (line 15). That more talk (and/or body movements) is going to follow is, however, projected by Elinda's gaze up just before the micro-pause (line 14). Being part of the depiction of her negative affective stance, her reenacted gaze may project the continuation of her turn. Indeed, the speaker's reenacted direct reported thought follows: The organizer's request is not relevant for Elinda.

This reported thought continues to be enacted: Elinda lifts her hands, her palms facing up, and gazes into the air (fig2). Her embodied conduct depicts the message that she verbally conveys: This information is not important for her because of her type of hair (she has very

curly hair). While the verbal message remains a simple declarative, the prosodic features and Elinda's gestures and facial expression allow her to display her affective stance through bodily means, thus assessing the organizer's last-minute request as unnecessary. The direct reported thought allows Elinda to embed her affective stance into the telling by framing her negative assessment of the organizer's request as part of the story world. By doing so, the speaker may provide guidance to her recipient as to which stance she is expected to affiliate with.

Indeed, Ekti affiliates with Elinda's stance: First, she utters a *mais oui* ('of course') in a very low pitch and with slight laughter (line 17), thus conveying that what Elinda said is an evident fact. Then, Ekti reacts to Elinda's turn extension at line 19 with laughter, thus orienting to Elinda's story as amusing. Ekti's display of affiliation overlaps with Elinda's closure of the explanatory side sequence with *enfin voilà* (for the closing function of 'voilà' in side sequences, see Hailaselassie 2015). The continuative final intonation, however, projects more to come and Elinda indeed continues her telling, as her directive at line 21 demonstrates.

The next excerpt illustrates that affiliation can be displayed by joining the depicted interaction, a phenomenon that Cantarutti (2020) describes as co-animation: The current interlocutor takes over the role of the speaker. In this specific case, the quote introduced by *j'étais là* introduces an ironic quote that hovers between thought and speech.

Prior to this excerpt, Frank shows Mathilde a photo of one of his watches on his cell phone, which she negatively assesses as "disgusting." Frank then defends the watch in a long turn by explaining that it is his grandfather's old watch, which he gave to him. At line 01, Mathilde responds to Frank's justification and continues her negative assessment:

Ex. 6.13: merci papi / thank you grandpa (Pauscaf_02_17, 02min54s)

```
MAT: Mathilde, FRA: Frank

01 MAT: [ouais mais] MÊ:+me-
         yeah but even that
02      fin t- tu VOIS le déjà d- des dia+mants à la place des
         well y you see the first diamonds instead of the
        inDEXES,+
        hands
03      (1.2)
04 MAT: °c'est° pas$ très$ mascu[LIN?      ]
         it's not very masculine
   mat                   $gz at FRA-->l.07
   fra           §closes bottle with r hand-->l.25
05 FRA:                        [°non mais°] c'est dégueuLASSE,
                                 no but    it's disgusting
06      (0.3)
07      nON+$ c'est dégueuLASSE,
        no it's disgusting
   fra    +gz at MAT-->l.09
   mat  -->$gz at cell phone-->l.10
08      (0.3)
09      >non chuis d'a+cc§↑ORD,<=
         no I agree
```

```
fra                 -->+gz at his cell in MAT's hands-->l.10
                      -->§reaches fw with r hand-->l.10
10      =moi j'l'ai r'çue§ j'tais là +↑m::$:,# ((click))+ (0.2)
          when I  received it   I was there m:::
fra                 -->§puts bottle on table w l hand,
                        puts imaginary object with r hand into
                    l hand with open palm-->l.11
                              -->+gz at his l h---+gz at MAT
                                                  -->l.14
mat                           -->$gz at FRA-->l.13
fig                                   #fig1a,b
```


fig1a   fig1b

```
11      su§per mer%CI% tu vois-§
          great thanks you know
fra    ->§closes l hand------>§opens both palms-->>
mat          %lifts eyebrows%
12      [bah ça m'faisait  %ça m'$fai     %]sait&
          well it made me it made me
13 MAT: [((click)) merci pa%PI;  $ £hh°_in£%]
                  thanks grandpa
mat                     -->$gz at cell phone-->>
                        %lifts eyebrows%
14 FRA: &plaisIr+ parce que c'est sa MONtre-
          happy    because it's his watch
fra        -->+gz down-->>
```

Mathilde responds to Frank's prior justifications (not in transcript) with a dispreferred turn by stating that the design of the watch is not very masculine (lines 01–04). Frank affiliates with this opinion by uttering twice the same negative assessment (lines 05 and 07). At line 09, he then explicitly agrees with Mathilde's aesthetic judgment. Note that the initial *non* ('no') preceding the agreement is an affiliative one (Sandager Sørensen 2021: 108f). The next intonation unit (line 10), attached to the preceding one with latching, initiates the storytelling *in medias res* and can be divided into two parts: a short description, *moi j'l'ai reçue* ('I received it'), and then the immediate reported response to the described event of receiving the watch, introduced with *j'étais là*.

First, the speaker reports a lengthened *m* with a high pitch movement, mimicking the intonation contour of surprise (Selting 1996b), and a click (line 10), which has been demonstrated to be regularly involved in affective stance-taking (Ogden 2018). Verbally, Frank first ironically assesses receiving the gift with a positive *super* ('great'; line 11) and thanks his

grandfather with *merci* ('thanks'). The subsequent *tu vois* ('you see') is preceded by a shift in footing as it addresses Mathilde in the ongoing conversation. *Tu vois* ('you see') may be a way to mobilize her response to Frank's ironic turn, especially since he also shifts his gaze to Mathilde just before (line 10; Stoenica & Fiedler 2021).

As thanking someone is the preferred relevant next after receiving a gift, *j'étais là* could introduce, in this case, direct reported speech. However, the reported speech is blended with irony which displays Frank's (probably unsaid) negative stance toward the gift. Consequently, while the thanking-turn may be analyzed as reported speech, the negative stance toward the gift may be analyzed as reported thought. Alternatively, the irony could be analyzed as not reported at all but as an addition to the ongoing storytelling to display an affective stance.

The ironic character[182] of Frank's response is primarily conveyed through bodily means: He depicts how he takes the watch with his left hand, presses his lips together to a forced smile (fig1a and fig1b), and gazes at the imaginary object while maintaining this holding gesture until the beginning of Mathilde's turn (line 13). With an exaggerated high-pitched movement, his bodily conduct allows him to display his negative stance toward the watch, which would be a highly dispreferred response to receiving a present.

This excerpt illustrates how bodily and verbal conduct are laminated (Goodwin 2013), thus allowing the speaker to convey incongruent stances: while Frank verbally reenacts a preferred response to his grandfather's gift (thanking and positively assessing the gift), his bodily and prosodic conduct demonstrate his negative stance toward the watch. This division of labor between layered resources is recurrent with the pattern [*j'étais là* + quote] and stresses the importance of analyzing not only the verbal, but also the multimodal layers of quotation.

In overlap with Frank's turn-continuation (line 12), Mathilde joins the staging of the scene and takes Frank's role, producing another thanking turn toward the grandfather (line 13). Through this repetition, she displays her affiliation by "displaying the addressee's [here: Frank's, SF] emotional stance toward a given utterance" (Svennevig 2004: 508) or action, like the grandfather's handing Frank the watch. Mathilde's prosody displays a similarly ironic thanking as Frank's self-quotation did. Note that Mathilde also produces a turn-initial click, parallel to Frank (line 10), which contributes to her mimicking turn having a "central evaluative component" (*ibid*.). Additionally, she lifts her eyebrows and closes her turn with laughter, thus orienting to Frank's reported scene as non-serious (Goffman 1981). In response to Mathilde's co-animation, Frank continues his telling with a next part: How he was happy about the gift because the watch was also something special for his grandfather (lines 12 and 14).

This excerpt shows (other than the lamination of multimodal resources) that another way of affiliating with the speaker's stance displayed through the pattern [*j'étais là* + quote] is to join the reenactment itself and take over one of the roles in the story. However, in some cases, recipients do not affiliate with the speaker's stance. The next excerpt shows how recipients disaffiliate with the speaker's reported thought.

Excerpt 6.14 illustrates one way of disaffiliation in response to the pattern [*j'étais là* + quote]: disattending a complaint (Mandelbaum 1991). Joana is telling her friend Ekta about a

---

[182]   I understand irony following Kotthoff (1998) as "doubly-voiced speech that transmits a cleft of evaluative perspectives as its main message" (*ibid*.: 1).

mutual acquaintance, Kelly. Prior to this excerpt, the two friends have already discussed Kelly apparently becoming arrogant since working as DJane. In the following excerpt, Joana delivers, through a constructed dialogue, a further example for why this fellow student has become smug. Prior to this excerpt, Joana describes that she saw Kelly at a party one year ago, which is where the reported exchange took place.

Ex. 6.14: trop de sport / too much sports (Pauscaf 05, 04min35)

```
JOA: Joana, EKT: Ekta

01 JOA: puis elle m'a DIT^euh (0.6)
        and then she told me uh

02      j'ai dit ↑ah t'as perdu du POIDS >tout ça<,=
        I said   oh you lost weight       etcetera

03      =>pi elle m'fait< ouAI:s je fais TROP de sport mais-
        then she was like yeah I work out too much    but

04      (0.3) elle a pas énormément perDU,
            she didn't lose a lot

05      parce qu'entre le lycée et pis mainteNANT, (0.3)
        because between college  and   now

06      §y a comme quoi qu/ maximum cing +ki↑LOS tu vois?§
        it's like what   maximum five kilos you know?
   ekt  §shakes her head 'no'-------------------------->§
                                    +gz at her coffee->l.08

07      [°pis + elle m-°]

08 EKT: [>mais+ elle a  ] pas le §TEM:PS;<§
          but she doesn't have time
   ekt     -->+gz at joa-->l.15
                            §lifts eyebrows§

09 JOA: >pis elle me fait<,=
        and then she goes

10      =ouais je fais trop de sport avec la perRINE,
        yeah I work out so much with      perrine

11      %on §va: au ma§chin maCHIN-%=
        we go to the thingy thingy
   joa  %moves torso left&right--->%
            §gz up--->§gz at ekt-->l.12

12      =#>ch'tais là< §°*eu#::((creaky))*°§%(0.5)%
          I was there uh
   joa            -->§closes eyes------->§
                                    %drops r hand%

   fig  #fig1                 #fig2
```

fig1

fig2

```
13        °°$juste c'est BON alors;$#°°
            just  okay then PART
          $rolls eyes---------->$gz right--->l.14
   fig                           #fig3
```


fig3

```
14        (1.3)$
   joa  ---->$gz at EKT-->l.17

15 EKT: $perrine la copine à ludoVIC?+
        perrine ludovic's girlfriend?
                            -->+
   joa  $lifts r hand, puts chocolate in her mouth-->l.16

16        (0.4)$
   joa      —>$

17 JOA: EX  (0.2)$ copine à ludovic.
        ludovic's  ex  girlfriend

   joa          -->$

18 EKT: ils sont plus enSEMble?
        they are not together anymore?

19        (0.3)

20 JOA: non: (0.2) c'est (fini.)
        no        it's over
```

Joana starts reporting a dialogue between her and the fellow student, Kelly (lines 01 to 13). Prosodic imitation (lines 03), an exaggerated stress on *trop* ('so much'; line 03) and a gaze up allow Joana to display her negative stance toward Kelly's response. Subsequently, Joana assesses Kelly's reported conduct negatively by saying that she actually cannot have lost that much weight (lines 04–06). After an aligning headshake at line 06, Ekta affiliates by displaying shared knowledge about Kelly (line 08) with *mais elle a pas le temps* ('but she doesn't have the time'), thus supporting Joana's negative assessment.

Without responding to Ekta's further argument, Joana continues her telling by repeating Kelly's response to her initial question (line 10). Again, Joana's embodied conduct displays her negative stance: During *on va au machin machin* ('we go to the thingy thingy'), Joana moves her torso left and right as if to mimic someone who wants to show off and gazes briefly up (lines 10 and 11) before gazing back at her interlocutor.

The reported stretch of talk is immediately followed by *j'étais là*, which works as a projector-construction for Joana's response to Kelly's utterance. I argue that this instance of [*j'étais là* + quote] reports thought, which concludes the reported (and partially repeated) dialogue with an assessment, as it is typically the case in the evaluation phase of a storytelling (Labov & Waletzky 1967). Joana thus frames her quote as something that she has thought in response to Kelly's boasting. Because the reported scene is part of an exemplification that Joana packages as storytelling, I analyze her quote as being actually reported and not as an emerging stance-taking in the ongoing telling, that she simply formats as thought. One reason for this analysis is that both interlocutors have already established their shared negative stance toward Kelly *before* this excerpt. It may thus not be necessary, for Joana, to distance herself from her negative stance through the format of quotation.

In contrast to the preceding excerpts, where bodily and verbal conduct were produced simultaneously, they occur, in this excerpt, one after another: First, Joana displays her stance through a bodily and prosodic depiction. She closes her eyes and lifts her eyebrows (fig2), produces an exaggerated creaky sound, and drops her right hand on the table. This multimodal assembly allows Joana to foreshadow early on her negative affective stance, thus indicating already to her interlocutor which stance she may need to affiliate with. At line 13, Joana then verbalizes her negative stance and reports her (verbal) response to Kelly's utterance. The formulaic expression *juste c'est bon* ('just it's fine') is accompanied by Joana rolling her eyes (fig3), thus displaying being annoyed. The utterance is produced very quietly and appears like an afterthought followed by a longer pause (line 14), at the end of which Joana gazes back at her interlocutor, thus handing over the turn to Ekta also through bodily means (Kendon 1967). Incipient sequence closure is also displayed by Joana lifting her hand and starting to eat the chocolate bar that she has been holding in her hand during almost the whole telling (starting four lines before this excerpt).

It is only after Joana's display of her strong negative stance with [*j'étais là* + quote] that Ekta reacts more elaborately. Ekta asks an information question about Kelly's friend who is mentioned at line 10. Instead of reacting to Joana's strong affective stance, which strongly projects a response (on the suggestion of a continuum of projective force, see Stivers & Rossano 2010), preferably an affiliation, Ekta topicalizes a detail of Joana's telling, a practice that has been described by Mandelbaum (1991) as *disattending complaints*. Mandelbaum argues that the non-uptake of a complaint happens when the complaint is not formulated explicitly but inferred. This may also apply to Excerpt 6.15, where Joana's bodily conduct is the primary indicator of her affective stance, which may not be sufficient for Ekta to affiliate with. In this excerpt, the social action format *j'étais là* introduces *first* a bodily conduct allowing the speaker to display her negative affective stance, and *only then* its verbalization.

The last excerpt will not be analyzed in detail. It shall simply illustrate that co-participants may demonstrates their understanding of the pattern [*j'étais là* + quote] as direct reported

thought. In a depictive storytelling, Frank tells a friend, Mathilde, why it was impossible for him to concentrate the last time he tried to study at the university library: The student sitting next to him took too much space, extending his elbows left and right, and made too much noise by constantly opening and closing his highlighters. The excerpt starts when Frank depicts the fellow student opening and closing his highlighters (line 01) and the fellow student's body posture (line 02).

Ex. 6.15: son coude / his elbow (Pauscaf_02_17, 10min32s)

```
FRA: Frank, MAT: Mathilde

01 FRA: >donc ça faisait< clic clic clic $CLIC-%
        so it made click click click click
   fra                                    -->%moves l elbow
                                              left->l.04
   mat                                   $leans fw,head down,
                                          moves left
                                          hands tw mouth-->l.04

02      pis il [avait son cou$de] comme ÇA;
        and he had his elbow like that
   fra                     $gz at MAT-->l.07

03 MAT:        [h↑IN           ]

04 FRA: pis deux$ trois ↑fois j'ai %tapé son ↑COUde,=
        and two three times I tapped his elbow
   mat        -->§
                             -->%taps in the air w r elbow-->l.05

05      =>pis %j'tais °l°à< mais MEC-+
        and I was there but man
   fra     -->%moves both elbows back to middle-->l.06
   mat                            +gz at FRA-->l.06

06      t'%as ta #place là à: DROIte donc eu:h-+
        you have your space there on the right so uh
   fra ->%reaches hand wide to right-->l.21
   mat                            -->+gz at her cookie
                                        -->l.07

   fig              #fig1
```


fig1

```
07      limite:^euh si on est tro:p^$euh coince$ cOINce+ ben:
        PART  uh    if we are   too    squeeze squeeze   well then
   fra                           -->$gz down---->$gz at MAT-->l.08
   mat                                    -->+gz at FRA
                                              -->l.10

        (0.5)%(0.7) dé↑RIVE.*%
                move
```

```
    fra    -->%depicts moving
             something to the r
             w hands------->%both h on table-->l.10
                          *smiles-->l.08
08      (1.4)*$(0.2)$
   fra    -->*
              -->$gz at his bottle-->l.10
   mat       $shakes with laughter-->l.09
09 MAT: £HH°£$
   mat    -->$

((15 lines omitted))

25 MAT: mais ouais mais t'aurais d- t'aurais dû dire quelque CHOSE,
        but yeah but you should have you should have said something
26      (2.4)
27 FRA: [ouais.  ]
         yeah
28 MAT: [t'aurais] dû dire excuse-↑moi: tes stabi↑lo:s tu peux faire
          you should have said sorry your highlighters can you make

        moins d'BRUIT?
        less noise
29      (0.5)
30 FRA: ouais.
        yeah
```

In lines 01 to 07, Frank complains about a fellow student sitting next to him in the library and taking up too much space with his elbows. He reenacts the scene by gesturing with his right elbow as if he was tapping the student next to him while narrating what happened (lines 04 to 05). At line 05, Frank introduces a direct reported thought with *j'étais là*. The quote reports what he purportedly thought in the moment of the described scene. With his right hand, Frank points toward the spot where the fellow student sat (fig1), subsequently reenacting his (imaginary) directive reminding the student that he has enough space on the right side (line 06) and that he should move (line 08). The direct reported thought format allows Frank to display his strong affective stance as part of a telling by reporting an *imagined* response in a situation of potential conflict (for proof, see below).

After a minimal laughter from Mathilde (lines 08 and 09), Frank describes that finally the fellow student left, thus closing the sequence. Until then, Mathilde has only reacted with (silent) laughter (lines 01 to 03, and line 09). At line 08, her response has been strongly projected but remains noticeably absent, which may be why Frank continues his telling. After another short side sequence (omitted lines), Mathilde reacts in a more elaborate way. She disaffiliates by telling Frank that he should have said something (lines 25–29), thus displaying her recipient understanding of Frank's self-quotation with *j'étais là* as thought, not talk. Mathilde being able to infer that Frank's response toward the fellow student's behavior remained unspoken clarifies that recipients draw on sequential, bodily, and linguistic information emerging through turns-at-talk to make meaning of everyday conversation.

Summary

In this section, I have investigated *j'étais là* as a hybrid quotative projecting direct reported thought or speech. My analysis has shown that the quote projected by the social action format allows speakers to take an affective stance in response to something that has been reported or described before. Forty-five tokens out of 57 are in some way responsive to a preceding action which the speaker reports prior to *j'étais là*. This coincides with what Goffman (1974) conceptualizes in his seminal book *Frame analysis* as 'minor reflexive frame breaks'. He defines them as follows:

> [T]he speaker often finds cause for minor reflexive frame breaks, turning to his own just finished verbal behavior as something to which he now directs exegetical or apologetic asides. Such self-generated, self-referential, inwardly spiraling grounds for response are necessarily somewhat cut off from the ongoing *inter*action, for here the actor all on his own provides at one moment the response to which he himself reacts at the next. (*ibid*.: 502, original emphasis)

What Goffman (1974) describes is a short change in the narrative stream from the speaker's telling activity to their commenting on a reported (or, in Goffman's terms, 'replayed' event). Speakers can thus break with the frame "telling," return to the frame of the ongoing interaction and utter something similar to a meta-comment on what has been narrated before.

Sequentially, *j'étais là* occurs often toward the end of tellings[183] and seems to make a recipient's response, and preferably an affiliative one, relevant. In 32 cases, the speaker ends their telling after the turn with *j'étais là* and the recipient's response follows. In 12 further cases, the recipients' conduct demonstrates that turns with *j'étais là* display an upcoming transition relevance place (TRP): In these cases, even if the recipient displays their understanding by responding just after the turn with *j'étais là*, the speaker continues (max. 4 further TCUs).

Due to the emotionally charged character of the tellings, the recipient's affiliative response (i.e., demonstrating emotional involvement) seems to be made relevant by the speaker through multiple resources. The affective stance is displayed not only verbally but also in an embodied way, thus making the speaker's stance audibly and visibly accessible for the recipient (especially Excerpt 6.14). Only some bodily resources are, however, recurrently used with *j'étais là*, as the following table shows:[184]

Multimodal conduct	Frequency
gaze at (one of the) interlocutor(s) at the end of the quote	65.0% (37/57)
enacted gaze	91.2% (52/57)
eyebrow raise	15.8% (9/57)
frown	3.5% (2/57)
facial expression	50.9% (29/57)
indexical gesture	17.5% (10/57)
dropping hand gesture	10.5% (6/57)
headshake	12.3% (7/57)
shrug	7.0% (4/57)

---

[183]   In 44/57 cases for *j'étais là*, the turn-taking takes place in less than 4 intonation units.

[184]   The sum of these occurrences does not add up to 57 because, of course, speakers may simultaneously perform several body movements. If this is the case, I counted them once in each category.

freezing	5.3% (3/57)
other gestures	5.3% (3/57)
change in posture	14.0% (8/57)
not visible	8.8% (5/57)

Table 25: Multimodal conduct accompanying [*j'étais là* + quote]

Despite the restricted informative value of the quantifying results, the table illustrates that speakers use facial expressions in about one half of the occurrences of *j'étais là* (50.9%). As the speaker's use of facial expressions, other than recipient gaze, may be considered independent from the number of interlocutors, this high number may be a meaningful result. Interestingly, even though the data is highly diverse, this result is similar to what Table 17 shows for German *ich dachte (mir)*. With *j'étais là*, 'enacted gaze'[185] is also recurrent (91.2%). The rather regular use of eyebrow raises and frowns, headshakes, indexical gestures, and changes in posture (together in about 68.4% of all cases)[186] confirms that speakers are involved in a depicting stance-taking activity while reporting their purported thought with *j'étais là*. The recipient's systematic gaze at the speaker, sometimes just after *j'étais là*, may indicate the recipient's orientation toward the necessity to *see* what speakers quote (Excerpt 6.15).

Taking an affective stance may be a tool for the speaker to make more explicit with which stance recipients are invited to affiliate subsequently. Indeed, in 28 cases, the recipient's response is affiliative. If the delivered response is not sufficient or dispreferred (i.e., disaffiliative), speakers seem to work toward eliciting the preferred response (Excerpt 6.13). There are also 9 disaffiliative responses: Instead of affiliating, co-participants request more information concerning the prior telling (*n* = 4; see Excerpt 6.14) or react to a detail of the prior turn and not to the speaker's stance (*n* = 1). Co-participants can also provide an explanation or additional information referring to the cause of indignation reported by the speaker (*n* = 3). There is also one case where the disaffiliation with the speaker is an affiliation with the third party that the speaker reported on. In another seven cases, recipients are only aligning by delivering a structurally fitting response to the prior turn (Stivers *et al.* 2011) through verbal or bodily acknowledgement tokens.

The 37 cases in which recipients display affiliation or disaffiliation show that the recipients emotionally engage with the speaker's telling. This is in line with Goffman's (1974) observation that

> what the individual spends most of his spoken moments doing is providing evidence for the fairness or unfairness of his current situation and other grounds for sympathy approval, exoneration, understanding, or amusement. And what his listeners are primarily obliged to do is to show some kind of audience appreciation. They are to be stirred not to take action but to exhibit signs that they have been stirred. For what a speaker does usually is to present for his listeners a version of what happened to him. (*ibid.*: 503f)

---

[185] Like in German, this category comprises instances where speakers reenact gazing at imaginary interlocutor on their left or right, when they depict an affective stance through gaze (looking up, squinting their eyes), or when they reenact a gaze away while reporting a dispreferred action during the quote.

[186] There are only two combinations of different embodied resources that occur more than twice: [facial expression + dropping hand gesture] (*n* = 3) and [facial expression + change in posture] (*n* = 4).

What Goffman describes as "being stirred" can probably be interpreted as displaying affiliation. With *j'étais là*, speakers provide, through their self-quotation, their own response toward what happened to them in the past. By doing so, they replay the effect of the events on themselves, thus exhibiting to the recipient what they must affiliate with.

My analysis has shown that *j'étais là* does not allow speakers to clearly display reported thought, which makes it a hybrid quotative, compared to *je me suis dit* ('I said to myself') or German *ich dachte (mir)* / *ich habe (mir) gedacht* ('I (have) thought (to myself)'), where speakers can at least clearly *display* their reported material as thoughts. The excerpts show, however, that, at least in my data, this ambiguity does not seem to cause interactional problems. When quotes introduced by *j'étais là* occur in the middle of a constructed dialogue with a third party, it seems relatively clear that it reports speech. When occurring in the evaluation phase of a telling, [*j'étais là* + quote] can also close a storytelling by reporting thought.

Why do speakers use the social action format *j'étais là* to display their affective stance and not a declarative format of the type [*je pense / je trouve* ('I think / I find') + assessment]? There are several possible explanations to that, which are not mutually exclusive: Making the affective stance part of the telling in the past lowers the speaker's responsibility and thus also accountability for their actions because they are presented as part of the story world. Framing the stance as such keeps the conditional relevance of a response in the ongoing conversation ambiguous: It may not be entirely clear whether the strong affective stance projects a response in the telling or also in the ongoing conversation. The numbers show that in most cases, recipients do, however, respond soon after the pattern [*j'étais là* + quote]. The ambiguity about whether *j'étais là* projects reported thought or speech adds an additional layer: Because speakers use *j'étais là* to take a strong, mostly negative affective stance, the ambiguity allows them to adapt their talk accordingly in case of the recipient's sanctioning response (i.e., by framing the quote retrospectively as thought) even though this may not correspond to reality.

Last but not least, my analysis has shown that the high degree of bodily depiction seems to have mainly two functions: (1) indicating shifts in footing, also on a bodily level (see specifically Excerpt 6.15), which are interactionally projected by the deictic *là* ('there'), and (2) either foreshadowing a negative affective stance by first depicting it through prosodic and bodily means, or by depicting a negative stance embodiedly while remaining (at least in the beginning) rather vague on a verbal level.

Throughout my analysis, I have gathered several features of *j'étais là* that point to its routinization as a social action format projecting a verbal or embodied quote:

(1) *J'étais là* is exclusively used to introduce a direct quote, or it is systematically followed by a reenactment of past talk or thought without deictic shifts.

(2) *J'étais là* seems to have a fixed position regarding the quote—namely, preceding the quote that it introduces. The quotative cannot grammatically project more to come according to Auer's (2002) criteria (see Section 2.2) because it is syntactically complete. As *j'étais là* is nevertheless recurrently used to introduce a quote, it may have developed into a projector-construction that interactionally projects "more to come."

(3) The material that *j'étais là* projects is, in my data, always a quote which can be verbal, embodied or both. Despite the high degree of variation of the quoted material, all quotes *depict* prior talk or thoughts.

(4) A last argument for *j'étais là* routinizing is its apparent loss of grammatical information regarding tense and aspect. In its use as a quotative, the social action format can be preceded by a telling in narrative present tense and followed by a quote in narrative present tense while remaining, itself, in the imperfective tense. I will further discuss this point in Chapter 7, where narrative present tense is treated in more detail.

The functioning of *j'étais là* as a social action format projecting a quote in everyday talk differs in many ways from *je me suis dit*. In what follows, I will therefore compare both quotatives regarding their formal, positional, sequential, actional, and embodied features.

6.3.3. Summary and Comparison of the French *Je me suis dit* and *J'étais là*

In Sections 6.3.1 and 6.3.2, I have shown that even though both social action formats, *je me suis dit* and *j'étais là*, introduce direct reported thought (or, in the case of *j'étais là*, also speech), they seem to be functionally different and occur in different kinds of sequences. The two social action formats project distinct actions, which tend to co-occur with different degrees of embodied conduct. This finding clarifies that the multimodal analysis of mundane interaction may be necessary to paint a detailed (and potentially more comprehensive) picture of certain grammatical formats, such as direct reported thought—a suggestion that Keevallik (2018: 14) has already made for direct reported speech.

If we compare the two quotatives *je me suis dit* and *j'étais là* in more detail, several differences became manifest throughout my analysis:

(1) Compared to *je me suis dit*, *j'étais là* can introduce a wide range of linguistic and non-linguistic material through which speakers depict a past scene, thus allowing the recipient(s) to (re-)experience what happened to them. According to Streeck (2009), depictions can be primarily separated into two categories of bodily conduct: *mimesis*, which depicts actions, and *handling*, which depicts "objects, or instruments, or accessories of action" (*ibid.*: 144). In my data, speakers use both kinds of bodily conduct: *mimetic* conduct to display their affective stance, often through facial expressions (Excerpt 6.14), and *handling* gestures (Streeck 2009) and posture to imitate the manipulation of objects or the like (Excerpt 6.13). These body movements stand in contrast to the ceiving or ception gestures (*ibid.*) that often accompany direct reported thought introduced by *je me suis dit* and help speakers to structure their talk. Recipient gaze may give us further evidence for that difference: compared to quotation with *j'étais là*, recipients do not orient the same way to the speaker's bodily conduct with *je me suis dit*; in 30 of 41 cases, recipient gaze is already established before the quotation with *je me suis dit* starts. There is only one instance where the recipient shifts her gaze to the speaker just after *je me suis dit*—compared to 13 instances with *j'étais là*.[187]

---

[187]   In 3 other cases with *j'étais là*, recipients gaze at the speaker at some point during the direct reported thought. In the 9 remaining cases, there is no recipient gaze, thereof 3 cases where recipients gaze at an object instead and one case where she turns away to cough. Whereas the comparison of quantifying multimodal results between German and French might need to be treated with caution, the comparative

(2) While *je me suis dit* allows speakers to display purported past thoughts only, *j'étais là* is an interactional tool that often leaves open whether what is projected has been said or only thought. Despite hints through contextual information, neither interlocutors nor analysts can know exactly what has been said or thought in the past. This ambiguity may be an important interactional tool, as Haakana (2007) has pointed out for a similar pattern in Finnish, to utter strong assessments without bearing its potential consequences: Speakers may still back down in case of a clear disaligning or disaffiliative stance of the recipient. Because it is not always clear whether the speaker's affective stance is a shared one, the ambiguity between thought and speech allows speakers eventually to downgrade, such as an utterance from the frame "said" to "only thought." If a common affective stance has been established beforehand, it seems, however, possible, to report thought while taking full responsibility for the strong affective stance that is displayed (Excerpt 6.14).

(3) In contrast to *je me suis dit*, *j'étais là* introduces a response to something very specific that has been depicted just before, for instance someone's reported speech. The characteristic of being reactive to a preceding action or event comes with the direct reported thought being shaped as emergent responses that are produced "on the fly." Speakers report strong assessments, curses, or judgements—all emotionally charged—that are staged as unfiltered responses in situations where the speaker has been somewhat snubbed or surprised. These characteristics make the social action format *j'étais là* an efficient means to enhance intersubjectivity by letting recipients relive how the speaker has been literally affected by what they report. Following Goffman's (1959) notion of performances,

> [w]hen an individual plays a part he implicitly requests his observers to take seriously the impression that is fostered before them. They are asked to believe that the character they see actually possesses the attributes he appears to possess, that the task he performs will have the consequences that are implicitly claimed for it, and that, in general, matters are what they appear to be. (*ibid*.: 28)

Accordingly, a dramatizing, affective dimension may make the speaker seem more reliable because the "impression" that speakers perform is designed to appear truthful by being staged as unfiltered, spontaneous, and depicted as immediate as possible. In contrast, with *je me suis dit*, speakers seem to report situations where they are less affected on an emotional level. Instead, they report reasonings that are staged as emerging from the speakers' logical and rational reflections—the overall stance thereby being a non-emotional one.

(4) One thing that both social action formats have, at least partially, in common is their position in the turn. They both occur toward the end of multi-unit turns and can indicate upcoming turn-closure.[188] In the closing environment of a sequence, both actions (i.e., the epistemic and the affective stance-taking through direct reported thought/speech) project a response as a relevant next. This leads me to the next major distinctive feature between *je me suis dit* and *j'étais là*.

---

results of the qualitative analysis suggest that, because *je me suis dit* is less followed by bodily depictions, visual attention form the recipient may be less necessary.

[188] In contrast to *j'étais là*, *je me suis dit* occurs equally often in the middle of a telling sequence. I have outlined in Section 6.3.1.3 which consequences the position of *je me suis dit* in the sequence has on the conditional relevance of a response.

(5) [*Je me suis dit* + direct reported thought] makes various responses by the co-participant relevant, depending on its position in the telling sequence. In contrast, [*j'étais là* + quote], due to its high degree of affectivity, is mostly responded to by the co-participant's affiliation. The different responses clarify that through next turn proof-procedure (Sacks *et al.* 1974), the distinct actions that speakers carry out with direct reported thought become visible. Additionally, the different responses according to the position of the reported thought within the sequence index the positional sensitivity (Schegloff 2007) of the social action format.

Analyzing the speakers' verbal *and* bodily conduct contributed in a crucial way to the finding that *je me suis dit* and *j'étais là* seem to be functionally distributed. This fact rekindles the debate about the grammatical role of quotations (see Chapter 2), thereby addressing the recent discussion of a potential grammar-body interface (Bohle 2014; Couper-Kuhlen 2018; Keevallik 2013, 2018; Mondada 2016; Pekarek Doehler 2019; Pekarek Doehler *et al.* 2021).

So far, I have only compared the two French quotatives with each other. In the following section, I will move on to the cross-linguistic analysis, where I compare my findings in German to the findings in French. The comparison will show that there are multiple similarities, on a linguistic, sequential, and multimodal level, between French and German direct reported thought, even though the quotatives are lexically distinct.

6.4. The Comparison of French and German Direct Reported Thought in Past Tense: Formal and Functional Results

In what follows, I present three pairs of excerpts, each consisting of one German and one French piece of data. Each pair illustrates one converging function between the social action formats projecting direct reported thought in both languages. The comparison concerns the sequential placement of the direct reported thought, its multimodal features, and the action that it allows speakers to accomplish. I will always present the German and French converging excerpts in a row before analyzing them. The following three pairs, each containing one quotative, will be discussed:

1) Excerpts 6.16 and 6.17: *ich habe gedacht* vs. *je me suis dit* allowing the speaker to make publicly available their emergent rationalization that builds up toward a climax.
2) Excerpts 6.18 and 6.19: *ich habe mir gedacht* vs. *j'étais là* allowing the speaker to display their affective stance in the closing environment of a storytelling.
3) Excerpts 6.20 and 6.21: *ich dachte* vs. *j'étais là* allowing the speaker to display their strong affective stance, which is reenacted through bodily means, especially through gaze, as dispreferred response within the reported storytelling.

The first two excerpts show converging instances of direct reported thought, once with German *habe ich gedacht* (lit. 'have I thought'), once with the French *je me suis dit* ('I said to myself'). The French and German quotatives, even though being lexically distinct, are both used to display the speaker's emergent rationalization in a past situation. In both excerpts, the reported thought is part of the building process for recipients to understand the upcoming climax. Excerpts 6.16 and 6.17 show instances where rationalizing direct reported thought is inserted into a storytelling as a multi-unit turn containing an important part of the plot.

The German excerpt is a conversation between a consultant, his client, and an assistant (the latter does not interact). The client is jobless and came to the consultant to learn strategies on how to get a new job fast. Both client and consultant work (or have worked) as people selling medication to physicians for pharmaceutical companies. To illustrate that "asking the right questions" to the physicians is important, the consultant tells the following exemplifying story about an eureka-moment when he was still working as an agent himself:

Ex. 6.16: allgemeinmediziner / general physician (FOLK_E_00174, 25min29)

```
THO: Thomas, DOM: Dominik

01 THO: un ich bin von arzt zu Arzt gelaufen einunNEUNzig-=
        and I went from doctor to doctor in ninety-one

02      un habe gefrAgt (.) setzen sie thyroXIN ein;
        and asked          do you use thyroxin?

03      (.) JA=JA=JA setz ich ein setz ich ein setz ich EIN;=
            yes yes yes I use (it) I use (it) I use (it)

04      un ich hab mich gewundert warum meine +UMsätze nich
        and I wondered why my sales didn't
   tho                                         +gz at DO-->l.05

        hochgingen- (.)
        go up

05      °hh (.) +%un nach +%!EI!nem jahr in (.) in% Irgend$eim
                and after a year    during during some
   tho       -->+.........+gz down-->l.09
                %.........%3 fingers to forehead,
                               shakes head %
   dom                                          $gz at
                                                TN->l.17

        gespräch °h °ha°_ich gedAcht-
        conversation I have thought (lit. have I thought)

06      %(0.45) %
   tho  %freezes%

07 THO: ich frAge diese ganzen allgemEInmediziner ob sie
        I ask     all these general physicians   if they

        thyroXIN einsetzen-=
        use     thyroxin

08      naTÜRlich-
        of course

09      °h die %frage% is viel zu +allgeMEIN-
            the question is way too general
   tho         %shakes
                head%
                        -->+gz at DO-->l.11

10      °h

11      >(mein) ich< (.) +herr DOKter-
         I say          sir (lit. mister physician)
                   -->+gz at LI-->l.12

12      (0.23)+
   tho       -->+gz at DO-->l.19
```

```
13 THO: stellen sie !SEL!bst auf thyroxin ein; (.)
        do you adjust (the patients) to thyroxin yourself

14      <<creaky>NEE->
                  NO

15      (0.21)

16 THO: das [macht der LUNG+enarzt;¹⁸⁹ (.) nEe das]
        the pulmonologist does that    no that

17 DOM:     [((laughs))    +das war klar] ˚h
                           that was clear
   dom                     +gz down-->l.18

18 THO: MACH +ich nich;
        I don't do
   dom    -->+

19      ja an thyroxin ˚>trau ich mich nich RAN;<˚
        yes I don't dare to touch thyroxin

20      [˚h]h okay:,
             okay

21 DOM: [˚h]

        [((laughs)) ]

22 THO: [aber sie se]tzen sie doch EIN-
         but you use PART it

23      ja aber NUR auf >empfehlung des lUngenarztes;<=
        yes but only on the recommendation of the pulmologist
```

In the French excerpt, Marie and Josiane are talking about two mutual acquaintances, Leila and her ex-boyfriend, who maintains a toxic friendship with her. Starting at line 01, Marie tells Josiane how she came to realize the ex-boyfriend's manipulative character.

Ex. 6.17: se faire manipuler / to get manipulated (Pauscaf_02_05, 19min03)

```
MAR: Marie, JOS: Josiane

01 MAR:                                     [j'te pro  ]METS-
                                             I promise you

02      (0.4) j'me suis rappelé d'↑ÇA l'aut' jour
        I have remembered that the other day

        quand j'suis [allée à klÉber avec ma MÈre] pa'ce quE, ˚hh
        when I went to kléber ((name of bookstore)) with my mother
        because

03 JOS:              [((sings a high note))    ]

04 MAR: +%j'ai vU le LIvre;% (0.4)
         I saw the book
   mar  +gz at JOS-->l.05
         %memes an open book with both hands%closes palms-->l.05

05      >qu'elle elle<% lui avait acheTÉ,+
        that she she had bought him
```

---

¹⁸⁹ Due to anonymization, the medication name has been changed to Thyroxin. I am aware that Thyroxin is a thyroid medication and that the specialist 'pulmonologist' does not fit here. In the original data, which I do not have access to, the medication probably has been a medication for the lungs, which is why the corresponding specialist here is the pulmonologist.

```
   mar              -->%keeps palms closed-->l.07
                            -->+gz right-->l.08
06 JOS: [mais incroy#A:ble,]
          but unbelievable

07 MAR: [˚hh       #h     ] ˚je me suis dit%# mais en FAIT-˚
                            I told myself      but in fact
   fig            #fig1                   #fig2
```


fig1

fig2

```
08      (0.6)+
   mar    -->+gz at JOS-->l.13
09      le tYpe il lui a dIt qu'il voulait ce LIvre (.)
        the guy  told her that he wanted this book

10      pour éviter d'se faire manipuLER,
        to avoid getting manipulated

11 JOS: mais [il veut peut-être (la MÊme).]
        but   maybe he wants the same

12 MAR:      [mais c'est ↑!LUI!          ] qui mani[PUle. ]
              but  it's him who manipulates

13 JOS:                                           [ah non]
                                                    oh no

        mais elle/ c'est un gros +ma#nipulaTEUR,=
        but  she  it's    a  big  manipulator
   mar                         -->+
   fig                      #fig3
```


fig3

```
14      =et j'le sais/ et on le di[t£ depuis l'dé:BUT; ]
        and I know it and we say that from the beginning

15 MAR:                          [il sait eXA          ]ctement
                                   he knows exactly

16      quoi lui dIre pour qu'elle !PAR!te (.)
        what to tell her so that she runs

        [direct vers LUI,           ]
        directly toward him
```

```
17 JOS: [(xx), c'est °incroyA:ble°-]
             it's unbelievable
```

Both social action formats project a direct reported thought consisting of a multi-unit turn which allows the speakers to report their post-hoc rationalization of certain circumstances (for a similar instance of *je me suis dit*, see Bangerter *et al.* 2011). Positioned in the middle of the telling, the direct reported thought reproduces the moment where speakers shifted their epistemic state from K- to K+ (Heritage 2012), thus displaying their sudden understanding of the problem they described before.[190]

In the German excerpt, after depicting through a constructed dialogue the confusing situation between the general physicians' claim (line 03) and the sales numbers (line 04), Thomas uses the format [*habe ich gedacht* + multi-unit quote] to make publicly available his eureka-moment when he finally understands the situation (lines 07–09). Subsequently, Thomas illustrates his rationalization with a reported dialogue (lines 11–23), which functions as an account for the rightfulness of his change of epistemic state.

In the French excerpt, Josiane first describes the situation in which she realized that Leila's ex-boyfriend is manipulative (lines 01–05). Even though her co-participant, Marie, already displays her (inferential) understanding through an affiliative *incroyable* ('unbelievable'; line 06), Josiane uses the format [*je me suis dit* + long quote] to make publicly available her post-hoc reasoning that made her realize the ex-boyfriend's manipulative character. She first displays her change from K- to K+ with *mais en fait* ('but in fact'), and then reports her purported thoughts leading to her conclusion, thus delivering an account for her preceding negative assessment (lines 07 to 12).

Note that, in French and German, speakers multimodally stage their emergent thinking process prior to verbalizing it. These gestures are part of the staging in the *ongoing* conversation. They cannot be part of what is displayed as quoted thought, which appears, as an inherent characteristic, without gestures. In German, Thomas moves his fingers to his forehead (line 05) while shaking his head and then freezes this posture during the pause after *ha ich gedacht* (line 06). In French, Josiane first makes a skeptical facial expression (fig1) and then moves into a thinking posture and gazing into space (fig2). This posture is also held through the pause between *je me suis dit* and the multi-unit quote. In both languages, the speakers' bodily conduct foreshadows, through an iconic depiction, the action that is about to follow—namely, reflecting upon some event from the past that they only understood retrospectively. Note that the pauses between the quotative and the quote evidence the projective force of the social action format, because in neither of the excerpts does the interlocutor take the turn. This projective force may not only be due to the grammatical projection of the quotatives *ich dachte (mir)* / *ich habe (mir) gedacht* ('I (have) thought (to myself)') and *je me suis dit* ('I have said to myself'), but also to the maintained depicting bodily conduct of both speakers which allows

---

[190]   In this function of displaying a discrepant assumption, the German occurrence is similar to what Deppermann and Reineke (2017) described for German *ich dachte*. However, one main difference between this practice and mine is the form of delivery. In my occurrences of *ich dachte (mir)* / *ich habe (mir) gedacht* where speakers display their sudden understanding of a situation, the shift of epistemic state is staged as such with prosodic and bodily resources. This is not the case in the excerpts that Deppermann and Reineke (2017) present in their article.

them to interactionally project "more to come" (i.e., the verbalization of their rationalization). Recipients seem to recognize that the reported thought is an important part of the narrative illustrating a shared stance from the speaker's perspective.

The cross-linguistic approach allows me to demonstrate that, for the format of direct reported thought, action ascription seems to function in a similar way in French and German—*despite* the lexical difference of the quotatives—with recipients being able to understand the differences in use thanks to the precise verbal, sequential, and multimodal design of the telling activities and the social action formats projecting direct reported thought.

The next two excerpts illustrate two cases where direct reported thought allows speakers to take a strong affective stance in storytellings with bodily resources. In German, the reported thought is introduced with *ich habe mir auch schon gedacht* ('I have also PART thought to myself'), in French with *j'étais là*. In the following German excerpt, two friends are having lunch together. Carola talks about her job: She is about to finish her practical year and a half at school (after university) to finish her degree as a high school teacher. Both participants know that the job perspectives for German teachers (in Germany) are particularly bad, especially in the city where they live, which explains Sonia's emphatic display of surprise at line 06.

Ex. 6.18: Deutschmangel / lack of German (teachers)[191] (DeJA-VI_014, 41min36)

```
CAR: Carola, SON: Sonia

01 CAR: unser schulleiter is Echt SUper;
        our headmaster is really great

02      (1.2)$
   son     >>$gz at her plate-->l.06

03 CAR: heute hat er mir sein LEID geklAgt- (--)
        today he lamented his sorrow to me

04      es fEhlt !IM!mer noch chemie und +D(h)EUTSCH;
        they are still missing (a teacher of) chemistry and german
   car                                  +gz at SON-->l.14

05      h°um_h°um

06 SON: [d!EU!T    $]SCH?
         german

07 CAR: [(ich soll)$]
          I should
   son         -->$shifts gz quickly to CAR-->l.08

08      ((1 nod))$ ((2 nods))
   son        -->$gz at her fork-->l.10

09      die ham DEUTSCHmangel-=
        they have a lack of german (teachers)

10      =ich hab mir AUCH schon >ged$acht<,
         I have also PART thought to myself
   son                         -->$gz at CAR-->l.12

11      #(0.4)%(0.5)#
   car   --->%sits straight,smiles-->l.12
   fig   #fig1a,b--->#fig2a,b
```

---

[191] Because of the data being eye-tracking data, the images 1b and 2b include small circles indicating where the co-participant, here Sonia—is gazing at this moment.
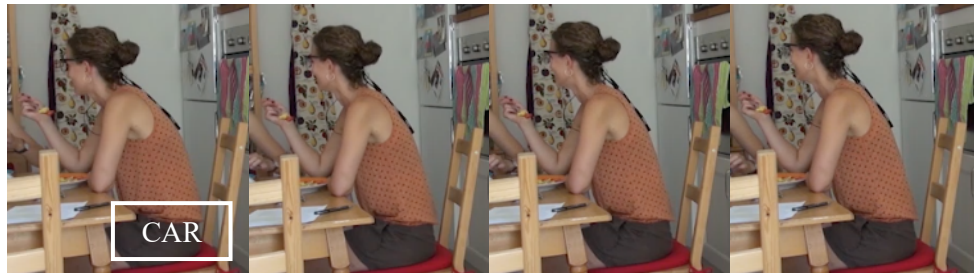
```
12 CAR: hehehe%he;+
   son          -->+gz front-->l.15

13 SON: ↑↑JUh↓u;
         yay
14      (0.3)+
   car     -->+gz in front of her-->l.16

15 CAR: °hh+ aber (.) ich mein:- (.)
            but      I mean
   son  -->+gz at her plate-->>

16      ich glaub ich hab es ganz geSCHICKT gemacht dass ich
        I think that I've been quite clever to

        jetz nächstes jahr acht STUNden +deutsch unterrichte- (..)
         teach now eight hours of german next year
   car                            -->+gz at SON-->>

17 SON: [<<creaky>§Ωhm.>§ ]

18 CAR: [des     §Ωheißt§] (.) wenn ich ΩgEh BRAUchen se jemand;
         that means     when I leave they will need someone
   son           §lifts eyebrows§
                 Ωnods--------------->Ω
```

The French excerpt is a conversation between two sisters, Liliane and Carole, and Liliane's boyfriend, Pascal. They talk about a newsletter that they all receive via email. Just prior to this excerpt, Pascal said that he never reads this newsletter because it is usually not interesting at all. The excerpt starts with Liliane storytelling about an exception when she read the mail. The email's subject said that one could win two tickets to Europa-Park, a large amusement park in the south of Germany.

Ex. 6.19: europa-park (Pauscaf_02_08, 02min47)

```
LIL: Liliane, PAS: Pascale

01 LIL: mais+ hier j'ai LU;
        but yesterday I read
   lil    >>+gz at PAS—>l.02
```

```
02      >parce que< %l +le tit$re de: +du MAIL% c'était;
        because th- the title of the email was
   lil            %points on table with
                        index finger-->%hands in front of
                                        her chest-->l.05
               -->+gz down at
                 her gesture-->+gz at PAS—>l.04
   pas                   -->$gz at LIL's gesture-->l.03

03      gagn$ez deux places pour europa-PARK-
        win  two tickets to europa-park and
   pas  -->$gz at LIL-->l.05

04      et du coup +[>j'tais §°là°< (.) #OH$:-          ]
        and so   I was here
   lil            -->+closes eyes,then gz up-->l.05
   fig                              #fig1
```


fig1

```
05 PAS:            [eh ÇA  $c'est un      $beau titre] de +mail
                    eh that's a nice title for an email
   pas                  $nods-->>
                                -->$gz at CAR-->

   lil                                      -->+gz at
                                               PAS-->

    pour %que tu LIses+$ ouais.
     to make sure that you read (it) yeah
   lil               -->+
        -->%
                     -->$

06 LIL: <<laughing>ouais,>
        yeah
```

In both excerpts, the speaker reports an event that is against an expected course of action. In the German excerpt, Carola just reported about the school looking for German teachers, which is usually not the case (line 04 and reformulated at line 09), which is evidenced by Sonia's response (line 06); in French, Liliane reports reading an email which all co-participants agreed upon never being worth reading (not in transcript). The direct reported thought introduces a potentially new information for the co-participants in an environment of already established shared knowledge or experience (both German speakers studied German for being a high school teacher, all three French participants receive the newsletter). The social action format projecting the reported thought is produced with latching to the prior intonation unit in German (lines 09 and 10), while in French, *j'étais là* is preceded by *du coup*, a particle that expresses the following action to be consequential. That speakers move, with the reported thought, toward the unexpected part of their telling is emphasized, in French and in German, through an

acceleration of tempo the closer speakers move toward the quote itself (accelerated *gedacht* in German (line 10) and *j'étais là* in French (line 04)).

The format of the quote is rather similar in both languages. Its high degree of affectivity is primarily conveyed through the speakers' embodied conduct. In the German excerpt, Carola does not say anything. Instead, her reported thought consists of her changing her posture by straightening her back, sticking her head up in the air as if to attract the headmaster's attention (fig1a and fig1b, and fig2a and fig2b).[192] She thus reacts, within the reported dialogue, to the headmaster's 'lamenting' with curiosity that is displayed through bodily means, thus manifesting her interest for a future position at the school. Through Carola's additional smile and her posture of heightened attention, she positively assesses the lack of German teachers. Sonia affiliates with Carola's embodied display of happiness through a response cry (line 13), thus co-animating Carola's positive affective stance.

In the French excerpt, Liliane displays her positive affective stance through the emphasized response cry *oh* (line 04), her body posture and her facial expression (fig1): she leans forward onto her arms on the table, her hands form a prayer-like gesture, her head and her gaze are directed to the ceiling. In contrast to the German excerpt, her co-participants, however, do not affiliate with her positive stance. Instead, Pascal mocks her falling for the title of the email (line 05). His mocking nevertheless demonstrates that he understood Liliane's display of enthusiasm about the possibility to win those tickets, thus making Liliane's reported thought a successful tool to display her affective stance.

In the German and the French excerpt, the (main) recipient's gaze is either already on the speaker (French excerpt) or shifts to the speaker during the quotative (German excerpt), thus demonstrating the recipient's orientation toward the speaker's visible embodied conduct. In both cases, recipient gaze is retrieved shortly after the redemption of the visible quote (German: line 12; French, line 05, where Pascal shifts his gaze to the third participant, Carole).

Compared to Excerpts 6.16 and 6.17, the direct reported thought is staged, in Excerpts 6.18 and 6.19, as a sudden, spontaneous, unfiltered response to prior events. This display is achieved through prosodic and bodily features only, which do neither occur in that same intensity nor with such a depictive character when speakers use direct reported thought to make publicly available their emergent reasoning. Considering the highly depictive sequential environment in which the direct reported thought occurs in Excerpts 6.18 and 6.19, the direct reported thought is thus well embedded into the storytelling.

The last two excerpts focus specifically on the speakers' and recipients' multimodal conduct. There are 53 of 57 occurrences of *j'étais là* where speakers express their affective stance through a reenacted gaze and/or a facial expression. For German, this holds for 80 of 123

---

[192] Because Carol's quote consists of a body movement only, it was not easy to decide upon the transcription of the body movement. The quote is prosodically projected by the quotative which has final semi-rising intonation (line 10). The movement is thus projected as being part of the quotation and consequentially 'belongs' to Carol's turn. The annotation of Carol's name 'CAR' for her multimodal conduct should, in that sense, not be necessary. For the sake of clarity, I decided, however, to mark 'car' at the beginning of line 11. Another question is whether the pause should not go on the same line as the quotative. Because Carol's body movement does not start immediately after the quotative but only after a 0.4s pause, I decided to put the body movement on a separate line, thus understanding the quote as a separate TCU.

occurrences[193] with *ich dachte (mir)* or *ich habe (mir) gedacht*. This depiction does not only concern the action itself (e.g., an affective stance display), but also its interactional features, such as its preferred or dispreferred status within the reported dialogue. The next two excerpts illustrate how speakers depict the dispreferred character of their direct reported thought in the original telling on a verbal and multimodal level.

The following excerpt has already been discussed in detail in Section 6.2.5. Three roommates, Sina, Nadja, and Marius have invited Fiona, a potential new roommate, to a "roommate casting." The excerpts starts with Nadja's reenacting Marius while eating a dish that she cooked for him but messed up.

Ex. 6.20: Wie gummi / like rubber (FOLK_E_00253, 20min35)

```
FIO: Fiona, NAD: Nadja, MAR: Marius, SIN: Sina
18 NAD:  [%(.) a$(h)h°#aß er so+ DA$ °ge°,
              ((laughter)) he like sa(t) there
   nad  ->%hand down-->l.19
                           -->+gz in front of her-->l.19
   fio       -->$gz at NAD------------>$gz down-->l.22

19         %gE+[$H::t,]
             (it's) okay
   nad  ->%slowly l h to mouth as if to eat-->l.21
         -->+gz left-->l.21

20 FIO:        [$gut, ]
                 good
   fio          $imitates MAR eating w r h,lifts eyebrows-->l.22

21 NAD:    %+hu[$hum::,+
   nad  ->%l h slowly down-->l.23
           +gz at FIO+gz down-->l.24

22 FIO:        [$(ja$ha,) °h
                  yehes
   fio       -->$gz at NAD-->>
                 -->$

23 NAD:    %und ich dAcht so    %HM,
             and I   thought like hm
   nad  ->%l h up,chest height%holds l h-->

24     (.) %<<high pitch>°ZUldigu:ng°>,+#
                          sorry
   nad  -->%covers mouth with l h-->l.28
                           -->+gz briefly l+gz down-->>
   fig                          #fig1
```

fig1

```
25 FIO: [((laughs))        ]
26 SIN: [((laughs))        ]
27 MAR: [((laughs))        ]
28 NAD: [°hab ich %(echt)°] verKACKT. (.) °h
          I've (really) fucked up
    nad        -->%l h down, manipulates pen-->>
```

In the French excerpt, Elinda tells Ekti about her best friend who apparently has not been sure about her sexual orientation for a while.

Ex. 6.21: un doute / a doubt (Pauscaf 20, ca. 41min10)

```
ELI: Elinda, EKT: Ekti

01 ELI: d'ailleurs ma meilleure amie-
          by the way my best friend

02      elle pensait+ qu'elle avait un DOUte pendant un moment,=
          she thought   that she had   a doubt for a moment
    ekt              +gz at eli-->>

03      =#j'étais là #(0.5) faudrait peut-être sa↓VOIR hein-
          I was there        might be good to know
    fig   #fig1        #fig2
```
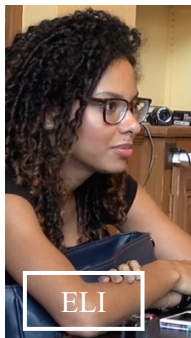

fig1
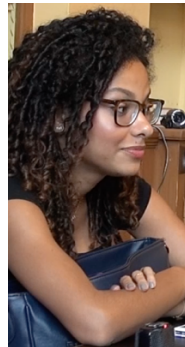
fig2

```
04      parce que§ chais !PAS! si ça va être paREIL; h°
          because I don't know if it's going to be the same
    ekt          §grabs bottle with l hand-->l.05

05 EKT: £h°£^$hoHO,
    ekt        $opens bottle, moves it twd mouth-->l.09

06 ELI: j'te jure quand elle m'a dit ↑ça >j'°é°tait LÀ;<
          I swear when she told me that I was there

07      (0.3)#(0.7) a#tTENDS là;# h°
                    wait a minute
```

186

#fig3   **#fig4**   **#fig5**

fig3   fig4   fig5

08   (0.6)

In both the German and the French excerpt, the direct reported thought allows the speaker to display their affective stance. In the German excerpt, Nadja's excuse (line 24) displays her shame for her failed dish, whereas in the French excerpt, Elinda displays her despise (line 03) and her shock (line 07) toward her friend telling her being insecure about her sexuality. Compared to Excerpts 6.18 and 6.19, where the affective stance has been primarily conveyed through bodily resources, speakers use verbal *and* bodily conduct in these last two excerpts. Note, however, that the verbal quotes are rather brief with two fixed expressions (*zuldigung* ('sorry'; line 24) in German, *attends là* (lit. 'wait here'; idiomatically 'wait a minute'; line 07) in French), and one abbreviated clause in French (line 03) where the expletive *il* of the imperative *il faudrait* ('one should') is left out.

In both excerpts, the speakers reenact how they respond to a first pair part of someone else: In German, Nadja responds to Marius' display of not liking the dish, in French, Elinda responds to her friend telling her that she might be lesbian. The speakers' multimodal conduct fits their respective verbal display: Nadja covers her mouth with her hand and looks left and down, while all the recipients gaze toward her (line 24, fig1). Elinda displays her negative stance through facial expressions (fig2 and fig3–fig5), including raised eyebrows and, both times, her gaze to her left.

The gaze away, which both speakers perform, has been evidenced to be a recurrent multimodal feature of dispreferred responses to questions (Kendrick & Holler 2017), also accompanying specific routinized verbal patterns like French *chais pas* ('dunno'; Pekarek Doehler 2019). Even though the responses are not second pair parts to questions, it may be possible that in similarly orderly adjacency pairs like [assessment + response to assessment], this multimodal conduct is comparable.

Both speakers reenact their responses through bodily means as dispreferred responses. This conduct demonstrates that speakers seem to treat their co-participants as if they were characters of the ongoing telling (Pfeiffer & Weiß 2022): Gazing away in a reenacted scenario where the interlocutor is not the one at whom the dispreferred response has originally been addressed demonstrates that speakers create intersubjectivity through the dramatization of events, and thus also through direct reported thought, in a way that co-participants can re-experience the original event as if they were a part of it.

These six excerpts have demonstrated that in both French and German, the two types of stance-taking are not only achieved using specific social action formats. A complex interplay of position in the telling sequence, prosody, and the bodily conduct of the speaker allow the speaker to recipient-design direct reported thought in a way that interlocutors understand its function within the telling.

The cross-linguistic analysis showed that affective and epistemic stance-taking are recurrently displayed through direct reported thought in French and German. The quotatives that introduce direct reported thought are lexically distinct in both languages. In French, the affective stance is almost exclusively taken with direct reported thought introduced by the social action format *j'étais là*; In German, speakers preferably use *ich dachte (mir)*. In contrast, to make publicly available an emergent reasoning or post-hoc rationalization, French speakers use *je me suis dit*; In German, speakers tend to prefer the analytical past form *ich habe (mir) gedacht*. On an actional level, German and French speakers do not only use direct reported thought in a comparable way, but they also seem to exploit multimodal resources similarly: The actional distinction reflects in an enhanced use of facial expressions and reenacted gaze patterns with direct reported thought for affective stance taking, while fewer resources occur during emergent reasonings.

In both languages, the pattern [quotative + quote] seems to be treated, by both speakers and recipients, as a bipartite structure: speakers use the social action format of the quotative as a projective device that introduces "more to come," thereby relying potentially on the grammatical projection of the CTP, which is the quotative, or, in the case of *j'étais là*, on the underspecification of the deictic. Recipients orient to the bipartite structure by not interrupting speakers during pauses between quotative and quote and, in the case of *j'étais là*, by shifting their gaze to the speaker just after the quotative ($n = 13$). Of course, this orientation, from the speaker and the recipient, should be verified with a larger data set.

To conclude, the comparison of French and German has proven highly beneficial to investigating direct reported thought; I was able to show that lexical convergence is not sufficient to determine (inter-)actional parallels between the German and French social action formats introducing direct reporting thought. I have shown that the German format of direct reported thought with *ich dachte (mir)* / *ich habe (mir) gedacht* ('I (have) thought (to myself)') does not find its "actional equivalent" in the French *je pensais* / *j'ai pensé*. Starting from the action of reporting thought made it possible to detect two French social action formats as functionally parallel quotatives to the German *ich dachte (mir)* / *ich habe (mir) gedacht*: *je me suis dit* and *j'étais là*.

7. Direct Reported Thought with Quotatives in Present Tense

7.1. Introduction

In this chapter, I investigate direct reported thought in French and German projected by quotatives in narrative present tense. The narrative present tense—also called "historical present tense"[194]—describes the phenomenon of speakers using present tense to report events that have happened in the past. Because this kind of present tense often occurs in tellings that bear certain dramatizing effects, prior literature has argued that narrative present tense has a function of "Vergegenwärtigung" ('presentification'/'visualization'; for a critical discussion of presentification, see Fleischman 1986; see also Casparis 1975: 22; Zeman 2013: 237).

Diachronic research based on French and German narrations suggests that the narrative present tense may have developed out of "the evaluative function of meta-narrative narrator comments" (Zeman 2013: 237, my translation). For Middle High German, Zeman (2013) shows that in conceptually oral texts of narratives, the narrators' comments are carefully placed between episodes of tellings. Consider the following example taken from Zeman (2013: 250):

(10) mich **dunket** sie **hân** bêde reht.
      ‚Mir scheint, sie haben beide recht.'
      [*Parzival*, 264, 25, um 1200]

This excerpt (lit. 'it seems to me they are both right') is inserted into a telling in past tense. While telling the story, the narrator steps out of the role as narrator and addresses the listener (or reader) by commenting on the prior episode. According to Zeman (2013), such a shift "implicitly sets relevant the relation of the teller to the telling" (*ibid.*: 250). However, in Zeman's excerpts, this evaluative function is restricted to cases where the teller switches *only* for this evaluation from past to present tense. In my data, this is not always the case.

Similar results have been obtained for French. Fleischman (1986), inspired by Labov and Waletzky's (1967) work on everyday conversational storytelling, shows for medieval "performed stories" (i.e., conceptually oral texts like the *Chanson de Roland*) that the narrative present tense is used in the evaluation phase of a telling (Fleischman 1986: 221–224), when a speaker "momentarily steps out of diegetic time to address his listeners directly" (*ibid.*: 213). Especially in the genre *chansons de geste*, a genre of orally told stories about heroic acts during battles, the internal evaluation (Labov & Waletzky 1967) of a telling by the narrator is common (Fleischman 1986: 224). Even though Zeman's (2013) and Fleischman's (1986) findings are based on narrative written texts, it appears as if speakers similarly use narrative present tense in my spoken data to evaluate (parts of) their tellings. This may be due to the conceptually

---

[194] Because the term *historical present* is problematic (for a discussion, see Zeman 2013: 240ff), I will use *narrative present tense* instead. Despite the different degrees of emotional involvement in the tellings I analyze, they all have in common that speakers structure their multi-unit turn, to work either toward a climax or toward a conclusion. Because of a certain variability of the phenomenon in my data regarding the position in the sequence where the shift to narrative present tense occurs, I follow Koch and Oesterreicher's (2011) larger conceptualization of narrative present tense, as they include cases where present and past tense alternate in a telling and cases where present tense is exclusively used throughout the whole telling (for a more restricted understanding of narrative present, see Fludernik 2002).

spoken character of those written texts, which are first and foremost oral stories that have been written down only later. The present chapter will weigh the findings that have been made based on written tellings against everyday conversational data.

On a syntactic level, the same clause-combining features that I have presented for past-tense quotatives apply to present-tense quotatives (see Section 6.2.3). As with past tense, I conceptualize the pattern [quotative + quote] as a bipartite structure with the following features:

- *Ich denke (mir)* ('I think (to myself)') and *je me dis* ('I say to myself') are complement-taking predicates that still grammatically project a complement.
- When introducing a direct reported thought, this empty argument position is filled by the quote. The quote is not syntactically linked through the complementizer *dass* or *que* ('that'). There is, however, a pragmatic link between the two juxtaposed units.

One important linguistic parallel between German past and present-tense quotatives is the co-occurrence of *ich denke (mir)* with additional lexical items. As I have shown in Section 5.4.1.3, my data suggests that the predicate systematically co-occurs with prepositioned items like *und* ('and'), *da* ('there'), or *dann* ('then') and/or with post-positioned *so* ('like') when used as a social action format projecting direct reported thought (for an overview, see Tables 33-34). This feature is distinctive when it comes to separating *ich denke (mir)* as a quotative from *ich denke (mir)* functioning as an epistemic parenthetical.[195]

In the remainder of this chapter, I show how direct reported thought in narrative present tense is used in German and French everyday talk. I first analyze the German quotative *ich denke (mir)* (Section 7.2) and then the French quotatives *je me dis* (Section 7.3.1) and *je suis là* (Section 7.3.2). The chapter closes with a comparative analysis of both languages (Section 7.4).

7.2.   Reporting Thought with the German *Ich denke (mir)* ('I think (to myself)')

In this chapter, I will analyze the social action format [*ich denke (mir)* + direct reported thought]. Morpho-phonologically, the quotative in first-person singular is constructed as follows:

ich	denk-e	(mir)
	think-1SG.PRS	PRO. REFL.1SG

I think (to myself)

The reflexive pronoun is optional and is placed either before or after the conjugated verb, depending on the overall syntactic configuration of the clause.

---

[195]   Other features also contribute to the differentiation. When *ich denke (mir)* functions as a quotative, the quote is often bodily reenacted, the pattern [*ich denke (mir)* + direct reported thought] occurs in the sequential environment of a telling, and the social action format *ich denke (mir)* does not seem to be positionally flexible.

7.2.1.   A General Picture: A Quantifying Analysis

In my data, 33 occurrences of *ich denke* and 19 occurrences of *ich denke mir* introduce direct reported thought. As mentioned before, one feature of *ich denke (mir)* as social action format projecting direct reported thought is that it co-occurs with additional lexical items, in verb-first syntax or as coordination ellipsis.[196] These features hold for 30 cases of *ich denke* and all occurrences of *ich denke mir*. The most frequent co-occurring items are the following:

Form	Frequency
ich denke (mir) + so ('like')	32.7% (17/52)
und ('and') + *ich denke (mir)*	28.9 % (15/52)
wo ('where') *ich (mir) denke*	17.3% (9/52)

Table 26: Lexical items co-occurring with *ich denke (mir)*

For cases with *und* ('and'), I also included combinations with *da*, *dann* ('there'/'then') or *deswegen* ('this is why'), or with *so* ('like').[197] Another recurrent syntactic structure is *wo ich (mir) denke* ('where I think (to myself)'), which can also co-occur with *so* ('like').

On a phonological level, some occurrences of *ich denke (mir)* introducing direct reported thought bear typical phonological features of frequent use. Stein (1995) mentions the morpho-phonological reduction of the first-person present tense due to dialectal and familiar speech, such as *ich denke* becoming *ich denk* through the elision of the final phoneme schwa (*ibid*.: 141). In my data, almost all instances of *ich denke (mir)* functioning as a quotative show this reduction, which is why I did not count it in the category of "morpho-phonological reduction or clitization." The following table shows these typical prosodic and morpho-phonological features:[198]

Feature	Frequency
morpho-phonological reduction / clitization (e.g., *deng_ich* or *un_denk ich*)	30.0% (11/33)
accelerated	42.4% (14/33)
no accentuated syllable	66.7% (22/33)
pause after *ich denke (mir)*	21.2% (7/33)

Table 27: Prosodic and morpho-phonological features of *ich denke* as a quotative

As the table illustrates, only the feature of non-accentuation occurs in more than half of the cases. Morpho-phonological reduction, one of the main possible indicators of routinization, affects only one third of the occurrences of *ich denke* as a quotative. This changes with the reflexive form:

Feature	Frequency
morpho-phonological reduction / clitization (e.g., *deng_ich mir* or *un_denk ich mir*)	94.7% (18/19)

---

[196]   For a detailed explanation of the difference between V1 and coordination ellipsis, see Section 6.3.4.

[197]   Consequently, there are four cases that are counted twice because they fulfill two of the three criteria from the table.

[198]   As with the past form, the basis of this analysis is an auditory analysis, following the idea of *impressionistic hearing* (Couper-Kuhlen & Selting 1996: 45; Kelly & Local 1989). No software such as Praat has been used to investigate the forms more in detail.

accelerated	84.2% (16/19)
no accentuated syllable	100% (19/19)
pause after *ich denke mir*	42.1% (8/19)

Table 28: Prosodic and morpho-phonological features of *ich denke mir*

With 19 occurrences, no generalizing conclusions can be drawn. The numbers show, however, that there is a higher tendency for the reflexive form to be reduced on a morpho-phonological and prosodic level than for the non-reflexive form.

In present tense, the reflexive form has a higher relative frequency in use as social action format projecting direct reported thought than does the non-reflexive form: In 79.2% of the 24 total occurrences in my corpus, *ich denke mir* is used as a quotative. This is a higher ratio than for the non-reflexive form (33 out of 130 total occurrences, or 25.4%). However, because my total number of occurrences is low, this tendency should be verified with a larger data set.

7.2.1.   Syntactic Features of *Ich denke (mir)* ('I think (to myself)')

Similar to the past forms of *denken*, the social action format *ich denke (mir)* can occur in various syntactic shapes.[199] In Section 6.2.3, I outlined which clause-combining patterns are possible for the CTP *denken*. For the present work, only the format [matrix clause + dependent main clause] and cases in which the quotative is combined with multimodal conduct are relevant. It is only in such bipartite structures that *ich denke (mir)* ('I think (to myself)') as social action format projecting direct reported thought occurs. As a reminder, in the syntactic pattern [matrix clause + dependent main clause], there is no syntactic linkage with the complementizer *dass* ('that'), and the verb in the subordinate clause is in second position instead of last position. The following table shows the frequency of the dependent main clause pattern:

Quotative	Frequency[200]
ich denke	45.5% (15/33)
ich denke mir	68.4% (13/19)

Table 29: Frequency of [*ich denke (mir)* + dependent main clause]

In some rare cases, the quotative is followed by a question (word) or an imperative.[201]

Quotative	Question (words)	Imperatives
ich denke	1/33	1/33
ich denke mir	2/19	1/19

Table 30: Frequency of *ich denke (mir)* introducing an imperative or a question

---

[199]   For positional issues concerning the order of *ich denke (mir)* and its quote, see Section 6.2.3.2.
[200]   These numbers also include cases in which the dependent main clause is preceded by a response cry.
[201]   The number of question (words) and imperatives include cases where the quote is preceded by a response cry. The combined numbers of Tables 29, 30, and 31 do not add up to the total number because questions and imperatives are counted twice, as dependent main clauses and as questions or imperatives (see Section 6.2.3.1). There are also 4 cases of *ich denke* that do not appear in these tables because they are followed by a repair and are parenthetically inserted or interrupted after the quotative.

Most of the remaining occurrences are cases in which *ich denke (mir)* introduces only response cries, bodily movements, or verbless structures. The following table shows the non-clausal material following *ich denke (mir)*:

Quotative	Response cries only	Bodily conduct only	Verbless structures
ich denke	18.2% (6/33)	6.1% (2/33)[202]	6.1% (2/33)
ich denke mir	21.1% (4/19)	-	5.3% (1/19)

Table 31: Other formats of quotes introduced by *ich denke (mir)*

Another configuration in which *ich denke (mir)* functions as a mere speech-situating device is when it is followed by a multi-unit turn. As with the past forms, *ich denke (mir)* allows speakers to subsequently produce a new episode of the telling or a more elaborate evaluation. However, this format is rare, as the following table shows:

Quotative	Frequency
ich denke	9.1% (3/33)
ich denke mir	10.6% (2/19)

Table 32: Frequency of the pattern [*ich denke (mir)* + multi-unit turn]

Tables 30–32 demonstrate that *ich denke (mir)* as social action format projecting direct reported thought can be followed by several types of quotes other than complete clauses. The recurrent use of *ich denke (mir)* as quotative and its formation of chunks with additional lexical material (see Section 6.2.4) can be an indicator of *ich denke (mir)* being on its way to routinize. In what follows, I will present combinatory aspects concerning *ich denke (mir)*.

7.2.2. Combinatory Aspects of *Ich denke (mir)* ('I think (to myself)')

The following tables shows the most frequent items preceding *ich denke* (Table 33) and *ich denke mir* (Table 34):

Form	Frequency
und ich denke (so) ('I think (like)')	27.3% (9/33)
aber ich denke ('but I think')	10.0% (3/33)
(und) wo ich denke ('(and) where I think')	21.2% (7/33)
da(nn) denke ich ('there/then I think')	10.0% (3/33)

Table 33: Combinatory pattern with *ich denke* introducing direct reported thought

Form	Frequency
und (da/nn) denke ich mir (so) ('and (there/then) I think (like)')	15.8% (3/19)
und ich denke mir ('and I think')	26.3% (5/19)
wo ich mir denke ('where I think')	15.8% (3/19)

---

[202]   There is one additional occurrence in which an explanatory dependent main clause follows the body movement.

da(nn) denke ich mir ('there/then I think')	10.5% (2/19)

Table 34: Combinatory pattern with *ich denke mir* introducing direct reported thought

As with both past forms, there is one recurrent post-positioned item that co-occurs primarily with one of the previously mentioned adverbs or particles: the adverb[203] *so* ('like'), which occurs 12 times with *ich denke* and four times with *ich denke mir*.[204] According to Wessels (2019), *so* is, in combination with *denken*, not a pointing deictic, but a deictic used "to demonstrate something […], namely the positioning-activity as such" (*ibid*.: 66, my translation). She analyzes *denke + so* as an A-part that projects a B-part (reported thought) used by speakers to "imply their positioning, without the need to make their evaluation explicit" (*ibid*.: 66, my translation).

The tables show that the combination with *und* ('and') is the most frequent. Interestingly, the combination with *da/nn* ('there/then') is not as frequent as it is with the past form (see Section 6.2.4). One reason for this could be that, with *ich denke (mir)*, speakers place their direct reported thought on the interface of the original situation and the ongoing conversation, thus leaving it up to the recipient to decide whether a response is projected as a relevant next. In such instances, speakers would not need temporal adverbs because their reported thought would not be presented as a next step in their telling.

Similar to the past forms, two additional syntactic patterns in which *ich denke (mir)* occurs are verb-first syntax ($n = 6$), and the coordination ellipsis ($n = 8$). In both formats, the front field may remain empty or may be filled with *und* ('and'). The subject pronoun is either unrealized or positioned after the verb, which does not correspond to the typical syntactic format in German, in which the subject precedes the verb. Verb-first syntax (V1) is part of the German standard grammar. It occurs in written language in several formats (*cf.* Auer 1993: 19), such as in the protasis of a non-introduced conditional clause. In spoken language, the most recurrent formats with V1 are imperatives or questions. Research on everyday talk, however, has shown that there is a specific context in which speakers resort to V1: tellings.

Günthner (2006) investigates the syntactic format of V1 in tellings specifically in narrative present tense tellings. Combining IL and Construction Grammar, she argues that V1 leads to "dense constructions" with a specific interactional meaning; V1 constructions introduce "new action steps of the telling or even initiate a punchline" (*ibid*.: 103, my translation; for similar observations, see Günthner 2000), thus being integral to dramatizing. She argues that

> [w]ith this interconnection of verbal positioning and narrative present tense, the narrator has the possibility not only to focus on the 'action-relatedness' in a general way, but also—in contrast to the use of the perfect tense—on the concrete action itself (Günthner 2000): The utterances begin with the semantically central verb part and thus with the mention of the narrated action. (*ibid*.: 101, my translation)

---

[203] *So* ('like') may also be categorized as modal deictic/adverbial (Stukenbrock 2021), allowing the speaker to draw attention to the following reported thought. To stress the lexical category of *so* regardless of its use, I keep, for now, the category of adverb.

[204] Because these numbers also include *ich denke (mir)* with V1 or as a coordination ellipsis, they are not equivalent to the numbers in Tables 33 and 34.

In my data, the quotation in narrative present tense does not co-occur significantly more often with V1 and coordination ellipsis than with the past tenses; the present tense constitutes 27.0% (14/52) of all cases, compared to 23.9% (33/138) for both past tenses (26.5% for the synthetic past form and 23.0% for the analytic past form). Notwithstanding whether the present-tense quotative is also in V1, the narrative present tense contributes, according to Günthner (2006), to the dramatization and staging of a telling.

In my data, V1 and coordination ellipsis make up 26.9% of the 52 total occurrences of *ich denke (mir)* introducing direct reported thought. This result suggests that the pattern [*ich denke (mir)* + direct reported thought] has a specific function when occurring in tellings. Consider the following example:

7.1: waffelform / waffle mold (FOLK_E_00329, 58min28)

```
18 DAR: ich hab se auf der ähm oben aufs ceRANfeld gestellt,
        I put them on the uhm up on the ceramic hob

19      hab die waffelform beFÜLLT, (0.26)
        filled up the waffle mold

20      °h un_dEnke, (.)
         and think

21      ↑SO:↓:,
         so

22              [und WIE kriegst du jetz die beFÜLlte waffelform]
                 and how do you get the filled waffle mold now

((two lines of overlapping responses omitted))

25 DAR: auf dein BACKblech?
        on your baking tray?
```

In this excerpt, in line 19, the speaker uses V1 to accelerate the progression of the telling (Günthner 2006). At line 20, she continues the telling with her direct reported thought, introduced by *un_denke* ('and think'). The omission of the subject pronoun *ich* ('I'), which remains latent from line 18, and the syntactic linkage with *und* ('and') makes lines 20–25 a coordination ellipsis. In line with Günthner's (2006) findings (*ibid*.: 103), the format allows the speaker to introduce a next part of the telling—here, its gist: Daria does not know how to move the silicon waffle mold onto the baking tray.

The following interactional analysis will present four excerpts containing direct reported thought introduced by *ich denke (mir)*. Through a sequential and multimodal analysis, I aim at showing how the social action format *ich denke (mir)* is used in everyday talk.

7.2.3.   Interactional Functions of *Ich denke (mir)* ('I think (to myself)')

With the following four excerpts, I show that, similar to the past forms, *ich denke (mir)* frames direct reported thought. Although many features between the past and present forms of *denken* as a quotative converge—syntactically, sequentially, and multimodally—there is one major difference between them: direct reported thought with *ich denke (mir)* is, in my data, almost exclusively used to display an affective stance in storytellings.

The first excerpt illustrates the most frequent pattern with direct reported thought in my data. [*Ich denke (mir)* + response cry and/or body movement + clause] occurs in 23.1% ($n = 12$) of

the cases in my corpus. Three friends talk about Polish women coming to Germany to work as caregivers. They specifically address the problems of such foreign caregivers who live with an elderly person but do not speak their language, which often leads to misunderstandings. In her telling, Selina reports a situation in which she and a team from the institution where she works went to an elderly man's house and discovered that the Polish woman taking care of him did not understand much of what he was saying.

Ex. 7.2: kraft aus polen / worker from poland (FOLK_E_00228, 49min40)

```
SEL: Selina, MAR: Marina, ANN: Anna

01 SEL: als ¢wir da mal hingegangen sind um die zu beSUchen-
            when we went there once to visit her
    mar    >>¢gz at SEL-->l.16

02        °hh $ham_mer erst mal fünfhundert sachen AUF↑gedeckt; (.)
                we first uncovered five hundred things
    ann       $gz at SEL-->l.14

03        also °h (.) nee das meint er nich +SO,=
            well        no he doesn't mean it like this
    sel                                    +gz at ANN-->l.05

04        =er sacht >des un DES;<=
            he says this and that

05        =>un die nur< ach+↑!SO!::,
            and she like oh okay
    sel                 -->+gz up-->l.06

06        ich dachte immer ich soll_s so un ↑SO +ma#chen;
            I always thought I have to do it like this and that
    sel                                   -->+gz at MAR-->l.08
    fig                                      #fig1
```


fig1

```
07        ∫↑JETZ verst#eh ich_s;∫
            now I understand it
    mar    ∫headshake---------->∫
    fig           #fig2
```


fig2

```
08        °hh+ (0.3)# deng_ich auch HÄ,
          I   think PART what
   sel  -->+gz right-->l.09
   fig           #fig3
```


fig3

```
09        >also ich mein< +wEnn die schon ne kraft aus polen
          well I mean if they PART bring a worker from poland
     sel               -->+gz down------------------------->

          HER+bringen; °h
          here
   sel  -->+gz to ANN-->l.11
10        (0.43)
11 SEL: dann+ (0.39) muss die+ auch ↑DEUTSCH können;
        then          she has to speak german
   sel   -->+gz down-------->+gz at ANN-->l.12
12       $(1.04)+$
   sel      -->+gz to MAR-->>
   ann  $3 nods $
13 SEL: also,
        well
14 MAR: ja: da ka_ma$ sich [gleich$ SE]Lbst versorgen.
        yeah then one can also just care for oneself
   ann          -->$
15 ANN:                   [hm=   $hm.]
   ann                           $reaches tw glass,drinks->>
16 MAR: also [DA  ] kann¢ste auch ins ALtersheim;=
        well then  you can also go to a retirement home
   mar            -->¢
17 SEL:     [ja. ]
```

Between lines 01 and 07, Selina reports a dialogue between her, her team, and the Polish caregiver. The depicted scene serves as an account for her opinion, which she uttered prior to this excerpt: that it is problematic when foreign caregivers do not speak the language of the people they care for. The exemplifying character of the dialogue transpires through the extreme-case formulation at line 02 and the subsequent vague formulations that the speaker uses; without naming concrete misunderstandings, Selina uses only approximative deictics when referring to unspecified events: *so* ('like this'; line 03), *des und des* ('this and that'; line 04), and *so und so* ('like this and that'; line 06). She reports how she and her institution explained to the Polish caregiver what she misunderstood (lines 03 and 04), how the caregiver reacts to this explanation (using a change-of-state token; line 05), and what she originally thought (line 06). While reporting, Selina exaggerates the scene bodily and prosodically, with high-pitched movements and stressed syllables that intensify the display of her emotional involvement (lines 05–07).

Marina, at whom Selina gazes at the end of line 06, produces an affiliative headshake (line 07), displaying her emotional co-involvement. The other recipient, Anna, still has not responded at this point in the telling.

The reported dialogue is followed by an inbreath that projects turn-continuation (Schegloff 1996: 105) and a short pause during which Selina turns her head and gazes to her right (line 08). Her facial expression displays non-understanding (fig3) and allows her to express her negative affective stance even before its verbalization. The subsequent quotation is formatted as follows: under the same intonation contour as the social action format *deng_ich auch* ('I think PART'),[205] the speaker utters a prosodically stressed response cry. The preceding facial expression and the response cry form a multimodal assembly that allows the speaker to display her negative stance economically; the pattern is verbally short, also due to the V1 of the quotative itself (Günthner 2006), and dense in semiotic (including bodily) resources that allow Selina to also visibly convey her attitude toward the reported situation.

There is no syntactic relation between the quotative *deng_ich auch* and the quote because neither *hä* ('what') nor the speaker's facial expression can constitute a complement of the CTP *denken*. However, the complement-taking predicate grammatically projects an open argument position, which is filled by this response cry and the facial expression. *Deng_ich auch* frames the following multimodal assembly as direct reported thought, thus establishing a pragmatic relation between the two parts. The prosodic attachment of the response cry to the quotative illustrates the difficulty of determining the boundary between quotative and quote: Such a delimitation may be possible on a syntactic level because *deng_ich auch* can be determined as a clausal unit containing a predicate with *auch* modifying this predicate, whereas *hä* does not. On a prosodic level, however, this delimitation is hardly possible because of *hä* forming one intonation unit with the quotative (see Section 3.1.2 for a problematization).

In the following TCU, Selina uses *ich mein* ('I mean') as a discourse marker, which functions as "discourse-organizing phrase" (Günthner & Imo 2003: 11): It is followed by a *when*-clause and has lost its status as matrix clause, now functioning as a structuring particle (Günthner & Imo 2003: 9ff) that is prosodically backgrounded through acceleration. The complex *when*-clause repairs Selina's strong affective stance from line 08 by delivering an explanatory turn (lines 9–11). The repair also allows her to re-emphasize her stance (line 11), thus verbalizing what she has bodily depicted before (line 08).

In this specific case, it seems the direct reported thought format is exploited by the speaker to depict, clearly and illustratively, her affective stance in response to the previously reported scene. The combination [quotative + *hä*] does not necessarily report a thought from the past; instead, it seems indexical for the speaker's scandalized stance resulting from the described circumstances. This brief format of quotation not only is an economic way of taking a stance, but it also fits the sequential environment of the preceding constructed dialogue.

Note that, during the *when*-clause, Selina gazes and re-gazes at Anna (lines 09, 11, and 12), the only participant who has not reacted so far. Gaze is a powerful resource for mobilizing

---

[205]   Note that *auch* functions here as a modal particle and not as an adverbial. I thereby follow Diewald and Fischer (1998), who argue that the fact that *auch* cannot be placed in the front field excludes its grammatical categorization as an adverbial (see also Imo 2007: 52ff for a discussion of this characteristic, and Thurmair 1989).

response (Stivers & Rossano 2010). During the subsequent pause (line 12) which indicates story completion, Anna finally produces three (dispreferred) aligning nods (Stivers 2008: 51), which may be why Selina shifts her gaze to Marina at the end of the pause, who subsequently produces a more elaborate, affiliative response (lines 14 and 16). Anna, in contrast, disattends from taking an explicit stance by reaching for her glass and drinking (see Hoey 2018) during her second alignment, using a minimal response (line 15).

This excerpt has shown a typical pattern of direct reported thought with *ich denke (mir)*: speakers first display their strong affective stance through prosody, bodily means, and a response cry, and only then verbally. Direct reported thought thereby is an economic way to display a stance that speakers present as emerging from the prior situation. At the same time, the strong affective stance may need further explanations to make recipients respond affiliatively, as illustrated by the additional work Selina invests to obtain a response from Anna. The pattern of an explanation following a strong affective stance has been evidenced for tellings in which "disagreement has surfaced in the interaction or is in the air" (Couper-Kuhlen & Thompson 2005: 260), as displayed by Anna. In triadic conversations, such as this one, even if one participant does not affiliate, the speaker can still orient to the other recipient, who may.

The next excerpt shows a dyadic conversation in which the speaker uses a similar bipartite structure—namely, *wo ich so denk so hä* ('where I like think like huh'). In this case, she does not receive the recipient's affiliation after the response cry. However, the speaker keeps extending her turn until affiliation is displayed, which demonstrates that the pattern [*ich denke (mir)* + direct reported thought] may be initially deployed for mobilizing the recipient's affiliation, which is pursued if not delivered immediately. Hannah tells Jonathan about the mountain bike accident that a friend had when both were on a mountain bike tour. The excerpt starts when Hannah tells how either the mountain rescue or the ambulance (she cannot remember) did not find her and her friend in the forest.

Ex. 7.3: bergrettung / mountain rescue (KAFKU_04, 15min20s)

```
HAN: Hannah, JON: Jonathan

01 HAN: =ja oder ich weiß nich ob_s §die +bErgrettung war oder der
        yes or I don't know if it was the mountain rescue or
   han                                       +gz at JON-->l.02
   jon                                 §nods-->

        NOTarz§§§t-=
        the ambulance
   jon     -->§
            -->§closes eyes-->l.11

02      =>+%#auf jeden fall< §ham ses nich +ge%@#!FUN!@dn;
            anyway           they didn't find it
   han  ->+closes eyes------------------->+gz at JON-->l.04
           %................................%l PUOH,hold-->l.04
                                            @raises eyebrows@
   jon                   -->§gz at HAN-->l.08
   fig        #fig1                            #fig2a,b
```

fig1   fig2a   fig2b

```
03 JON: h° [h°      ]

04 HAN:    [>wo ich] so+ denk so<#@HÄ?#+%
              where I like think like huh
   han             -->+gz left------->+gz right-->l.06
                                @moves head fw-->l.05
                                -->%gesticulates w
                                   left PUOH-->l.05

   fig                        #fig3#fig4
```


fig3   fig4

```
05     (0.6)          @
   han  -->head back@

06 HAN: des+ [kann              ] doch nich +SEIN;
        that is just impossible
   han  -->+gz at JON-------------------->+gz l,down-->l.09

07 JON:      [((clears throat))]

08 HAN: also $grad/ %(0.4)    $(0.2)
        I mean especially
   han            -->%gesticulates w both h PUOH-->l.18
   jon     -->$gz twd HAN's h $gz at HAN-->l.19

09     +da passiert jetz +wirklich +%nich (.) SUper
        it's not super rare that
   han  ->+gz at JON------->+gz up--->+gz at JON-->l.11
                                    -->%

       selten was- $(0.3)
       something happens there
   jon             $ 1 slow nod back, 4 rapid nods-->l.10

10     [und (.) +grad] DIE müsstn doch ungefähr wissen wO-$
        and     especially they should know roughly where
                                                 -->$

11 JON: [mh        +m-  ]
   han          -->+

12 HAN: >%oder +wie< sie da hin%komm;
        or how they get there
   han   %shrug--------------->%
               +gz at JON-->l.14
```

200

```
13        $>aso< ds sin ja breite WEge-
           I mean these are broad paths
     jon  $nods-->l.16

14        ds sin ja+ (0.5) wirk+lich FORSCHDwege;
           these are PART really forestry roads
     han        -->+gz at
                   her gesture+gz at JON-->

15        (0.4)

16 JON:   hm.$+
     jon  -->$
     han    -->+gz down-->>

17        (0.6)

18 HAN:   na[$JA.]
           well

19 JON:   [$ja:] des is halt eigentlich au wert$volle ZEIT-
           yeah  it is PART PART also valuable time
     jon  -->$gz down------------------------->$gz at HAN-->>
```

At line 02, Hannah reaches the gist of her telling, which is depicted through an eyebrow raise, a deictic gesture to an imaginary place (fig1), and a palm-up open-hand gesture (fig2a and fig2b), thus emphasizing the obviousness of the problem (Marrese *et al.* 2021). So far, Jonathan has only minimally reacted with nods (line 01) and a choked laughter (line 03). In overlap with Jonathan's breathy laughter, Hannah continues her telling with a follow-up (Selting 2017) in the format of direct reported thought. Introduced by the pattern *wo ich so denk so* ('where I like think like'), the direct reported though allows Hannah to comment on the gist of her telling by displaying her negative affective stance—here, her incomprehension of the ambulance not finding her and her friend. The format of an adverbial clause with temporal *wo* ('where') indexes an additional temporal link between the reported thought and the previously described circumstances from lines 01 and 02.

The quote itself resembles the one from Excerpt 7.2: It seems to exaggerate the speaker's stance from the past, thus conveying Hannah's non-understanding from the past situation. Hannah's embodied conduct is also similar to Selina's in Excerpt 7.2: she first gazes away (to the left, line 04), then, at the end of the response cry, at her interlocutor (fig4). Hannah simultaneously moves her head forward and makes another palm-up open-hand gesture, thus depicting her disbelief in the ambulance not finding them. Note that the response cry also carries the focus accent of the entire intonation unit. This assembly of multiple resources allows Hannah to display her negative affective stance economically, as in the prior excerpt. However, despite Hannah using multiple resources to display her negative affective stance toward the described situation, Jonathan (still) does not respond.

The response cry is followed by a 0.6 s pause during which Jonathan's response remains noticeably absent (line 05). Note that, during this pause, Hannah retracts her body from the interaction space by moving her head back at the end of the pause as if also to leave the interaction physically. The lack of uptake may be why she then continues her turn with a more elaborate explanation of her negative stance (lines 06–14). After a cutoff at line 08 and a pause during which she gesticulates as if to accelerate her self-repair (for a similar bodily conduct

during word searches, see Goodwin & Goodwin 1986), Hannah pursues an affiliative response from Jonathan with the help of several resources. First, she delivers an account for being so scandalized about the doctors not finding them (line 09), then she reaffirms the problem at lines 10 and 12, which she extends, after Jonathan's overlapping minimal response (line 11), by giving further evidence about the paths (lines 13 and 14). During Hannah's extended account, Jonathan first displays his access to Hannah's stance through several nods (line 09), thus possibly projecting an upcoming affiliation, and aligns with an overlapping acknowledgement token (line 11). Jonathan's long nodding from line 13 to line 16, however, which continues even during a pause and a backchannel (line 16), seems to be treated as disaligning, as it orients toward the non-completion of Hannah's story, and disaffiliative.

The pause at line 15 indicates, once again, that Hannah moves toward closure. Just before another pause (line 17), she gazes down (line 16), which may indicate her yielding the turn (Kendon 1967). With *naja* ('well') in post gap turn-extension and falling final intonation, Hannah closes her turn. This time, Jonathan responds with an extended aligning response (line 19), which displays his agreement with Hannah's negative stance toward the doctors. However, note that his response (line 19) remains vague, carrying signs of dispreferredness: the lengthened initial *ja* ('yes'), the three modalizing particles *halt eigentlich auch*, and the delay in the delivery of his response to Hannah's long stance-taking and account.

Excerpt 7.3 shows that, despite Hannah's attempt to mobilize Jonathan's affiliation, she is only partially successful in obtaining the preferred response to her stance display with direct reported thought. In contrast to the previous excerpt, there is no third interlocutor she can turn to for a preferred response. The direct reported thought allows Hannah to display her strong negative affective stance while complaining about the mountain rescue's incompetence. In line with what Haakana (2007) describes for reported thought in complaint stories, the format allows Hannah to make her own stance available, thus suggesting to her recipient the stance he should affiliate with. By shifting her footing from the telling in past tense (line 02) to her commenting on it through direct reported thought in present tense (line 04), Hannah situates her affective stance in the conversation of the here-and-now. By doing so, her own stance-taking becomes more immediate, which may increase the pressure on Jonathan to respond affiliatively. That Hannah pursues this affiliation after a lack of uptake demonstrates that a response to her direct reported thought would have been expected.

The next excerpt illustrates what happens when the teller is confronted with two divergent stances from recipients. It also shows how the matter of reporting thought can be topicalized by a recipient. In Excerpt 7.4, the social action format functions as projector-construction for a longer turn—a format that only concerns 9.6% of *ich denke (mir)* introducing direct reported thought in my data (*n* = 5). Three friends (the couple Zoe and Nils and their friend George) are talking about being at a salon. Prior to this excerpt, Nils complains about hairdressers who ask him how he wants to have his hair cut. George affiliates and says, referring to a similar situation that he recently experienced, that the hairdresser should be the one to know how to cut his hair. The account he delivers for his stance is only vaguely related; he justifies his opinion with the fact that the hairdresser "looks gorgeous" and adds that hairdressers also "mostly look a bit stupid" (not in transcript), which provokes the side sequence at line 01.

Ex. 7.4: frag mich do net dumm / don't ask me stupidly (FOLK_E_00293, 1h32min20s)

```
ZOE: Zoe, GEO: George, NIL: Nils

01 ZOE: +immer diese $VORurteile.
          always these prejudices
    zoe  +shifts gz to NIL-->l.02
    geo            $gz at ZOE-->l.04

02 GEO: isch +mag ¢aber SO was-
        but I like that kind of thing
    zoe    -->+gz at GEO-->l.06
    nil          ¢gz at GEO-->l.08

03     (0.44)

04 GEO: °i$sch° bin net >so $unbedingt< (0.22) für die (.)
         I'm not so much                 into the
    geo  ...................$gz at NIL-->l.06

        HELlsten,
        most clever ones ((speaking about women))

05     (0.23)

06 GEO: die (.) $isch mag eher so (0.37)+
        the     I prefer more like
    geo      -->$gz at ZOE-->l.08
    zoe                        -->+gz at NIL-->l.09

        [de$¢ BITchische typ,]
        the bitchy type

07 NIL: [((laughs, 1.41s))

08 ZOE: [is$¢ in ORdnung; IS ] in Ordnu[ng-]
         it's okay        it's okay
    zoe -->$gz at NIL-->l.10
    nil -->¢gz at ZOE-->l.09

09 GEO:                          [äh-]+:¢
                                  uh
    zoe                          -->+gz at GEO
                                     -->l.12
    nil                             ¢gz at GEO
                                       -->l.11

10     (0.22)$
    geo   -->$gz at ZOE-->l.11

11 GEO: IS ja$ au e¢g[  a+l-    ]
        never mind
    geo   -->$down-->l.12
    nil        -->¢gz down-->l.14

12 ZOE:         [spr+isch]d$ für DIE.+
                 speaks for them ((the women))
    zoe         -->+gz at NIL---->+gz down-->l.14
    geo             -->$gz at NIL-->l.13

13     (0.34)$(0.2)
    geo    -->$gz down-->l.15

14 GEO: un (.) >deng¢_i§sch mir<-+ (.)
        and I think to myself
    geo              $.......-->l.15
    nil        -->¢gz at ZOE-->l.16
    zoe                     -->+gz at NIL-->l.15
```

```
15       jetz +FR§Ω$AG ↑misch do net dumm;+
         don't ask me stupidly now
   geo        -->$open hand,palm left-->l.16
                -->Ωshrug-->l.17
                 -->$gz at NIL-->l.16
   zoe    -->+gz at GEO's hand------->+...->l.16

16       schn+eid sO dass [du$ mit mir heut¢ $abend WEG§gehen
         cut them so that you would go out with me tonight
   geo                   -->$down-->l.17
                                      -->$2 beats  $,,,
                                              ->l.17
   zoe    -->+gz at GEO-->l.17
   nil                             -->¢gz at GEO-->l.21
         würdest.§Ω]

17 NIL:                   [((laughs, 2.5s))

             §Ω] $haha↑HA +$°HH%
   geo ,,,,,,,,,$
             -->Ω
              -->$gz at ZOE$gz at NIL-->l.19
   zoe                   -->+gz at NIL-->l.20
                              %headshake-->l.20

18 GEO: [$JA- #des is mal- ]
        well that is PART

19 NIL: [$he  #°h hast du_s]  m§al geS[A $GT?]
           did you say that at some point?
   nil  $shoulder&hand shrug->$grasps table-->>
   geo        #fig1a,b
```


fig1a

fig1b

```
20 GEO:                    [ne$ih ]+%ein.
                            noho
   geo                   -->$gz down-->l.21
   zoe                     -->+gz at GEO's
                              hand-->l.21
                            -->%

      [  °h$h                ]
```

```
21 NIL: ¢[des$ wär eiglich die p]er+fEkte ANTwort.
            that would be PART the perfect answer
     geo    -->$
     zoe                         -->+gz at NIL-->>
22         (2.96)
```

In the side sequence (lines 01–13), George clarifies that even though he had negatively assessed the hairdresser as stupid prior to this excerpt, which led to Zoe's dispreferred response at line 01, he prefers this type of women over intelligent women (lines 04–06). While his male friend, Nils, responds to this concession with affiliative laughter (line 07), Zoe attempts to cut this slightly macho positive assessment of unintelligent women short by interrupting George with repeated *is in Ordnung* ('it's okay'; line 08). George subsequently aligns with Zoe's orientation toward a closure of the side sequence, which revealed that Zoe and George do not agree on judging female hairdressers as stereotypically unintelligent. At lines 09 and 11, George closes the side sequence with *is ja auch egal* (idiomatically 'never mind,' lit. 'doesn't PART PART matter'), which has been shown, in English, to be a practice for "[accomplishing] sequence and topic closure in part by proposing a summarizing assessment, the matter is not worth concern ('Don't let it trouble you'), but also by displaying prosodic disengagement" (Couper-Kuhlen 2004: 213) in complaints or trouble sequences. George's closing of the side sequence is successful even though the prosodic pattern does not fit Couper-Kuhlen's (2004) description. The continuative final prosody projects "more to come," and George in fact continues the larger sequence about the hairdresser asking him how he wants his hair cut.

During the subsequent pause, George gazes down (line 13), which may be an early indicator for his upcoming negative affective stance (lines 14–16).[206] At line 14, George returns to the larger complaint sequence, skip-connecting (Sacks 1992) his subsequent direct reported thought back to his talk prior to the side sequence (not in transcript) with *und* ('and'; line 14).

With *deng_isch mir* ('I think to myself'), he projects his upcoming purported thought from his visit to the salon, which is, in this case, an imperative directed at the hairdresser. The imperative contains two parts: the negative assessment of the hairdresser's question (line 15) and George's (potential) indications of how she should cut his hair (line 16). Note that the quotative is in V1, thus accelerating the tempo of the telling while approaching its climax. The quote is prosodically separated from the quotative by a micro pause (line 14), thus increasing the drama of the upcoming twist in the story.

The imperative is not syntactically attached to the quotative. If it were syntactically linked, the imperative would have to be reformulated to a clause with the modal verb *sollen* ('should'). The lack of subjunctive mood and deictic shifts in the quote points toward its status as a dependent main clause. However, because the empty argument position of *denk ich mir* is filled by two intonation units containing two imperatives, a syntactic linkage is impossible. One way to explain the social action format projecting a larger unit than one clause is to analyze the whole pattern as a bipartite structure in which the quotative functions as projector-construction

---

[206]   The gaze down to yield the turn (see Excerpt 7.3) is probably not the case here. The side sequence has just been closed by George (line 11) and by Zoe (line 12). Because George has paused his telling (not in transcript) after Zoe's assessing his prejudices, it is probably clear to Zoe that George is going to continue his telling.

that introduces a unit exceeding the size of one clause. The additional prosodic separation between quotative and the following larger spate of talk may support the projective force of the quotative, thus demonstrating the emergent nature of this bipartite structure: the social action format prototypically projects one clause, which is then adapted to the ongoing speech situation—here, by adding a second imperative.

The following quote allows George to display his negative affective stance toward the hairdresser's question ('how do you want your hair cut') while also implying that he finds her attractive (line 16). While reporting his purported thoughts, George gazes at Nils (line 15), then down during the potentially controversial part (line 16), and then briefly at Zoe, thus establishing a moment of mutual gaze (line 17). Zoe, however, shifts her gaze toward Nils, as does George shortly after, at the end of Nils's affiliative laughter (line 17). These quick gaze shifts may indicate that both participants, Zoe and George, orient to their disagreement on the matter by avoiding mutual gaze. The subsequent turn confirms this assumption.

Zoe responds to George's affective stance-taking with a long headshake (lines 17–20), a skeptical facial expression (fig1a and fig1b), and a head tilt (Debras & Cienki 2012), pursed lips, and a gaze at her boyfriend, Nils, as if seeking his affiliation. Nils, instead, laughs loudly (line 17) and asks an information question, thus disattending George's complaint (Mandelbaum 1991), and circumventing the projected affiliation. With his question (line 19), Nils orients toward the fact that George displaying reported thought with the social action format *ich denke* may not mean that he has not said the thought out loud some other time (expressed through *mal* ('at some point'), line 19). The defensive prosody and the doubling of the one syllable of *nein* ('no') to two syllables by adding the *h* in the middle (*neihein* ('noho')) allow George to display that his direct reported thought has not been originally designed to be pronounced. Nils's positive assessment of George's response, even if it was said out loud (line 21), closes the sequence and demonstrates Nils affiliation with George's stance.

This excerpt has shown how a speaker may deal with two different responses to an affective stance display with direct reported thought. In contrast to Excerpts 7.2–7.4, in Excerpt 7.5, it is not the prosody or bodily conduct of the speaker but the reported thought itself that is used to display an affective stance through lexical choices (*dumm fragen*, 'to ask stupidly'; line 15) and the speaker's action of transforming a professional relationship into a potential dating situation. This excerpt also differs regarding the placement of the direct reported thought within the telling sequence: the pattern [*ich denke mir* + direct reported thought] constitutes the climax itself instead of commenting on the climax.

Excerpt 7.5 is one of the rare instances in which direct reported thought introduced by *ich denke (mir)* is part of a reasoning (*n* = 5). The same three interactants from Excerpt 7.5 are talking about prejudices. Zoe says prior to this excerpt that everyone can have prejudices but that every individual about whom one has prejudices should be given the chance to prove one wrong. Nils thus initiates a telling to convince his interlocutors that—at least, concerning Indians at his university—prejudices are right and generalizable. The sequence does not follow the typical structure of a storytelling; instead, Nils reports an inner monologue that is nurtured by external events that happen during his typical day.

Ex. 7.5: drei inder / three indians (FOLK_E_00293, 1h37min20)

```
NIL: Nils, ZOE: Zoe, GEO: George

01 NIL: $¢wenn ich jetz morgens AUFstehen würde un würde
        if I PART got up in the morning      and thought
   zoe  $gz at NIL-->l.09
   geo   ¢gz at NIL-->l.19

        mir denken,
        to myself

02      °h ↓ja +eigentlich (.) sin die inder bei
           yeah actually     the indians at our
   nil        +gz up-->l.03
        uns an der Uni-
        university

03      +eigentlich sin die wie WIR.+
         actually they are like us
   nil  +gz at GEO---------------->+closes eyes,gz up-->l.06

((three lines omitted))

07      dann sitz ich im BUS,
        then I sit in the bus

08 ZOE: he [he ich ha °h $(glück) noch]
        hi  hi I am (luckily)

09 NIL:    [scho ma    di$REKT,       ]
            already PART directly
   zoe               -->$gz at GEO-->l.11

10 ZOE: kein +INder [erwi   ]$scht.    ]
        no Indian yet
   nil    -->+gz at ZOE-->l.13

11 NIL:          [scho ma] $direkt d]rEI
                  already PART directly three
   zoe               -->$gz at NIL-->l.24

12 GEO:          [j(h)a. ]
                  ye(h)es

13 NIL: inder um mich +RUM, (.)
        indians around me
   nil             -->+gz at GEO-->l.14

14      stinkt nach +CURry.#   +
        stinks of curry
   nil             -->+gz at ZOE+gz left-->l.15
   fig                      #fig1
```


fig1

```
15      >denk_(i)sch mer schomal<+% ↓JA::- °h
         I think PART to myself PART yeah
```

```
     nil                        -->+gz up-->l.17
                                  %mvs head l to r->l.16
```

16   dafür <<all>können die ja% nich unbedingt WAS.>
*   they can't necessarily do PART anything about that*
```
      ----------------------->%
```

17   (.) dann +(.) bin ich in der VORlesung-#
*   then   I am in the lecture*
```
           -->+gz down-->l.20
```
fig   **#fig2**

fig2

((6 lines ommitted where Nils describes another alleged 'characteristic' of Indians during the lecture))

24   °h $dann# fra¢gen se$ no∫ch ma irgendne behinderte
*   then they ask yet another PART random stupid*
```
   zoe  -->$down,l--------->$gz at NIL-->>
   geo           -->¢gz front,l-->l.34
                           -->∫
```
fig   **#fig3**

fig3

      FRAge,=
*   question*

25   =die der prof vorher +auch schon erKLÄRT hat,
*   that the professor also already explained before*
```
                           +gz at ZOE-->l.27
```

26   °h %ja.   %
*   yes*
```
   nil     %head back%forward-->l.26
```

27   °h +%dann (.) äh% fährste wieder HEIM,#
*   then   uh you go home again*
```
   nil  -->+gz down-->l.28
           %headshake %
```
fig   **#fig4**

fig4

```
28      hast WIEder des gestInke +von @dEnen,
        have again the bad smell of them
   nil                       -->+gz tw ZOE-->l.34
                                    @.....

29      un dann denkst@e   @DIR,        @
        and then you think to yourself
   nil  ..............@h up@h falls down@

  30      h° <<all>wenn de abends heimkommst>,
              when you come home in the evening

31      ↓ja eigentlich sin die inder DOCH scheiße.#
         yes PART the indians are indeed shitty
   fig                                          #fig5
```


fig5

```
32      (0.55)

33 GEO: [°h+h  ¢ ((smack)) °ich i i i°           ]
                              I

34 NIL: [un+ un¢ des is nich ¢immer der GLEI¢che] inder
          and and it is not always the same indians
   nil  -->+gz down-->l.35
   geo      -->¢down-------->¢gz at NIL---->¢down-->l.36

        den ich da seh,
        that I see there

35      des sin ALle +i[nder. ]
        that's all indians
   nil             -->+gz at ZOE-->>
```

Nils frames his telling as reporting hypothetical events using an initial *when*-clause (line 01) and subjunctive mood (*würde* 'would'; lines 01–02). The telling sequence covers a typical university day for the speaker: he gets up in the morning (line 01), takes the train (line 05; not in transcript) and then the bus (line 07), sits in a lecture at university (line 17), drives home

(line 27), and arrives at home (line 30). These time slots are filled with descriptions related to whether Nils meets Indians and how their conduct alters or confirms his prejudices.

Starting (in the morning) with the hypothesis that Indians at his university *eigentlich sin die wie wir* ('are in fact like us'; line 03), Nils gazes at George, who, contrary to Zoe, shares Nils's negative opinion about Indians as he has previously displayed (not in transcript). In overlap, George confirms Nils's description with emphatic prosody, probably because, so far, Nils's telling matches George's project to reduce his prejudices (not in transcript). The next part of the telling, however, already implies Nils's negative stance: He frames not having crossed any Indians on the train yet as "luck" (lines 05–06; not in transcript).

During the next part of his telling, when the membership categorization (Stokoe 2012) and prejudice of Indians smelling of curry arises (lines 07–14), Zoe is still orienting to Nils's telling as funny and responds to Nils's formulation at line 06 with laughter (line 08) and repetition (line 10). She holds her smile even during the strong negative assessment of the Indians' smell (fig1), which illustrate to Nils how prejudices can also be correct. In the next part of the telling follows a direct reported thought that allows Nils to downgrade his negative stance in the described bus scenario and position himself as someone who can rise above such things by stating that *dafür können die ja nich unbedingt was* ('they can't necessarily do PART anything about that'; line 16). The overall structure of the telling sequence shows that the direct reported thought serves as a concessive turn (Couper-Kuhlen & Thompson 2000), as a stepping stone toward the upshot of the telling.[207] According to Couper-Kuhlen and Thompson (2000),

> the speaker doing the conceding is doing so in the context of going on to make his/her own potentially contrasting point. Thus, 'conceding' can be understood as one way in which speakers can deal with expressing 'contrast'. (*ibid.*: 382)

Nils's conceding is marked with the particle *schomal*, indicating that this is just one side of the coin, and the prosody of *ja* ('yes'; line 15), which is lengthened and pronounced with a low pitch movement, thus expressing his (possibly reluctant) tolerance of Indians smelling of curry. The account for his concession is subsequently verbalized at line 16 and accompanied by Nils swaying his head, displaying his reluctance through embodied means.

Nils continues his telling with the third part of his concession—namely, the "contrasting point," which he elaborates on with several exemplar situations during the lecture at university (lines 17–25) and the description of his ride home (lines 27–28). He closes his reasoning with the conclusion that, because he does not always see the same Indian in these described situations (line 34), this must apply to all Indians (line 35). Note that the direct reported thought introduced by *denken* in second-person singular may allow Nils to increase intersubjectivity by inviting his interlocutors more explicitly to join his negative affective stance.[208] According to Auer and Stukenbrock (2018), such uses of *du* ('you') can have several functions, such as "[referring] to a collectivity of people to whom both the speaker and the addressee belong.

---

[207] Even though this concession is only produced by one speaker (instead of two, as described by Couper-Kuhlen & Thompson 2000) the pattern functions similarly—namely, to preface a disruptive disagreement.

[208] A similar phenomenon has been described by Dostie (2020) for the French *t'es là* ('you are there').

Usually, this occurs in statements that formulate rules, truisms, wisdoms, etc." (*ibid*.: 306)—here, *eigentlich sind die inder doch scheiße*[209] ('PART Indians are PART shitty') at line 31.

This excerpt illustrates that direct reported thought introduced by *ich denke (mir)* can also be used in reasonings. The speaker deploys the pattern in a different sequential position, namely the middle of the sequence, as part of a concessive structure. With *denke ich mir schonmal*, Nils introduces an appeasing assessment framed as direct reported thought, thus commenting on his prior assertion. During the quote, the speaker neither displays a high degree of affectivity nor depicts his purported thoughts, as is the case in Excerpts 7.2–7.5. Probably due to the argumentative character of Nils's turn, the speaker's bodily conduct also remains subtle: He gazes up and moves his head left to right, displaying through embodied means his reflecting on two sides of a matter. These kinds of gestures can contribute, according to (Kendon 2004), "to the referential meaning of the utterance of which it is a part" (*ibid*.: 197). Of course, such gestures have a multitude of forms. Nils's head movement and gaze fit Kendon's (2004) description in that they "supply components of meaning integral to the utterance of the moment" (*ibid*.: 198)—here, weighing two points of view against each other.

7.2.4.   Summary

In the last Section, I have analyzed the social action format *ich denke (mir)* introducing direct reported thought in tellings. I have shown that the quotation of purported thought is primarily used in storytellings to display an affective stance toward prior events (90.4% of all occurrences), with few exceptions where speakers make publicly available their emergent reasoning (9.6%). Speakers seem to use the pattern [*ich denke (mir)* + direct reported thought] primarily in the closing environment of a telling (Excerpts 7.2–7.4), with only few exceptions where the social action format occurs within the telling (Excerpt 7.5).

The different sequential positions of the pattern [*ich denke (mir)* + direct reported thought] make different responses from the recipient(s) relevant. Whereas direct reported thought in the closing-environment of a telling calls for affiliation, ideally from all recipients, the mid-telling position seems to make alignment or the display of access to the telling relevant (Excerpt 7.5). I have shown that, if an affiliation is not delivered in response to direct reported thought at the end of a telling, speakers work toward this preferred response (Excerpt 7.3). This is also the case for a second recipient when a first recipient has already affiliated (Excerpt 7.2).

Compared to instances in past tense, the narrative present tense of the social action format moves the reported thought into the here-and-now of the ongoing telling, thus emphasizing its degree of dramatization and creating its narrative density (Günthner 2006). The tellings are either entirely told in present tense, thus increasing the directness of the scene ($n = 34$), or the speaker shifts to present tense for only the reported thought ($n = 13$).[210] In the remaining three cases, the speaker shifts to present tense for the adjacency pair of a reported first pair part to which the reported thought responds as a second pair part within the telling. In those instances in which speakers shift from past to present tense with the quotative, the shift in footing allows

---

[209]   *Doch* functions here as an adverbial and not as a modal particle because it constitutes a part of speech and could be moved to the front field (Diewald & Fischer 1998: 78; see also Imo 2007: 52ff). I still marked it only as "PART" because of its difficult translation into English.

[210]   In two additional cases, what is quoted is in conjunctive mode or does not contain a predicate.

the speaker to make the direct reported thought immediate by moving only the pattern [*ich denke (mir)* + direct reported thought] into the here-and-now of the ongoing telling whereas further parts of the telling remain more "distant" in past tense. This is in line with research on medieval narratives in which narrators already marked "shifts in footing" through tense-shifts, thus commenting on their tellings in present tense while addressing their public (Fleischman 1986; Zeman 2010, 2013). By doing so, *ich denke (mir)* introduces an affective stance display that is not only valid in the respective telling but "reaches" into the ongoing conversation, thus creating intersubjectivity and, at the same time, increasing the pressure on the recipient(s) to respond. In some instances, this configuration may suggest that the affective stance display is actually not a report of past thoughts. This concerns especially Excerpts 7.2 and 7.3, where the quotative is prosodically attached to the response cry *hä* ('huh'). These quotations may be simply an exaggerated depiction of a past stance without actually quoting anything; however, since the attribute "reported" does not necessarily claim to reflect what was actually thought, these cases still fall under the formal description of direct reported thought, because this is what speakers display.

Concerning (clause-)combining, my analysis has shown that quotes are mostly syntactically independent from *ich denke (mir)*. In those cases where the quote consists of a complete clause, the latter fulfills all formal characteristics of a dependent main clause structure. Consequently, two patterns, both forming a bipartite structure are possible: (1) When the quote is a dependent main clause, *ich denke (mir)* functions as a main clause; however, the degree of embeddedness of the dependent main clause (= the quote) is low (see Chapter 6 for a detailed explanation). (2) When the quote does not have the format of a clause or when the quotative first introduces a response cry that is only then followed by a clause, the quote cannot be categorized as dependent main clause. This makes any syntactic bind between *ich denke (mir)* and the quote impossible and leaves only a pragmatic relation between the two where the quotative functions as a mere projector-construction (Günthner 2008b) introducing quotes of different shapes.

As with the past forms, Brünner's (1991) suggestion of calling quotation *Redesituierung* ('speech situating'; see Sections 6.2.3.1 and 6.2.3.2) may be a possible conceptualization that allows an inclusion of response cries and body movements into a grammar of reporting formats in spoken language, as Keevallik (2018) has suggested. It also allows for an attribution of projective force to the social action format. My analyses have shown that *ich denke (mir)* still grammatically foreshadows "more to come" because it projects an open argument position that is expected to be filled. In talk-in-interaction, other than in written language, what fills the empty slot can consist of a variety of formats ranging from body movements to full clauses.

As a result of my qualitative analysis, I was able to retrieve the speaker's multimodal conduct while quoting.[211] The following table shows the frequencies of body movements and gaze behavior that I observed during the pattern [*ich denke (mir)* + direct reported thought].[212]

---

[211] As I have already pointed out in Chapter 6, these numbers must be treated with caution due to the diversity of data that my German corpus comprises.

[212] The total number of occurrences below does not add up to 52 because some of these body movements co-occur. Recurrent co-occurring assemblies have not been detected with the present-tense quotative *ich denke (mir)*.

Multimodal conduct	Frequency
speaker's gaze at (one of the) interlocutor(s) at the end of the quote[213]	55.8% (29/52)
eyebrow lift	3.8% (2/52)
facial expression	30.8% (16/52)
head movement left and right	1.9% (1/52)
dropping hand gesture	3.8% (2/52)
palm-up open-hand gesture (PUOH)	13.5% (7/52)
shrug	13.5% (7/52)
headshake	3.8% (2/52)
iconic gesture	9.6% (5/52)
change in posture	3.8% (2/52)
no specific body movement / not visible[214]	23.1% (12/52)

Table 35: Multimodal conduct co-occurring with [*ich denke (mir)* + direct reported thought]

Similar to [*ich dachte (mir)* + direct reported thought], facial expressions are recurrent with the social action format in present-tense (30.8%). Shrugs and palm-up open-hand gestures co-occur 14 times with direct reported thought introduced by *ich denke (mir)*, which corresponds to 27.0%. The third most frequent body movement are iconic gestures. I used the term "iconic gestures" as umbrella term to subsume different depicting gestures that correspond to Streeck's (2009) category of *mimesis*, or gestures that depict action, such as taking an object back from an (imaginary) person.

Facial expressions are used in about one third of the direct reported thought introduced by *ich denke (mir)*. Section 7.2 has demonstrated that these facial expressions allow speakers to display their affective stance not only economically but also *while* they quote their thought verbally. In this sense, the actions of quoting and taking an affective stance are, according to Goodwin (2013), *laminated actions* that are "organized as layers of diverse resources" (*ibid*.: 12). This result emphasizes the importance of a multimodal analysis because it shows that two actions can emerge in parallel by being distributed between resources—an observation that has not received much attention yet.

Thus far, I have not yet addressed whether *ich denke (mir)* in narrative present tense must be described as a separate practice compared to its past forms. I argue that, to a certain extent, this distinction may be useful. As my qualitative analysis and my quantifications have demonstrated, *ich denke (mir)* is primarily used for affective stance-taking. It rarely occurs as a social action format to display an epistemic stance like *ich habe (mir) gedacht* does. As a consequence, *ich denke (mir)* primarily shows parallels in its use with *ich dachte (mir)*: concerning their clause-combining formats, the actions that speakers implement with them, and the frequent use of facial expressions during the quote. One difference between *ich denke (mir)* and *ich dachte (mir)* as quotatives, however, remains: Using a quotative in present tense may lead to a stronger effect of immediacy of the affective stance-taking in the here-and-now of the ongoing conversation. Through this effect of immediacy, speakers may create more

---

[213] Nine occurrences are not visible due to the camera angle. In two additional instances, the speakers are cooking or driving and are thus unable to look at their recipient(s).

[214] This relatively high number is due to my German data, in which speakers are often involved in other activities, like driving a car or cooking (see Chapter 5).

intersubjectivity while increasing the pressure on their interlocutors to respond to their affective stance-taking.

The next section will focus on direct reported thought in French with quotatives in present tense. My quantifying and qualitative analysis will treat *je me dis* ('I say to myself') and *je suis là* ('I am there') in more detail. I will show that, similar to their past tense format, speakers use *je me dis* and *je suis là* for distinct practices in talk-in-interaction.

7.3.   Reporting Thought with the French *Je me dis* ('I say to myself') and *Je suis là* ('I am there')

7.3.1.   The Present Tense-quotative *Je me dis* ('I say to myself')

In this section, I analyze the social action format [*je me dis* + direct reported thought]. Whereas Section 6.3.1 illustrated how direct reported thought is projected the same social action format in past tense, the present section focuses on instances in which the quotative is in narrative present tense. The morphological structure of the quotative is as follows:

je	*me*	*dis*
PP.1SG	myself-REFL.1SG	say-PRS.1SG

'I say to myself'

The reflexive pronoun is thereby always placed between the personal pronoun *je* and the conjugated verb.

7.3.1.1.   A General Picture: A Quantifying Analysis

Among the 28 occurrences in which *je me dis* ('I say to myself') is used to introduce direct reported thought in my data, only few show signs of morpho-phonological reduction. Table 36 shows the frequency with which different phonological features and characteristics of morpho-phonological reduction occur:

Feature	Frequency
accelerated	39.3% (11/28)
no accentuated syllable	21.4% (6/28)
pause after *je me dis*	28.6% (8/28)
omission of one or both of the final 'e's in *je* ('I') and *me* ('myself')	10.7% (3/28)
omission of the personal pronoun *je* ('I')	3.6% (1/28)

Table 36: Prosodic and morpho-phonological features of *je me dis*

It is specifically interesting that the present form is less often accelerated than the past form (see Section 6.3, Table 19) and that the conjugated verb *dis* is more often prosodically stressed (78.6% (22/28) compared to 24.4% (11/45) for *je me suis dit*). There also seems to be a higher tendency for a pause to follow the quotative in present tense (28.6% (8/28) compared to 13.3% (6/45) for *je me suis dit*). The recurrent stress on *dis* could be an indicator that the quotative still carries semantic and pragmatic weight on its own instead of pragmatically foregrounding the quote. Fourteen of 28 cases of *je me dis* are preceded by additional lexical items. Items that co-occur more than once are shown in the following table:

Form	Frequency
pis ('then')	10.7% (3/28)
mais ('but')	7.1% (2/28)
donc ('so')	7.1% (2/28)

Table 37: Particles co-occurring with *je me dis*

The table shows that the present tense is rarely preceded by lexical items that connect the direct reported thought to what precedes it. One reason for this may be that [*je me dis* + direct reported thought] is less often embedded into longer telling sequences.

7.3.1.2.   Syntactic Features of *Je me dis* ('I say to myself')

Generally, the verb *se dire* is the reflexive form of a transitive verb that calls for a compulsory direct object. When *je me dis* introduces *indirect* reported thought, the quote fills the empty argument position of the predicate (i.e., of the main clause). When *je me dis* introduces *direct* reported thought, the quote is, however, syntactically independent (i.e., it is not syntactically linked through the complementizer *que* ('that')).[215] In this case, the quote cannot have the grammatical status of a complement. The prototypic syntactic configuration of [*je me dis* + quote] is thus the juxtaposition of two clauses, which form a bipartite structure without syntactic link, as in the following excerpt:

Ex. 7.8: la laisser partir / let her go (Pauscaf_02_05, 22min40)

```
14 MAR: [°h pis je me DIS- ] (--)
        and I say to myself
15      il a de !TOUTES! manières pas envie d'la laisser parTIR;
        he doesn't want to let her go anyway
```

The quotative, *pis je me dis* ('and I say to myself'), is the main clause (line 14), which introduces the quote *il a de toutes manières pas envie d'la laisser partir* ('he doesn't want to let her go anyway'; line 15) without linking it syntactically. Because the verb *dire* still misses its—obligatory—direct object at the end of line 14, the quotative grammatically projects more to come. The subsequent pause may indicate that the recipient orients to this grammatical projection by not taking the turn at this point. Even though what follows at line 15 is not a syntactically linked complement (i.e., a direct object), the grammatically projected open slot is nevertheless filled—namely, by a syntactically independent element. This element may be a clause, but it may also be a response cry or a body movement. Notwithstanding which elements speakers introduce with *je me dis*, there is, as I have already pointed out in Chapter 3, a pragmatic relationship between the quotative and quote.

---

[215]   For a detailed discussion of the syntactic status of the quote, see Chapter 3. Because the syntactic features of *je me dis* as a quotative are the same as for its past form *je me suis dit*, I simply refer here to Section 6.3.1.2 for further syntactic features and a more detailed discussion.

In my corpus, 20 of the 28 cases are followed by a complete clause.[216] In another 4 instances,[217] *je me dis* is followed by a complex syntactic pattern. In the following excerpt, for instance, the second part of the bipartite structure is a list:

Ex. 7.9: au fitness / to the gym (Pauscaf_02_09, 09min30)

```
QUE: Quentin, MAT: Mathis

01 QUE: dans dans une de mes révisions où j'avais Un peu
        during during one of my study times where I had a bit

        plus de TEMPS,
        more time

02      j'avais essayé de,
        I had tried to

03      je me dis !TOUS! les JOURS, (0.3)
        I say to myself every day

04 MAT: mhm-=

05 QUE: =à: !SEI!ze heure seize heure dix-huit heures je
         at four o'clock four o'clock six o'clock I

        vais au fitNESS-
        go to the gym

06      après je mange-=
        then I eat

07      =>et après< je rebosse un petit COUP >tu vois<,
         and then I study again a little bit you see
```

In this excerpt, the quotative is immediately followed by the temporal adverbial *tous les jours* ('every day'; line 03). The quote continues after a brief pause and a continuer from Mathis (line 04). The speaker, Quentin, informs his interlocutor about his studying day with three subsequent main clauses, of which each clause comprises one event: going to the gym (line 05), eating (line 06), and working in the evening (line 07). *Je me dis* introduces these syntactically linked clauses through the adverbs *(et) après* ('(and) then'; lines 06–07). Syntactically, what follows *je me dis* is a complex syntactic pattern that cannot constitute the complement of *je me dis* simply because a complex clause cannot occupy an empty argument position. This is why I conceptualize the pattern [*je me dis* + quote] as a bipartite structure.

The following table shows other formats of quotes in my data in addition to complete clauses (*n* = 20) or complex clauses (*n* = 4):

---

[216]　These 20 cases include four instances in which the clause is preceded by a response cry as well as two questions with main clause syntax. In French, to be subordinated, the question (without a question word) would need to be reformulated to a *si*-structure (*if*-structure).

[217]　All the numbers do not add up to 28 because the two questions are counted twice, once as main clause syntax and once in the category question (word) in Table 38.

Quotative	Response cries only	Bodily conduct only	Verbless structures	Question (words)	Infinitives
je me dis	1	0	0	4	1

Table 38: Other formats of direct reported thought introduced by *je me dis*

According to these numbers, *je me dis* tends less frequently to introduce other formats than complete clauses. Even compared to its past tense *je me suis dit*, *je me dis* is followed less by response cries or bodily conduct.

The delimitation of the two (grammatical) units that quotative and quote comprise is not always simple, in particular because of elements inserted between them. Similar to the past form, *particules d'amorces* ('initiating particles') are inserted between the quotative and the quote (*n* = 11). The following table shows these items:

Form	Frequency
tu vois ('you see')	7.1% (2/28)
'fin ('well')	3.6% (1/28)
euh ('uh')	3.6% (1/28)
mais ('but')	14.3% (4/28)
oh	3.6% (1/28)
ben ('well')	3.6% (1/28)
attends ('wait')	3.6% (1/28)
TOTAL	29.3% (11/28)

Table 39: *Particules d'amorces* occurring between *je me dis* and direct reported thought

The table demonstrates that initiating particles occur in about one third of all occurrences. The detailed sequential analysis in the following section shows that these particles may be used to mark the beginning of the quote and therefore a shift in footing.

In what follows, I will analyze four excerpts to show the interactional use of the pattern [*je me dis* + direct reported thought].

7.3.1.3. Interactional Functions of *Je me dis* ('I say to myself')

With the following four excerpts, I will illustrate that the social action format *je me dis* is used in sequences of reasoning (see Section 3.2.3). As a reminder, these sequences do not share many features with classic storytellings in which speakers report past experiences dramatically; instead, speakers share a (mostly rational) thinking process that leads to a certain decision or conclusion, often without much bodily animation or depiction. The social action format *je me dis* allows speakers to make rationalizations publicly available.

The first excerpt shows the most frequent pattern of use. [*Je me dis* + direct reported thought] occurs in the closing environment of a sequence in which a speaker makes the reasoning behind an opinion publicly available. In this specific case, Marie explains two sides of an argument, which she weighs against each other when thinking of her professional and personal future. This narration is uttered in response to Patrick saying he is career-oriented.

Ex. 7.6: ça me stresse / it stresses me out (Pauscaf 13, 09min40)

```
PAT: Patrick, MAR: Marie

01 PAT: =chais pas moi je::: (0.8)
          dunno I I

02      j'avance plutôt sans trop m'poser de quesTIONS,
        I rather move forward without asking myself too many
        questions

03      (0.6)

04 PAT: >mais< j'vise plutôt une carriÈRE que:-
         but  I rather aim at a career than

05      >de m'dire< AH purée i faut qu'tu trouves quelqu'UN:-
         telling myself oh damn you have to find someone

06      t'as ving- t'as bientôt$ [vingt-six] ANS$ euh:-
        you're twen you're almost twentysix years old uh
   mar                          $gz left-------->$gz at MAR-->l.09

07 MAR:                        [alors-   ]
                                well

08      (0.3) moi j'suis vraiment entre les DEUX en fait;
              for my part I'm really  between the two of them PRT

09      (0.2)$(0.2)
   pat     -->$gz down-->l.10

10      j'suis pAs là à$ fONd en train de viser la carriÈRE,
        I'm not   there totally focussing on aiming for a career
   pat              -->$gz at MAR-->l.15

11      (0.7)

12      mais j'ai quand même envIE d'avoir un beau méTIER,
        but  I would like to have a nice job though

13      (0.2)

14      pis j'suis pas là: non plus: 'fin: OUI# +j'me %dis#:
        then I'm not either there    well yeah I say to myself
   mar                                      +gz at left->l.15
                                              %eyebrow
                                              raise->

   fig                                    #fig1        #fig2a,b
```


fig   fig1

fig2a

fig2b

```
        ça m'stresse et %tOUt:,=
        it stresses me out and all
   mar              -->%
15      = +mais: sans plus non ↑PLUS$ %'fin: voilÀ?
        but not that much neither PRT PRT
   mar  ->+gz at PAT—>l.17
                                   %shoulder raise-->l.17
   pat                     -->$gz down-->l.17
16      (0.3)
17 PAT: mais si on te pro%pose $d'aller faire un job à l'étranger
        but if you get an offer to leave for a job abroad
   pat                     -->$gz at MAR-->>
   mar                 -->%
        euh: l'année prochaine tu PARS?
        uh   next year would you go?
        (0.3)+(1.0)
   mar    -->+gz left-->>
18 MAR: chais pas;
        dunno
```

After Patrick's tentative[218] explanation of how he imagines his future, Marie starts her inner monologue early with a pre-beginning in overlap (line 07). She first clarifies that she is torn between career and having a family (as mentioned in what precedes this excerpt; line 08). After a short pause (line 09), she delivers an account for this inner conflict. To illustrate her reasoning *ex negativo*, she uses twice the pattern *je suis pas là* ('I'm not there'). In a list structure with its typical prosody (Selting 2007), Marie enumerates the different arguments: She is not exclusively focusing on a career (line 10), but she wants a nice job (line 11). The continuation of her list is interrupted and self-repaired from a negative formulation with *je suis pas là* ('I am not there') to a positive formulation with the social action format *je me dis* (line 14) that projects her conclusion.

The subsequent pattern is syntactically parallel to lines 10 and 11 with two clauses connected with *mais* ('but'): *ça me stresse et tout mais sans plus non plus fin voilà* ('It stresses me out and all but not that much neither PART PART'; lines 14 and 15). Note that the format is preceded by a prosodically stressed *oui* ('yes'), which is accompanied by a short eyebrow raise (see shift between fig1 and fig2a/fig2b), allowing Marie to emphasize (Ekman 1979) what she verbalizes

---

[218]   He initiates his turn with the epistemic parenthetical *chais pas* ('dunno'; line 01; Pekarek Doehler 2016, 2019) and uses hedges like *plutôt* (rather; lines 02 and 03) and longer pauses (lines 01 and 03).

in the following direct reported thought.[219] During her quote, Marie starts gazing left (line 14), just to gaze back at Patrick at the end of her utterance (line 15). This gaze away could indicate embarrassment while admitting that she is stressed because she needs to decide between family and career.

Before closing her turn at line 15, Marie uses the turn-closing particles *'fin voilà* (Hailaselassie 2015) accompanied by an "attitudinal shrug"[220] (Debras 2017)—an assembly which may imply that there is nothing that she can do about this dilemma. While lines 08 to 14 allow Marie to present her point of view "as having been based on a rational decision-making process" (Bangerter *et al.* 2011: 208), [*je me dis* + direct reported thought] allows her to present "the consequences of that rationalization" (*ibid.*), which she displays through bodily means (a shrug) and verbally (line 15). Note again that the sequence in which the direct reported thought occurs does not report a past situation where a described event affects the speaker emotionally. Instead, Marie makes publicly available her personal reasoning about a state of affairs that is the product of her own thinking process.

Patrick's subsequent response may be foreshadowed by his gaze down before the closing particles (line 15). With his question at line 17, he also challenges Marie's reasoning: He displays his disbelief of her prioritizing a job over family—an assumption that is partially confirmed by Marie's dispreferred *chais pas* ('dunno'; line 18; Pekarek Doehler 2016).

In this excerpt, Marie's multimodal conduct during her direct reported thought is not a means to depict prior events through multiple resources, as is mostly the case with *être là*. Instead, through the eyebrow raise and the shrug, she seems to topically focus a certain part of the utterance (Krahmer & Swerts 2004) while moving toward closure with her quote (line 15). In this excerpt, Marie's turn closure is successful and her reasoning, even though subsequently challenged by Patrick, is complete.

Excerpt 7.7 is similar to Excerpt 7.6, only that the reported-thought format makes publicly available a decision that may affect the speaker's immediate future. This instance comprises several of the body movements recurring with *je me dis*: a palm-up open-hand gesture, a shrug, and a gaze at the recipient at the end of the quote. Lise has just told Thomas about her finally having found a place to live. Because of certain circumstances which are not entirely clear from context, Lise is happy to move even though the new apartment is not perfect:

Ex. 7.7: on cuisine dehors / we cook outside (Pauscaf 4, 5min00)

```
LIS: Lise, THO: Thomas

01 LIS: pis pour la ventilaTION-
            and for the ventilation

02      >tu vois< +c'est ça$ le proBLÈME-=
          you see  that's the problem
    lis        +gz at THO-->l.04
    tho            $gz at LIS-->l.09
```

```
03      =c'est que vu qu'%il y a pas une %cuisine (x)-
         it's that because there is no kitchen (x)
    lis                  %...............%PUOH,beat..-->l.04

04      %eh     j%e me dis on+:: %+au @PIRE on- +(0.6)+
        well I say to myself we worst case we
    lis %throwing
         away
         gesture %,,,,,,,,,,,,,,,%
                       -->+right+at THO----->+left +right
                                                    -->l.05
                          @shoulders up..->

05      pendant les +premiers mois on cuisine de@HORS hein-@#
        during the first months we cook outdoors huh
    lis          -->+gz at THO-->l.12
         ........................................@shrug---->@

06      (0.7)

07 THO: hh°# <<:-)>ça va être freeSTYLE>; h°
                  it's going to be freestyle
    fig     #fig1
```


fig1

```
08 LIS: mais moi je m'en ↑FOUS franchement.=
        but I couldn't care less honestly

09 THO: =>ouais ouais ouais-<$
          yeah yeah    yeah
    tho                    -->$turns l,gz l-->l.11

10 LIS: [j'étais prête %à aller  vi$vre au  cam%PING-]
         I was ready to go live at the camping site

11 THO: [t' as raison  %c'est juste$ c'est  jus%te    ] deux mois
         you're right   it's just    it's just        two months
    tho                          -->$gz at LIS-->l.16
    lis             %pointing right-------->%

        et tout euh:=
        and all uh

12 LIS: =mais# @↑OUAIS,@+
         of course yeah
    lis         @shrug  @
                    -->+gz right-->l.16
    fig     #fig2
```

221

fig2

```
13 THO: voilà.
        PART

14       (0.9)@(0.5)@(0.2)
   lis       @shrug@

15 LIS: 'fin  [+rien à BATtre; ]
        well   I don't care

16 THO:        [+et je je m'y fe]rais très $bien ausSI ouais;
                  and I I would be very fine with that too
   tho                                   -->$gz left-->>
   lis        -->+gz at THO-->>
```

In lines 01 to 03, Lise presents the problem with her new apartment: There is no ventilation in the kitchen. The consequence of these circumstances is already grammatically projected by the formulation *vu que* ('because'; line 04), which allows Lise to first present an account for an upcoming scenario. With the social action format *je me dis*, the speaker projects her upcoming reasoning on what she decided to do *au pire* ('at worst'), as an *ultima ratio* so to say, namely, to cook outside (line 04). Lise's quote closes her rationalization (lines 01 to 04) during which she makes accessible to her interlocutor the thinking process that led to her final decision to maybe cook outside. The direct reported thought is preceded by an indexical throwing away gesture (Bressem & Müller 2014a), which expresses, according to cognitive-linguistic studies, "getting rid of, removing and dismissing annoying topics of talk by throwing them away from the speaker's body" (*ibid*.: 3). Simultaneously, Lise gazes right (i.e., outside the window) after her pause, thus deictically referring to "outside" (line 04). At the same time, she lifts her shoulders with *pire* ('worst') and lifts them even higher to a visible shrug at line 05. The apex of the shrug matches the new information of the utterance—namely, *dehors* ('outside'). Starting at line 05, Lise also gazes at Thomas, thus being able to monitor his potentially upcoming response, which has been projected by Lise's reported thought.

Lise's shrug could be classified, according to Debras (2017), as part of "shrugs expressing affect" (*ibid*.: 24). This type of shrug may allow speakers, among other functions, to express indifference and could be glossed verbally as "I don't care about this" (*ibid*.).[221] This is essentially what Lise responds to Thomas' following laughter (fig1) and assessment *ça va être freestyle* ('it's going to be freestyle'; line 07) with *mais moi je m'en fous franchement* ('but I couldn't care less honestly'; line 08). Overlapping Lise's subsequent account for her decision (line 10), Thomas finally explicitly aligns and affiliates verbally with Lise (line 11). The sequence is closed shortly after with another shrug from Lise (line 14) during a long pause (line 14) and her reaffirmation that she does not care (line 15).

---

[221]   Streeck (2009) describes such shrugs as "[displays] of distancing and disengagement" (*ibid*.: 191).

In this excerpt, the direct reported thought is preceded and accompanied by several gestures that support the speaker's reasoning on the matter of cooking outside. The conclusive character of the pattern [*je me dis* + direct reported thought] may contribute to it being closing implicative, which is challenged by the recipient's disaffiliation. The speaker's subsequent working toward an affiliative response until its delivery clarifies that the recipient's initial response has been a dispreferred one.

The pattern [*je me dis* + direct reported thought] can also serve an interlocutor to close a sequence for the speaker. The next excerpt illustrates such a case. Ekti has just told a longer story about a (former) friend who is involved with Ekti's ex-boyfriend (not in transcript). As it transpires from the prior conversation, this ex-boyfriend cheated on Ekti which is why she is trying to convince her friend, in vain, to leave him. Just prior to this extract, Ekti accounts for her behavior toward her (former) friend by putting forward that this friend is important to her and that she only wants for her to see what is really happening (not in transcript). At line 01, Ekti starts moving towards closure.

Ex 7.8: bout du nez / tip of the nose (Pauscaf 20, 20min00)

```
EKT: Ekti, ELI: Elinda

01 EKT: @mais si elle euh elle voit PAS euh::-
           but if she uh she doesn't see uh
      ekt  @grabs her scarf,puts it around neck-->l.04

02       ff°h plus loin que son bout du +NEZ #euh$:-#+
           beyond the tip of her nose uh
      ekt                                  +gz at ELI-->+down->l.03
      eli                                            $gz left-->l.04
      fig                                      #fig1  #fig2
```
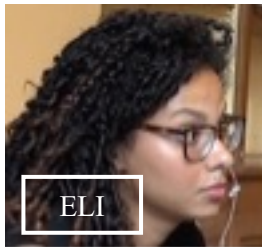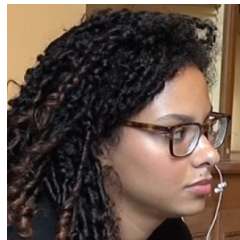

fig1   fig2

```
03       °et° pis qu'elle fait que+ tout à sa TÊte,
           and also that she only does everything her way
      ekt                      -->+gz at ELI-->l.06

04 ELI: °ben° $t'façons $ces   $ $histoires euh:- (0.3)@
          well anyway these stories uh
      eli     -->$gz at EKT-------->$left-->l.05
                       $eyebrw
                        raise $
      ekt                                           -->@

05       >je me dis< ces histoires de cUl
          I say to myself (that) these hook-up stories

         c'est toujours la même CHOse hein-$
          it's always the same (thing) huh
      eli                          -->$gz at EKT-->>

06       >tu peux pas faire< grand +CHOse=hein §°ma foi;°+§
           there is not much you can do about it huh y'know
```

223

```
    eli                                        $headshake $
    ekt                    -->+gz up--------------->+at ELI
                                                   ->1.08

07       [?HU +hh° °hh  ]
08 EKT: [°voi+LÀ quoi.°]
            PART PART
    ekt    -->+gz down-->1.17
09       (1.1)+(2.2)
    ekt     -->+
```

With initial *mais* ('but'), Ekti marks an upcoming contrast—namely, the negative assessment of the friend's conduct, who "does not see beyond the tip of her nose" (lines 01 and 02). At the end of her negative assessment Ekti gazes briefly at Elinda (line 02), which may suffice to see Elinda's facial expression displaying her affiliative negative affective stance (fig1 and fig2; Ruusuvuori & Peräkylä 2009). At the end of the turn-extension of her negative assessment (line 03), Ekti gazes at Elinda, which may foreshadow—despite the syntactic incompleteness of Ekti's turn at this point—upcoming turn-closure (Kendon 1967).

Indeed, Elinda takes the turn at line 04 and, after a self-repair, provides a conclusion to Ekti's dilemma, thus making publicly available her thought about the matter, framed by *je me dis* (lines 05 to 06). The reported thought dismisses Ekti's personal inner struggle through Elinda's generalizing stance about *histoires de cul* ('hook-up stories'). It puts the matter, which clearly affects Ekti personally, into perspective, thus potentially diminishing the gravity of the situation.

By framing her reasoning as direct reported thought, Elinda responds to Ekti's telling, while situating her response with *je me dis* as self-talk. This format allows Elinda to distance herself from her opinion about hook-up stories and may make her generalizing conclusion less offensive (as opposed to an assessment, for example). Elinda's gaze up and the expression *ma foi* (lit. 'my faith') further downgrade Ekti's responsibility in the matter by framing the circumstances as beyond her control. Displaying reduced commitment may be necessary because Elinda moves toward closure *instead of the speaker*, thus breaching epistemic rights which have been established in the preceding telling as being primarily Ekti's who experienced the situation. Nevertheless, with her reported thought, Elinda responds affiliatively to Ekti's telling—an affiliation that is confirmed by a sequence-closing third by Ekti (line 08).[222]

This excerpt has shown how a recipient may close a reasoning sequence with direct reported thought introduced by *je me dis* instead of the speaker, thus evidencing the intersubjective dimension of direct reported thought. Elinda suggests, with her direct reported thought, a solution to a dilemma that Ekti has described before. In that sense, it is an attempt initiated by the recipient to close a (long) sequence in which the speaker does not seem to come to a solution herself.

The last excerpt illustrates a case where the emergent reasoning that the speaker displays with the social action format *je me dis* is co-constructed. In contrast to the previous excerpt, both speakers present—together—different possibilities in a scenario that affects neither of

---

[222]   Note, however, that this affiliation comes in at a moment in the sequence when Ekti has (prosodically and grammatically) projected a continuation of her turn (line 11), which may be a delicate matter.

them personally but seems to preoccupy them because they know the involved parties. Josiane and Marie talk about a mutual acquaintance who is not able to break away from her ex-boyfriend. According to the interlocutors, the reason for this is her ex-boyfriend manipulating her. With the direct reported thought, Marie introduces her point of view in the conjointly developed reasoning.

Ex. 7.9: la laisser partir / let her go (Pauscaf_02_05, 22min35)

```
JOS: Josiane, MAR: Marie

01 JOS: pis c'qui ferait du bien à +long TERme j'dis- (.)
        and also what would do good in the long run I mean
    mar                             +gz at JOS-->1.07

02      il s'met $en COUple-
        he gets into a relationship
    jos         $gz at MAR-->1.08

03      elle est !O!bligée- (.)
        she has no choice

04      lui il veut !PLUS! rentrer en contact aVEC-
        he  he doesn't want to get in touch with (her) anymore

05      pis elle est obligée d'passer à aut' CHOse;
        and she has no choice but to move on

06      (0.4)

07 MAR: le problème c'est qu'ce TYpe-+
        the problem is that this guy
    mar                         -->+gz up-->1.10

08      (0.2)$(1.1)
    jos    -->$gz down-->1.09

09 JOS: il $s'rait encore capAble d'entretenir un [TR+U  ]C hein,
        he would still be able to maintain a thing huh
    jos  -->$gz at MAR-->1. 12

10 MAR:                                           [OU+AIS-]
                                                   yeah
    mar                                    -->+gz at JOS
                                              -->1.14

11      %c'est ÇA;%
         exactly
    mar %nod----->%

12      (0.5)$
    jos    -->$gz left-->1. 15

13 JOS: [ah +$ouais-$=ah ouAIS,]
         oh  yeah     oh yeah

14 MAR: [°h +$pis je$ me DIS- ] %(--)
            and I say to myself
    mar  -->+gz up-->1.15
                               %....
    jos       $eyebrw
              raise$

15      il a $de !TOU%TES!#mani+ères pas envie d'la# laisser parTIR;
        he doesn't want to let her go anyway
    mar  .............%pointing gesture-->1.16
```

```
                            -->+gz at JOS-->>
     jos    -->$gz at MAR-->1.16
     fig            #fig1                    #fig2
```


fig1   fig2

```
16     %(--) %$pa'ce que sinon%[il  fer%ait pas TOUT c'qu'il] fait
              because otherwise he wouldn't do all of what he does
   mar  %.....%flat hand------>%,,,,,,->%
   jos       -->$gz down-->1.17

17 JOS:                     [en fait (.)  il veut vraiment]
                             in fact     he really wants

18     [il veut vrai]ment la gar$der sur le cô[TÉ,    ]
        he really wants to keep her on the side
   jos                      -->$gz at MAR-->1.23

19 MAR: [pour euh     ]                    [↑ouais-]#
          to uh                              yeah
   fig                                        #fig3
```


fig3

```
20 JOS: (.) [comme si] tu l'amènes sur [le BANC-]
            as if you would take her to the bench

21 MAR:     [exact-  ]                   [comme ça]
             exactly                       like that

22 JOS: [il fait jouer les autres il fait: ouais- $(.) ↑hop-$]
        he makes the others play he makes yeah        hop

23 MAR: [exact ()                      ah ben $voilà;    $]
         exactly                        ah well PART
   jos                                  -->$gz down->$at MAR
                                                      -->>

24     [ouais,]
        yeah

25 JOS: [c'est ] (.) c'est incroyABLE;
         it's        it's unbelievable
```

At lines 01 to 05, Josiane suggests a solution to the problem of her mutual acquaintance being unable to break away from her ex-boyfriend. Subsequently, Marie projects an upcoming disagreement by means of a projector phrase of the type "N be that" (line 07; Günthner 2011b; for French, see also Pekarek Doehler 2011), thus allowing her to foreshadow an upcoming

multi-unit turn. Subsequently, Marie first embodiedly stages her thinking with a gaze up (line 07; Heller 2021), which she holds until line 10. In the subsequent pause (line 08), Josiane gazes down, which may indicate her being about to take the turn (Kendon 1967). Indeed, at line 09, she suggests a candidate completion of Marie's syntactically incomplete turn. This negative assessment of "the guy," meaning the ex-boyfriend, is emphatically confirmed by Marie with stressed *ouais* ('yeah'), a gaze at Josiane, and a second verbal confirmation and a nod (line 11).

In overlap with Josiane's affiliative repeated response (line 13), Marie initiates an additional argument to the conjoint reasoning (lines 14 to 16) with the pattern [*je me dis* + direct reported thought]. *Je me dis* is followed by a pause, which allows Marie to briefly interrupt her talk while initiating her pointing gesture (fig1 and fig2). The pause may also be a means to attract the interlocutor's gaze, which is established shortly after (line 15). Marie's eyebrow raise may signal the topical importance of what is about to follow (line 14; Krahmer and Swerts (2004): While establishing mutual gaze with Josiane (line 15), Marie makes publicly available her reasoning about the ex-boyfriend (lines 15 and 16). The reported thought somewhat downgrades its validity from an assertion to something more speculative. At the same time, the quoting-format allows Marie—in contrast to epistemic parentheticals like *je crois* ('I believe')—to stage her rationalization as such: With her pointing gesture,[223] her gaze up (fig1 and fig2), and with the pause between quotative and quote (line 14), she stages her rational thinking process as emerging in real time while she talks.

Marie's display of her emergent reasoning is interrupted by Josiane who selfs-selects in overlap (line 17): She suggests another assumption about the ex-boyfriend's intentions. *En fait* ('in fact') thereby frames the upcoming utterance as somewhat related to what precedes (line 17). Again, this suggestion is emphatically confirmed by Marie (line 19), who also affiliates with Josiane's negative assessment through her facial expression (fig3). The sequence is closed, after Josiane's further explanation, with an assessment (Goodwin 1986).

This excerpt has illustrated that the reasoning displayed through the pattern [*je me dis* + direct reported thought] can also be co-constructed by two interlocutors. In this case, the direct reported thought seems to allow the speaker to initiate one supposition among others within the reasoning activity. Interestingly, in this excerpt, the—usually highly dispreferred— interruption remains unproblematic as the interlocutors agreement concerning the Josiane's conclusion demonstrates (lines 23 and 25).

7.3.1.4.   Summary

This section has illustrated that the social action format *je me dis* projecting direct reported thought allows speakers to put their thinking process on display by making publicly available their thoughts *while they emerge in real time*. Occurring recurrently in reasoning sequences, the format is often used in its closing environment (Excerpts 7.6–7.8). However, it may also be used in a conjoint (and agreeing) reasoning process to initiate an argument (Excerpt 7.9).

The analysis of the speaker's multimodal conduct while quoting themselves has evidenced their use of ceiving gestures (Streeck 2009) and the palm-up open-hand gesture (Marrese *et al.*

---

[223]   Pointing gestures have been classified by Streeck (2009) as "gestures of ceiving," which Streek says may allow speakers to structure their talk (see Section 6.4.1.4 for detailed explanations).

2021; Müller 2004) to structure their talk, as well as shrugs (Jehoul *et al.* 2017; Debras 2017) and eyebrow raises (Krahmer & Swerts 2004) to mark their arguments as obvious or simply to structure their talk.[224] While these body movements play an important part in the individual excerpts, a quantitative overview of all occurring body movements in my corpus revealed that none of the body movements is overwhelmingly frequent other than the systematic gaze at the recipient (at the latest) at the end of the reported thought (*n* = 25) as the following table shows:[225]

Multimodal conduct	Frequency
speaker's gaze at (one of the) interlocutor(s) at the end of the quote	89.3% (25/28)
eyebrow lift	3.6% (1/28)
facial expression	3.6% (1/28)
head movement to the side	7.1% (2/28)
extended hand gesture (palm down/side)	17.9% (5/28)
palm-up open-hand gesture (PUOH)	10.7% (3/28)
shrug	10.7% (3/28)
headshake	21.4 % (6/28)
iconic gesture	3.6% (1/28)
other (eating, drinking, self-grooming)	21.4 % (6/28)
no specific body movement	7.1% (2/28)

Table 40: Multimodal conduct co-occurring with [*je me dis* + direct reported thought]

The high number of speakers' gaze at the recipient may indicate that speakers either monitor the recipient's attention—when the direct reported thought is positioned in the middle of the sequence (*n* = 8)—or seek to mobilize the recipient's response, which has been made relevant through direct reported thought in the closing environment of the sequence. Another interesting outcome of the quantification is that, compared to the quotatives *je suis là / j'étais là*, facial expressions occur only once after *je me dis*. Instead, headshakes, eyebrow raises, and extended hand gestures are more recurrent, as my qualitative analysis also demonstrated.

   A certain "multimodal assembly" (Pekarek Doehler *et al.* 2021), which does not show in the table above, occurs six times in my collection: a combination of resources seems to allow speakers to stage their thinking process. This includes gazing up or into space, frowning, squinting, and/or pursing their lips. This multimodal assembly seems to allow speakers to stage the action that they accomplish with *je me dis*—namely, making reasoning publicly available, presented as being assembled "on the fly" (Hopper 1987) while talking (e.g., Excerpt 7.9).

---

[224]   At the same time, eyebrow raises and shrugs also occur during assessments and affective stance-taking (Ekman 1979; Streeck 2009). As Krahmer and Swerts (2004) rightly point out, marking a topical focus and displaying an affective stance through shrugs or eyebrow raises are two actions that are separated by a fine line. It is difficult to say, from my limited data, which function is prevalent, especially because reasonings often hover between argumentation and affectiveness (specifically, Excerpts 7.10, 7.12, and 7.14).

[225]   As I have already pointed out on several occasions (see Chapter 6), these numbers deliver only rough tendencies. The different participant frameworks make especially the interpretation of speaker's gaze problematic. However, because I focus on the production side of multimodal conduct, this quantifying analysis may give first insights into the set of resources speakers use while quoting their thought, which may be further investigated by future research.

The low number of facial expressions and gestures more generally fits the sequential environment in which *je me dis* occurs. Speakers rather embed their direct reported thought in a reasoning that is presented as non-emotional or rational. Out of 28, only three sequences can be classified as tellings according to the criteria that Labov and Waletzky (1967) established. In the remaining 25 sequences, speakers are involved in narrations that are first and foremost reasonings, where speakers either make publicly available their rationalization on a certain topic ($n = 20$) or lay out arguments to convince their interlocutor of a certain point of view ($n = 5$).[226]

My analysis has also demonstrated that, when used at the end of a turn ($n = 15$), [*je me dis* + direct reported thought] allows a speaker to move toward closure. By doing so, speakers seem to mobilize the recipient's response, often together with a gaze at the recipient. This qualitative result confirms the quantitative outcome that speakers gaze at the recipient at the end of the quote in 89.2% of all occurrences. [*Je me dis* + direct reported thought] is also used in the middle of a turn ($n = 13$). In this case, speakers seem to use the pattern to introduce one argument among others, sometimes in response to a recipient's first pair part (Excerpt 7.8).

Bodily conduct accompanying [*je me dis* + direct reported thought] does not contain facial expressions. In contrast, shrugs, headshakes, and eyebrow raises are used recurrently (e.g., Excerpt 7.6, 7.7, and 7.8; see Table 40 for an overview). These kinds of embodied actions have been related in the literature to several phenomena, which I have also demonstrated in my analysis:

- Eyebrow raises may display skepticism (Ricci Bitti *et al.* 2014; Krahmer & Swerts 2004; Hutchby 1992)[227] (see Excerpts 7.6 and 7.9).
- Headshakes may express "a negation that is *implied*" (Kendon 2002: 150), even in positive statements (Excerpt 7.8).
- Shrugs "have to do with the expression of subjective negation ('I…not'), and the expression of the speaker's withdrawing from something because of inability, unwillingness or ignorance" (Debras 2017: 24; Excerpts 7.6 and 7.7).
- So-called gestures of ceiving or ception like PUOH (Streeck 2009) have been argued to help speakers structuring their talk (see Excerpt 7.8).

As Table 40 shows, the above-mentioned body movements occur in about 30% of all instances.[228] Nevertheless, this result indicates that, as with the past form, there seems to be a tendency for speakers to use body movements that only display a moderate affective stance, as none of the gestures is particularly expressive or exaggerated. One possible explanation for this result may be that reasonings, because they are less emotionally charged, and because they do not occur in tellings where dramatizations are recurrent, necessitate less staging and thus less embodied conduct.

---

[226] Excerpts 7.6–7.8 demonstrate this first function, and Excerpt 7.9 illustrates how a speaker may use the pattern [*je me dis* + direct reported thought] to share a downgraded personal assumption.

[227] Ricci Bitti *et al.* (2014) point out that eyebrow raises can be used for accentuation or to express emotionality. Both functions may be difficult to differentiate, and further systematic studies of natural conversation would be needed to distinguish them clearly.

[228] I am aware that the low number of cases does not allow for any generalizing conclusions. What I would like to point out are merely tendencies that should be verified with a larger data set.

The next section treats *je suis là* ('I am there'), which shows several differences to *je me dis*, especially concerning the speaker's degree emotional involvement.

7.3.2.   Reporting Thought with the French *Je suis là* ('I am there')

The present section focuses on the pattern [*je suis là* + quote].[229] Compared to my analyses in Section 6.3,1, where direct reported thought is introduced by the more frequent past-tense quotative *j'étais là* ('I was there'), this section presents three of the five instances of the quotative *être là* in narrative present tense. The quotative is morphologically structured as follows:

je	*suis*	*là*
PP.1SG	be-PRS.1SG	DEIK

'I am there'

All five tellings containing [*je suis là* + quote] are complaints (Drew 1998; Drew & Walker 2009), which have been shown to recurrently contain reporting formats (Skogmyr Marian 2023). As Haakana (2007) has shown for Finnish, reported thought seems to be recurrently used in complaint stories. It allows speakers to display their affective stance in an immediate, staged way, while distancing themselves from their reported thought (i.e., their stance-taking) by the means of merely "reporting" it.

According to Haakana (2007), using a quotative that leaves open whether something has been really said or only thought may be a useful interactional tool in such complaint stories where delicate matters are often addressed. As with *j'étais là*, *je suis là* allows speakers to leave open whether a quote has been only thought or actually said. Speakers may use this ambiguity interactionally because they leave it up to the recipient to make their best guess about the most probable case.

7.3.2.1.   A General Picture: A Quantifying Analysis

Je suis là ('I was there') bears, like the other quotatives that I have treated so far, some phonological and morpho-phonological features that may be of interest for its general analysis. The following table shows these features:

Feature	Frequency
accelerated	40.0% (2/5)
no accentuated syllable	80.0% (4/5)
pause after *je suis là*	-
morpho-phonological reduction to *chuis là*	80.0% (4/5)

Table 41: Prosodic and morpho-phonological features of *je suis là*

Because the overall number of cases is so low, it is difficult to draw any conclusions from the results. The only noticeable characteristic is that four of five cases are morpho-phonologically reduced to *chuis là*. This is, however, a recurrent feature of *je suis* ('I am') in spoken French

---

[229]   As with the past form, I will use for this pattern the denotation [*je suis là* + quote] instead of [*je suis là* + direct reported thought] to account for the ambiguity of the quote being direct reported thought or speech.

(for a similar morpho-phonological reduction of *je sais pas*, see Pekarek Doehler 2016, 2019; see also Pekarek Doehler *et al*. 2021).

The five occurrences are preceded by different items, as the following table shows:

Form	Frequency
après ('then')	20.0% (1/5)
du coup ('so')	20.0% (1/5)
et moi ('and I')	20.0% (1/5)
((pause))	40.0% (2/5)

Table 42: Lexical items co-occurring with *je suis là*

Concerning the items following *je suis là*, there are two cases where *je suis là* is followed by a response cry and two other cases where speakers first use body movements before verbalizing the quote. In the remaining case, the speaker simultaneously produces a response cry and a facial expression.

7.3.2.2.   Syntactical Features

Je suis là bears the same syntactic features as its past tense equivalent (for a detailed explanation, see Section 6.3.2): It can be considered as syntactically complete with its combination of a copula (*être*, 'to be') and the deictic particle *là* ('there'). The deictic can refer to a place or time of an imaginary space. This feature does not allow for any syntactic embeddedness of the quote even if it was a complete clause.[230] Because of the non-possibility of a syntactic relation between quotative and quote, I suggest, similar to the past form, an analysis of *je suis là* as projector-construction projecting the second part of a bipartite structure.

In the following sequential analyses, I will illustrate that speakers use the social action format [*je suis là* + quote] in complaint sequences where the—bodily staged—direct reported thought allows them to display their negative affective stance.

7.3.2.3.   Interactional Functions of *Je suis là* ('I am there')

In this section, I will present three of the five excerpts of [*je suis là* + quote] occurring in my corpus of spoken French. The social action format is always used in complaints that are formatted as storytellings but the direct reported thought varies regarding its formal characteristics. The following three excerpts contain three types of quotes: a complete clause, a response cry followed by a clause, and bodily conduct only, thus demonstrating the variety of bipartite structures in my corpus. All excerpts do have in common, however, that—regardless of its format—the quote is always bodily staged.

The first excerpt is the only case where *je suis là* is immediately followed by a complete clause. Mia and Alice are talking about their masters, which they assess rather negatively.

Ex. 7.10: professionnellement / professionally (Pauscaf_02_01, 04min39)

```
MIA: Mia, ALI: Alice
```

---

230   For a detailed discussion of the syntactic status of the quote introduced by *je suis là*, I simply refer to Section 6.3.2.2, where different possibilities are discussed in detail.

```
01 MIA: moi ce qui me saOULE- (0.4)
        me   what annoys me

02      tout ce qu'on FAIT (.) $pour euh: ce masTER-
        everything that we do   for our masters
   ali                           $gz at MIA-->l.03

03      j'ai l'impression que ça sera !PAS! valorisable
        I have the impression that it will not be valuable

        PROfessionnelle$+ment;
        professionally
   mia                    +gz at ALI-->l.05
   ali                --->$gz down-->l.05

04 ALI: ah [NON    +(sûrement pas)]
        oh no       surely not

05 MIA:    [et du +↑coup:        ] $%chuis# LÀ:-  %
            and so                    I'm    there
   mia                                %eyebrow raise%
          -->+gz right-->l.07
   ali                          -->$gz at MIA-->l.08
   fig                                     #fig1
```

fig1

```
06      ben ça fait #CHIER-=
        PART that's a pain
   fig              #fig2
```

fig2

```
07      =parce que tu +passes [du   +TEMPS, et $euh +:  ]
         because you invest         time   and uh
   mia             -->+gz left-->l.08

08 ALI:                       [bah r'+GARde- la $PREU+ve,]
                               well look    the proof
   ali                                     -->$
   mia                         -->+gz at ALI--->+down->l.09

09      j'ai fait mon cé +VÉ là;
        I've made my  CV    here
   mia             -->+gz at ALI-->>

10      mon (        )-
        my (        )

11      qu'est-ce que j'ai vanTÉ,
        what did I brag about
```

```
12      mes projets de groupe en éco de va/ de chaleur
        my group projects    in eco(logy) of va of heat

        et tout ÇA,=
        and all that
13      =des $trucs que j'ai faits$ en bacheLOR;
         things that I have done in my bachelor's degree
   ali      $shrug------------->$
14 MIA: ouais voiLÀ-#
        yeah exactly
   fig              #fig3
```

fig3

Mia is involved in a complaining activity about her master courses which is projected at line 01 by the pseudo-cleft construction *moi ce qui me saoule* ('me what annoys me'). Pseudo-clefts are a recurrent projecting format (Blanche-Benveniste 2010; Günthner 2008b; Pekarek Doehler 2011) that allows speakers to move a specific topic in the foreground of the ongoing activity— in this case, the complainable (line 03). Alice affiliates with Mia verbally (line 04) while the latter already continues her turn with her reported thought.

Shortly before *chuis là*, the speaker gazes right, thus averting mutual gaze during her strong negative assessment. One reason for Mia's gaze away may be the potential disagreement on the topic: Her friend attends the same class but may not have the same strong negative opinion on the master's degree. *Chuis là* is produced with continuative final intonation. It is the only of all five occurrences where the deictic particle is prosodically stressed and lengthened. Mia also raises her eyebrows just during the quotative (line 05, fig1), thus topically focusing on what is about to follow (Krahmer & Swerts 2004).

Projected by the social action format *chuis là*, Mia verbalizes her negative affective stance (line 06). Preceded by a particle, *ben*, that expresses obviousness (Mosegaard Hansen 1998: 250ff), thus marking shared epistemic access to the upcoming matter, Mia utters a strong formulaic expression: *ça fait chier* ('that's a pain'). She displays her affective stance also through bodily means: Her gaze up and her PUOH gesture (fig2) allow her not to mark her critique of the master courses and embodies the obviousness of her point (Marrese *et al.* 2021) toward her interlocutor, who shares her knowledge about the master.

In overlap with Mia's account (line 07), Alice takes the turn herself (line 08) and delivers an affiliative second story (lines 09–13). The sequence is closed with Mia's verbal confirmation and her facial expression (squinting and raising her eyebrows) allowing her to display her evidential vindication (Kendrick 2019) through embodied means (fig3).

This excerpt has shown that *je suis là* can introduce a complete clause with which speakers verbalize their affective stance. When speakers have, like in this excerpt, equal epistemic access to the complainable, the direct reported thought, and thus the speaker's strong affective

stance, may be co-constructed; recipients can join the stance-taking activity by delivering themselves an account for why the speaker's stance is justified.

Excerpt 7.11 shows an instance where the speaker's affective stance is at first displayed through a non-lexical vocalization and the speaker's bodily conduct, followed by a verbal explanation. Oréane asked Liv prior to this excerpt whether she and her mother sometimes go to the movies. Liv responds that her mother is very stressed and tired from work at the moment and prefers staying at home in the evening. This is followed by Oréane's question at line 01.

Ex. 7.11: dix minutes / ten minutes (Pauscaf 19, 17min22)

```
ORE: Oréane, LIV: Liv

01 ORE: et tu r'gardes genre vous r'gardez la télé enSEMble?=
        and you watch  like  you   watch   TV      together

02 LIV: =OUI 'fin elle a+dore# danse avec les STARS;+
         yes like she  loves 'dance with stars' (TV-show)
    liv              +gz up, rolls eyes--------->+gz up left
                                                  -->l.03
    fig                        #fig1
```


fig1

```
03      (0.4)+(1.2)
    liv    -->+gz at ORE-->l.05

04 ORE: <<high pitch>c'est trop [chou,>]
                      that's so cute

05 LIV:                      [enHEhe]hehe- +°H
    liv                             -->+gz left-->l.07

06      (1.0)

07 LIV: oui bon ça m'+va deux minutes enfin >ça m'va dix minutes<=
        yeah well it's fine for me two minutes or say it's fine for
        me ten minutes
    liv           -->+gz at ORE-->l.08

08      après +chuis là# %((sound, 0.4s))+ j'en peux PLUS;%#
        then I'm there              I can't take it anymore
    liv     -->+gz up, rolls eyes------->+gz left-->>
                        %headshake--------------------->%
    fig              #fig2                    #fig3,fig4
```


fig2   fig3   fig4

```
09      (   )
```

```
10      (0.8)
```
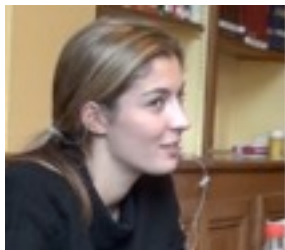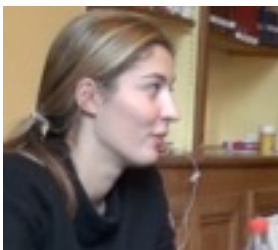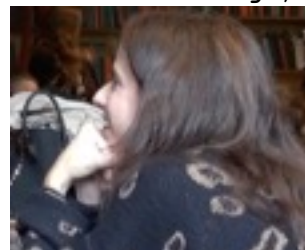
At line 01, Oréane produces a first pair part with a question. With *genre* ('like'), the question is marked as being one exemplar activity out of several. Liv responds affirmatively and subsequently adds which TV-show her mother loves to watch (line 02). The utterance accounts for the implied consequence that Liv would watch this TV-show with her mother. At the same time, Liv demonstrates that she does so only reluctantly. Her negative affective stance is bodily displayed: She gazes up and rolls her eyes (fig1; on eye rolls as a means of stance-taking from a linguistic anthropological point of view, see Goodwin & Alim 2010).

During the subsequent pause, where a recipient response remains noticeably absent, Liv starts gazing at her interlocutor. The mobilization of her interlocutor's response seems to be successful as Oréane responds with a positive assessment of the mother's TV-show preferences (line 04), which can be, in this specific context, also understood as slightly mocking—especially due to its production with high pitch and the lengthening of the evaluative adjective *chou* ('cute'; line 04). With *oui bon* ('yeah well'), Liv brushes off Oréane's potentially ironic assessment. Subsequently, she re-positions herself with her direct reported thought: After a self-repair (line 07), Liv uses the social action format *chuis là*, which is produced in a morpho-phonologically reduced way. The projected quote consists of, first, a non-lexical vocalization during which she gazes up again, rolls her eyes (fig2 and fig3), and produces a headshake, and second, of the formulaic expression *j'en peux plus* ('I can't take it anymore'; line 08).

The speaker's affective stance is, in this excerpt, first expressed through bodily means and a non-lexical vocalization. The verbal stance-taking is only produced subsequently. By doing so, the speaker can foreshadow her affective stance early on, thus guiding the recipient already toward the expected response. Compared to the first excerpt, the recipient only responds with laughter (fig4), not verbally. The subsequent pause (line 10), after a non-audible "after thought"–like turn-extension from Liv in very low volume (line 09), may indicate that the sequence is closed at this point. Indeed, after the pause, Oréane initiates another topic about what she has recently watched on TV (not in transcript).

The last excerpt is an instance where *je suis là* is followed primarily by bodily conduct. The two accompanying verbal items project "more to come," but the clause remains incomplete. Carole, Thibaud and Alba[231] talk about different kinds of greetings depending on their different cultural backgrounds. As it transpires from their prior conversation, all three are Swiss but with different origins. This information is important for the following excerpt, where the three make fun of a certain, apparently Arabic, way of greeting that Thibaud uses which consists of a high-pitched sound and a rapid move of the tongue between the lips and the palate (transcribed with 'brl' in the transcript). Prior to this excerpt, Alba characterized this greeting as 'unbearable'. In what follows, she describes a situation where one of her friends, who uses this way of greeting, justifies it toward Alba.

Ex. 7.12: signe / sign (Pauscaf_02_11, 15min17)

```
ALB: Alba, THI: Thibaud, CAR: Carole
```

---

[231] The entire videorecording is an interaction that takes place between four participants. One of them, however, is involved in a reading activity of the consent form and does not participate in the present interaction.

```
01 ALB: [il me sort d'sa] voiTURE il fait-
         he says out of the car (to me) he does

02 THI: [brlbrlbrl      ]
         ((greeting sound))

03 ALB: #nous au moins on a un SIGne-
         us at least we have a sign
    fig  #fig1
```

fig1

```
04      vous vous avez QUOI,
        you what do you have

05      je suis là mais# [euh-  ]
        I'm there but    uh
    fig                   #fig2
```

fig2

```
06 THI:                  [bah on] a [un pays STA#ble
                          well we have a stable country

07 CAR:                              [on a des   #(   ) hihi
                                      we have (   )
    fig                                           #fig3
```

fig3

```
08 ALB: uhinhinhinhinhinhin <<to THI>#ah merci->
                              ah thank you
    fig                      #fig4
```

fig4

While Carole and Thibaud are still involved in the reenactment of Thibaud's greeting (line 02), Alba reports the third party's speech (line 01): This third party (*il* 'he'; line 01) purportedly challenged the legitimacy of Alba's way of greeting (whether in Alba's home country or in Switzerland is not yet clear) by saying that, in his culture, they have at least *un signe* ('a sign'; line 03), which Alba subsequently reenacts (fig1). The third party's reported speech continues with a question directed at Alba: *vous avez quoi* ('what do you have'; line 04).

Within the reported dialogue, Alba's following direct reported thought is responsive to this question. The quotative *je suis là* is followed by a disruptive *mais* ('but'; line 05), which may be a placeholder displaying Alba's opposition to the third party's prior (reported) turn. The following *euh* ('uh'; line 05) stages Alba's speechlessness, while she takes a stance through bodily means: Covering her mouth and nose with both hands (fig2), she displays both embarrassment about apparently not having any sign (she also gazes right as if to reenact averted gaze) and bewilderment about this third party attributing so much value to his culture-specific way of greeting (fig3). The social action format *je suis là* projects a quote that only consists of Alba's depiction of her affective stance with bodily means. She can thus display her speechlessness—a possibility that speakers do not seem to have with the quotative *je me dis*. Not quoting anything verbal may be a way to keep her negative affective stance "low profile," thus leaving more room for interpretation to her interlocutors.

Switching his footing and talking on behalf of Alba, Thibaud responds within the telling with the reenacted utterance *bah on a un pays stable* ('well we have a stable country'; line 06), thereby referring to Switzerland. Alba affiliates with his response and orients to it as amusing through laughter (line 07) while thanking Thibaud for his affiliation verbally and bodily (fig4). The direct reported thought with *je suis là* allows Alba to stage her immediate response to a third party's prior turn as emerging spontaneously in the described situation. Her negative stance toward her interlocutor is thus presented as part of the reenactment. Displaying her affective stance primarily through her bodily conduct allows the speaker to increase the moment of dramatization by depicting her speechlessness through an actual absence of speech. This demonstrates that *je suis là* can introduce a large spectrum of resources, including demonstrations of "not talking." The format is also a possibility for Alba to leave open whether what she reports has been 'said' or thought in the reported situation. Both characteristics might be a conversational means enabling speakers to still back down in case of the (current) interlocutors' disaffiliation, which—in this excerpt—does not occur.

7.3.2.4. Summary

My analysis has shown that *je suis là* is, at least in my data, rare. When it occurs, however, it is consistently used as a social action format introducing direct reported thought or speech. It allows speakers to display their negative affective stance toward a prior matter verbally (Excerpts 7.10 and 7.11) and/or through embodied means (Excerpt 7.12) in complaint sequences. One of the most important convergences between the three excerpts is the speakers' use of embodied resources during the direct reported thought—a feature that the remaining two instances also have, as the following table illustrates:

Multimodal conduct	Frequency
speaker's gaze at (one of the) interlocutor(s) at the end of the quote	100.0% (5/5)
speaker's gaze away during the quote	80.0% (4/5)
facial expression	80.0% (4/5)
palm-up open-hand gesture (PUOH)	40.0% (2/5)

Table 43: Multimodal conduct accompanying [*je suis là* + quote]

Similar to the social action format in past tense, *je suis là* is recurrently followed by facial expressions, which allow speakers to display their affective stance (Ruusuvuori & Peräkylä 2009; Peräkylä & Ruusuvuori 2006, 2012). Of course, the small number of cases does not allow me to draw any general conclusions. However, it seems as if speakers use *je suis là* to prioritize the depiction of reported material and to "*show* others what it [the reported speech/thought, SF] looks or sounds or feels like" (Clark 2016: 342, original emphasis).

Both parts of the bipartite structure, quotative and quote, are syntactically independent and only linked through the presentational character of *je suis là*. Because the deictic *là* projects its specification, it is, like the past tense *j'étais là*, grammatically best described as projector-construction (Pekarek Doehler 2011) or *annonceur* (Authier-Revuz 2019). The quotative cannot grammatically project more to come, but it can do so pragmatically and interactionally. What is projected, however, tends to not be verbal material. If verbal material is introduced, it is either preceded or accompanied by a body movement (Excerpt 7.11) or it contains formulaic expressions (Excerpt 7.10).

One multimodal assembly occurs in four of the five excerpts: the speaker's gaze away during the quote coupled with a facial expression (see also Table 43). This assembly may be due to the actions that are projected by *je suis là* and to the sequential position in which it occurs. As my analysis has shown, speakers resort to *je suis là* when they stage their negative stance toward a previously described event which may be dispreferred or, generally, a delicate matter. Averting gaze while producing the reporting format may indicate that the speaker orients to the dispreferred character of their negative affective stance. Because [*je suis là* + quote] is responsive to previous events, it is not impossible that speakers reenact what is a recurrent bodily conduct in dispreferred responses to polar questions (Kendrick & Holler 2017).

The strong affective stance that speakers display through the pattern [*je suis là* + quote] projects a recipient response. Indeed, in four of five cases, recipients respond, sometimes in overlap. In the remaining case, the speaker closes her sequence with [*je suis là* + quote] and the recipient response remains noticeably absent. Recipients do, however, orient to the speaker's

telling, at least multimodally: In all five excerpts, recipients gaze at the speaker, at the latest at the end of the quote. This gaze conduct may either display the recipient's attention, or it foreshadows upcoming turn-taking (Kendon 1967).

Because speakers leave open whether their quote has been said or thought, the social action format *je suis là* allows them to blur the line between actual talk—which may have been sanctioned by the original interlocutor given the strong affective stance display—and unspoken stance-taking. This fuzziness may be exploited by speakers by leaving it up to the recipient to make a "best guess." Based on contextual information from the sequence, recipients seem to be able to do so as the matter is, in my data, not addressed specifically. The ambiguity of the speakers' reported speech/thought might also allow them to always keep "an easy way out" should it turn out that recipients disaffiliate.

After the interactional analysis of the two French social action formats *je me dis* and *je suis là*, I will compare them regarding form and function in the following section.

7.3.3. Functional and Formal Differences between *Je me dis* ('I say to myself') and *Je suis là* ('I am there')

The two preceding sections have demonstrated that speakers use *je me dis* ('I say to myself') and *je suis là* ('I am there') to project a direct reported thought or, in the case of *je suis là*, ambiguous reporting formats. I have shown that the distribution of the two social action formats seems to be motivated by action formational characteristics: While speakers use *je me dis* recurrently in reasoning sequences to make a post-hoc rationalization available to their interlocutors, *je suis là* is used in emotionally charged tellings where it allows speakers to take a strong affective stance. Interestingly, the *je suis là*-quotative is much rarer than the *je me dis*-quotative, which is not the case for their past tense equivalents (see Section 6.3.2).

It is striking that, compared to the quotative in narrative present tense, the past-tense quotative *j'étais là* is, in the same amount of data, over eleven times more frequent. It may be possible that narrative present tense is generally less frequent in French tellings. One argument in favor of this hypothesis would be that in my German data, speakers shift to narrative present tense in various moments of their telling (see Section 7.2), whereas French speakers use *je suis là* only in tellings that are exclusively in present tense, which might be generally rare.

One result of my analysis, however, contradicts this hypothesis. When looking at the tellings where [*j'étais là* + quote] occurs, I find that there are 13 cases[232] where speakers shift from narrative present tense *before* the direct reported thought to past tense only for *j'étais là* and shift back to present tense during the quote itself (if containing language). What this finding reveals, is that *j'étais là* seems to have routinized into a format that is sometimes used tense- and aspect-independently. Previous literature has so far shown that speakers may shift to narrative present tense during a quotation or story climax (Casparis 1975; Günthner 2006; Wolfson 1978)—the reverse shift to past tense in a narrative present tense telling therefore

---

[232]   In 10 of these 13 cases, reported speech precedes *j'étais là*, which is also framed as such by quotatives in narrative present tense, such as *elle dit* ('she says') or *elle me fait* ('she does to me'). In three other cases, the preceding description is in narrative present tense.

seems rather odd, or even unlikely. I argue that, in fact, what seems to be a temporal shift on a strictly formal level may be actually explained through tendencies of routinization.

Routinization is the process whereby grammatic patterns in use, emerging from interaction, are recurrently used to accomplish specific social actions. This routinization does not include characteristics that are claimed for grammaticalized or pragmaticalized linguistic structures, like positional flexibility, semantic bleaching, and the "reanalysis of a pattern in the domain of language use as a pattern in the domain of language structure" (Thompson & Mulac 1991a: 318; on grammaticalization in general, see Hopper & Traugott 2003; Traugott & Heine 1991b, 1991a). Criteria that have been established for grammaticalization (see, for instance, Lehmann 2015) do not necessarily apply to the linguistic and bodily formats that I describe herein.

Routinized forms are "*[r]ecurrent* turn-designs—including *recurrent* on-line adaptations—for accomplishing precise actions" with "*frequent* combinations of grammatical (and bodily) units, ultimately ensuing in the sedimentation of grammatical action formats from *frequent combinations in use*" (Pekarek Doehler 2021b: 837, my emphasis). What transpires from Pekarek Doehler's description of routinization, which I adopt, is the importance of a recurrent and frequent use of a (grammatical or bodily) format to become routinized. How crucial the process of repetition is to the development of fixed forms has been put forward by Haiman (1994), thereby using the terminology of *ritual language* and *ritualization*. Pointing toward the parallels between anthropological and ethnological observations on rituals, Haiman argues that automatization and the growing emptiness of meaning in rituals are processes comparable to grammaticalization in linguistic descriptions (*cf. ibid.*: 8).

It is well known that, when recurrent linguistic structures routinize to chunks because they are frequently used for similar actions in talk, they may lose their original grammatical status and partially also their original semantics (Imo 2012b; Kärkkäinen 2012, 2010; Pekarek Doehler 2016; Thompson & Mulac 1991a). My results indicate that the grammatical information of French imperfective tense remains morphologically encoded but is not semantically or pragmatically exploited. That means that neither tense nor aspect in relation to preceding or following parts of the telling seem to play a role anymore—at least in those instances where speakers shift from narrative present tense to *j'étais là* and back to narrative present tense during the quote. This finding is a strong indicator for the ongoing routinization of *j'étais là.* The rarity of *je suis là* may simply be a side effect of this routinization.

Syntactically, both quotatives considerably vary. Neither of them is syntactically linked to the quote that follows. However, *je me dis* at least grammatically projects an empty argument position, which can then be filled with different material as my analysis has demonstrated. *Je suis là*, in contrast, does not project anything on a grammatical level. If at all, the pattern interactionally projects the specification of its deictic *là* ('there / at that moment'), as it does in its past form. This deictic allows the speaker to depict subsequently their state or actions in—or rather: in response to—the situation that has just been described. By doing so, the speaker invites the recipient to relive the past scene with him/her. This is not the case with *je me dis*, which does not have this staging component.

Concerning the multimodal conduct of the speaker while quoting theirselves with *je me dis* and *je suis là*, it is difficult to draw meaningful conclusions due to the low number of

occurrences. There are some tendencies that seem to be related to the actions that speakers carry out with the respective formats, even though their actual validity must be treated with caution:

(1) The degree of bodily reenactment is higher with *je suis là* than with *je me dis*, as I have already stated for the respective past forms (see Section 6.3.3 for a summary). My quantifying analysis has shown that headshakes ($n = 6$) or hand gestures (e.g., PUOH gesture; $n = 8$) are recurrent with *je me dis*, whereas facial expressions are rare ($n = 1$). Compared to that, *je suis là* introduces in four of five cases a facial expression coupled with the speaker's gaze away during the quote, which may be related to the dispreferred character of the strong affective stance (see also Section 7.4).

(2) The speaker's gaze conduct varies between [*je me dis* + direct reported thought] and [*je suis là* + quote]. As for *je me dis*, 23 of the 28 occurrences are accompanied by the speaker's gaze at the recipient during their self-quotation or at the latest at the end of the quote.[233] As for *je suis là*, in four of the five occurrences, speakers gaze away during the quote and only gaze back at their recipient(s) at its end, which may be due to the delicateness of the telling during which speakers take a strong affective stance. In contrast, *je me dis* occurs in sequences of reasoning, where the described scenes are usually not emotionally charged and less delicate.

The comparison of *je me dis* and *je suis là* introducing direct reported thought or speech has shown that there are functional and formal differences between them. However, their functional distribution seems to be complementary in that they are, in my data, used according to the actions that speakers accomplish with them. In what follows, I will go one step further by comparing the French social action formats *je me dis* and *je suis là* with the German *ich denke (mir)*. Some of the specificities of each format will become more obvious when comparing it to another language. To do so, I will also use excerpts that have already been discussed in my preceding qualitative analysis. New excerpts will be introduced in more detail.

7.4.   The Comparison of French and German Direct Reported Thought in Present Tense: Formal and Functional Aspects

After analyzing each social action format in narrative present tense individually, this section is dedicated to their cross-linguistic comparison. I will present two pairs of data excerpts with one German and one French excerpt each. These pairs illustrate the main converging types of action formation that occur with direct reported thought with *ich denke (mir)* ('I think (to myself)') and *je me dis* ('I say to myself') and *je suis là* ('I am there'): taking an affective stance and taking an epistemic stance to display a reasoning.

The comparison will address formal and sequential aspects, the actions that speakers accomplish with the social action format as well as multimodal features. Because of my small data set and the multitude of characteristics that I investigated (prosody, syntax, sequentiality and multimodality), it is beyond the frame of the present work to quantify the exact number of converging cases. Further analyses with a larger data set, focusing on each of these characteristics individually, is necessary to confirm the following tendencies:

---

[233]   In three further cases, speakers gaze at an object involved in the ongoing conversation; only in the remaining two cases do speakers gaze away beyond the end of the quote. One may be because the interlocutors are seated next to each other instead of facing each other, which leads to a lower number of gazes at each other.

1) Excerpts 7.13 & 7.14: *ich denke* vs. *je me dis*, allowing the speaker to make publicly available an emergent rationalization in sequences of reasoning. In German, this use is rare in present tense (*n* = 5) compared to French (*n* = 28).

2) Excerpts 7.15 & 7.16: *ich denke* vs. *je suis là*, allowing the speakers to stage their responses toward a reported first pair part while displaying their negative affective stance in a storytelling. The stance-taking is mostly displayed through bodily means or a response cry before verbal resources are added.

Excerpts 7.13 and 7.14 are converging instances of direct reported thought, with the German *denk ich* (lit. 'think I') and the French *je me dis* ('I say to myself'). Although the quotatives do not converge lexically, they are used in a similar way within the sequence. In both excerpts, the speaker weighs two different possibilities against each other and makes publicly available her reasoning. In Excerpt 7.13, three friends who live and study in Augsburg[234] talk about the internship Rabea applied for. She seems to have a good chance to get an internship in Würzburg but does not display much enthusiasm about it prior to this excerpt.

Ex. 7.13: hierbleib / stay here (KAFKU_01, 49min48)

```
LIS: Lisa, RAB: Rabea, JUL: Jule

01 LIS: aber (.) hast du dir nur sachen in
        but      did you only look for things in

        würzburg rausgeSUCHT,=
        würzburg

02      =oder [auch in augsburg und so;   ]
         or    also in augsburg and stuff

03 RAB:       [NEE auch paar sache in AUGS]burg-
               no  also a few things in augsburg

04 LIS: o[keh-]
        okay

05 RAB: [°und] +ja.° (0.4) richtung donauwörth-
         and    yeah       near donauwörth
   rab         +gz down-->l.11

06 LIS: [°hmhm°,]

07 RAB: [°hh ich] weiß net-=
             I don't know

08      =einerseits wollt ich ja schon voll gern nach WÜRZburg-
         on the one hand I really wanted PART PART to go to
         würzburg

09 JUL: [((clears throat))]

10 RAB: [aber irgendwie   ] #(1.0) deng_ich halt au JA:-
         but  somehow              I think PART also well
   fig                      #fig1
```

fig1

```
11      %wenn ich% jetz noch +HIER$bleib-
         if I still  stay here now PART
   rab  %shrug-->%
                          -->+gz at LIS-->1.13
   lis                           $nods-->1.12
12      (0.3)$#(0.3)$
   lis  -->$shrug $
   fig       #fig2
```
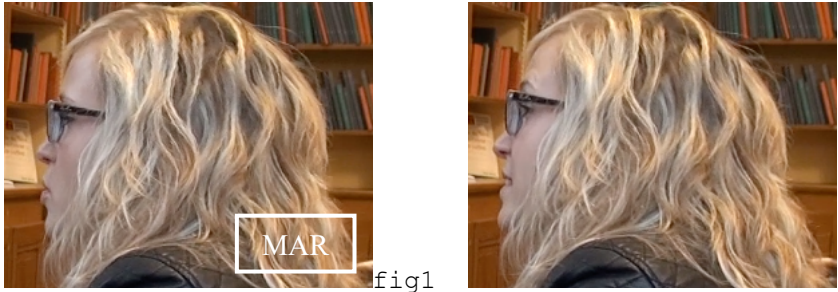

fig2

```
13 RAB: °dann is au nich° (.) +SCHLIMM- %HA[!HA!  %
         then it's also not    dramatic  ((laughter))
   rab                   -->+gz at JUL-->>
14 LIS:                                  [hehehe% [hehe-]
   rab                          %shrug--->%
15 JUL:                                          [nö;  ]
                                                  no
```

In the French excerpt, which I have analyzed in Section 7.3.1.3, Marie weighs having a good job against having a family. The excerpt starts with Marie's response to Patrick's description of his own stance—namely, that he does not really ask himself this question.

Ex. 7.14: ça me stresse / it stresses me out (Pauscaf 13, 09min40)

```
MAR: Marie, PAT: Patrick

07 MAR:                       [alors-   ]
                               well
08      (0.3) moi j'suis vraiment entre les DEUX en fait;
              for my part I'm really  between the two of them PRT
09      (0.2)$(0.2)
   pat   -->$gz down-->1.10
10      j'suis pAs là à$ fONd en train de viser la carrIÈRE,
         I'm not   there totally focussing on aiming for a career
   pat             -->$gz at MAR-->1.15
11      (0.7)
12      mais j'ai quand même envIE d'avoir un beau méTIER,
```

243

```
           but I would like to have a nice job though
13       (0.2)
14       pis j'suis pas là: non plus: 'fin: OUI# +j'me %dis#:
           then I'm not either there    well yeah I say to myself
   mar                                        +gz at left->l.15
                                                   %eyebrow
                                                   raise->
   fig                                       #fig1        #fig2
```

fig1

fig2

```
         ça m'stresse et %tOUt:,=
           it stresses me out and all
   mar               -->%
15       = +mais: sans plus non ↑PLUS$ %'fin: voilÀ?
           but not that much neither PRT PRT
   mar  ->+gz at PAT—>l.17
                                     %shoulder raise-->l.17
   pat                        -->$gz down-->l.17
16       (0.3)
17 PAT: mais si on te pro%pose $d'aller faire un job à l'étranger
           but if you get an offer to leave for a job abroad
   pat                         -->$gz at MAR-->>
   mar                   -->%
         euh: l'année prochaine tu PARS?
           uh   next year would you go?
         (0.3)+(1.0)
   mar    -->+gz left-->>
18 MAR: chais pas;
           dunno
```

Both speakers, Rabea (German) and Marie (French), display their uncertainty about a next step in their lives: for Rabea, choosing between different internships; and for Marie, deciding whether she wants to have a family first or focus on her career. In both excerpts, Rabea (lines 03–13) and Marie (lines 08–15) respond to first pair parts (in German: lines 01–02; in French: not in transcript) with an emergent reasoning which allows them to make publicly available the competing options that they think about.

In French and German, after presenting a first scenario (in German: going to Würzburg (line 08); in French: her stance toward a potential job (lines 08–12)), the second option, or—in French—the conclusion of the dilemma is uttered with the respective social action format *ich denke* and *je me dis* introducing direct reported thought.

In German, the second option is already projected at the beginning of Rabea's TCU with *aber irgendwie* ('but somehow'; line 10) and followed by a pause, which allows Rabea to stage

her thinking process as if it was made publicly available upon emergence. Note that during this pause, Lisa starts already smiling as if she had anticipated Rabea's reluctance to leave Augsburg (fig1). Rabea's direct reported thought is introduced with *deng_ich*, which is hedged by a particle and followed by a lengthened *ja* (here, 'well'), displaying Rabea's hesitation. The second option, to stay in Augsburg[235] (line 11), is formatted as a conditional clause. The protasis is accompanied by a shrug and followed by a positive assessment (line 13) as apodosis, which is closed with loud laughter. During the pause at line 12, Lisa shrugs and raises her eyebrows (fig2), thus already affiliating with Rabea's being undecided. So does Jule, who, for the first time in this interaction, takes the turn with an affiliative *nö* ('no'; line 15). The direct reported thought allows Rabea to close her reasoning with a conclusive thought that she makes available to her interlocutors. Her thinking process is thereby produced as emerging in real time with pauses (lines 10 and 12), lengthenings, and, until line 11, a gaze down. This multimodal assembly allows Rabea to stage her thinking process, which seems to prevent her interlocutors from taking the turn until her reasoning is closed (line 13). Lisa's bodily responses at lines 11 (nods) and 12 (shrug, facial expression) are a "silent" possibility to nevertheless display her affiliation. The format of direct reported thought allows Rabea to present her conclusion as something that "just came up," as one thought among others, thus lowering her commitment in terms of actually realizing these thoughts.

In French, Marie first presents two important arguments related to a hypothetical job (lines 08–14). The two intonational units are separated by a 0.7 s pause (line 11), during which the recipient does not take the turn, probably due to the rising final intonation at line 10. The two contrastive arguments are, after another brief pause (line 13), followed by a third argument (line 14) which remains syntactically incomplete and contains several lengthened syllables (on *là*, *plus*, *fin*, and *dis*). After this hesitant turn, Marie changes trajectory (marked by '*fin oui* ('well yes'); line 14) and introduces her reported thought. With *je me dis*, she makes publicly available the conclusion to her prior weighing activity. As in the German excerpt, the pauses (lines 11 and 13, but also the dispreferred initial pause at line 09) and lengthenings (lines 14–15) allow Marie to display her reasoning as emergent in the ongoing interaction. The reported thought thereby seems to be a means to close the monological reasoning sequence. Indeed, the turn-final *voilà* with rising intonation at line 15 is responded to by Patrick who, other than in the German excerpt, disaffiliates by challenging Marie's rationalization.

None of the quotes is bodily staged. Instead, Rabea (line 12) and Marie (line 15) only shrug during their quote, a movement which may here "[express] a meaning of lessened engagement on the part of the speaker" (Debras 2017: 24). In this specific case, it may display "lessened epistemic endorsement (ignorance, uncertainty)," as neither Rabea nor Marie seem to be able to know (or influence) what their future will look like. Note that both speakers also gaze away during the beginning of their quote, and gaze at their interlocutor(s) again when approaching the end of their turn (German: line 11 at Lisa, line 13 at Jule; French: line 15). This gaze conduct matches prior research on gaze patterns allowing speakers to close a turn and leave the floor to a potential next speaker (Auer 2021; Kendon & Cook 1969).

---

[235]   The German *hierbleiben* ('stay here') refers to the city where they live and where the recording takes place.

The two excerpts have illustrated that speakers may use direct reported thought with *ich denke (mir)* and *je me dis* to display their purported emergent reasoning toward the end of a sequence. In German, this format is used rarely (*n* = 5). In French, speakers use, in my data, exclusively *je me dis* and not *je suis là* to display their reasoning.

The next two excerpts show an instance of the German *dann denk ich mir* ('then I think to myself') and of the French *je suis là* ('I am there'). Both of them are used in complaint sequences where the speakers first report what a third party says and then stage their reported thought as spontaneous response to that first pair part. In French and German, the direct reported thought allows the speaker to take a negative affective stance toward this first pair part. In Excerpt 7.15, Daria tells her mother that when she goes on a date, she wants the man to plan and prepare everything.

Ex. 7.15: ganzes Leben / whole life (FOLK_E_00327)

```
DAR: Daria, PAT: Patricia

01 DAR: ich bin echt gern n TYP,
        I really like to be the kind of person

02      (0.65)

03 DAR: der (.) sich gern mItnehmen lässt auf_s DATE.
        who     likes to be taken to a date

04      (0.59)

05 DAR: wo einer dann sagt wir treffen uns DORT
        where one says we meet there

        und machen dann DAS,
        and then we do that

06      und wenn er dann sagt ja was willst DU denn mach↑en-=
        and when he then says well what would you PART like
        to do?

07      =>dann denk_ich mir< %NEE-%
         then I think to myself no
   dar                     %head
                            back%

08      (0.49)

09 DAR: ich plan mein GANzes l(h)eben j(h)etzt pl(h)anst
        I plan my whole life now it's your turn to plan

        %DU mal w[as;% he h°]
        PART something
   dar  %points tw
         mother-->l.10

10 PAT:          [geN%AU.   ]
                  exactly
   dar            -->%

11      (0.52)

12 PAT: geb ich dir vollkommen RECHT.
         I totally agree
```

In the French excerpt, Damiane responds to her mother's opinion about how to deal with the conflict in Syria.

Ex. 7.16: population / population (Pauscaf 27,12min25)

```
DAM: Damiane, GAE: Gaëtan

01 DAM: c'est plus pou- euh >par rapport< au conflit en syRIE-=
        it's more fo    uh   about the conflict in syria

02      =ma mère est vraiment LÀ-=
         my mother is really there (idiom. 'like')

03      =>mon dieu< mais c'est +horrible ce qui se passe pour la
         my god    but   that's horrible what is happening for the
   dam                           +gz at GAE-->l.05

        populaTION-=
        population

04      =il FAUT intervenir,
         we must intervene

05      (1.1)+
   dam    -->+gz right-->l.06

06 DAM: chuis# là% mais ↑↑NON,# hhh°+hinhin% °H
        I'm there  but   no
   dam                          -->+gz at GAE-->>
        %shrug------------------>%
   fig        #fig1---------->#fig2
```
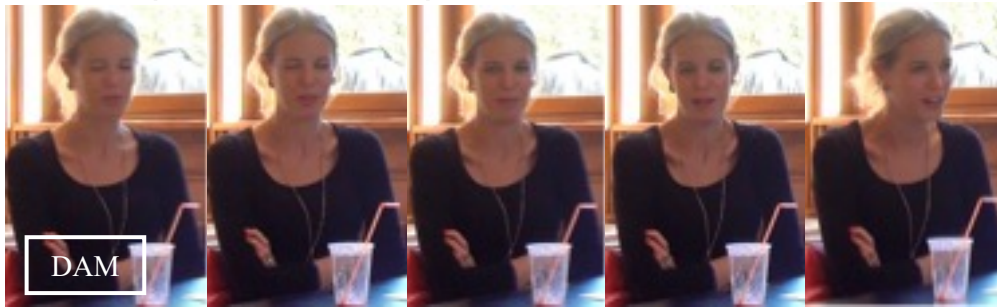


```
        fig1                                    fig2

07      de quel ↑DROIT?=
        by what right

08 GAE: =non c'est trop [TARD- c'est fiNI] >d'toute< façon.
         no it's too     late  it's over     anyways

09 DAM:                 [de quel DROIT-  ]
                         by what right
```

In both excerpts, the speakers first report speech from a third party. Daria imitates a man (a potential date; line 06), while Damiane reports her mother's utterance (lines 02–04). While in the German excerpt, Daria continues directly with her responsive thought (line 07), Damiane first pauses (line 05), thus increasing the dramatizing effect of her telling.

In German, the direct reported thought is projected by an accelerated *dann denk ich mir* ('then I think to myself'). Daria first produces a stressed *nee* ('no') and a head movement back, thus disaffiliating with the reported (hypothetical) question. After a short pause (line 08), she continues her imagined replica to the potential date with laughing voice (line 09), while pointing toward her mother who is standing in front of her. The pointing suggests that Daria makes her mother the imagined interlocutor and thus part of the narrated scene. In overlap, Patrizia strongly affiliates with her daughter's negative affective stance (lines 10–12).

In French, Damiane responds to her mother's stance with direct reported thought projected by the social action format *chuis là* ('I'm there'). It is followed, as it is in German, by a strong disaffiliation through a prosodically stressed *mais NON* ('but no'). With her facial expression and her head movement (fig1 to fig2), Damiane simultaneously displays her negative affective stance toward her mother's utterance. Similar to German, Damiane extends her turn, but with a rhetorical question which is followed by Gaëtan's verbal affiliation.

Note that Daria and Damiane start laughing during or just after their strong disaffiliation (German: line 09; French: line 06). Their laughter may function as a hedge for both speakers' strong disaffiliation (Clift 2016) with the first pair part in the telling. The format of the quote (bodily animation, prosodic stress, and pitch movement), and the quotative in narrative present tense may increase the pressure on the interlocutor in the ongoing conversation to respond as Patrizia's and Gaëtan's strong affiliation demonstrates.

Summary

My cross-linguistic analysis has demonstrated that there are formal and functional similarities between French and German direct reported thought with *ich denke (mir)* ('I think (to myself)') and *je me dis* ('I say to myself') / *je suis là* ('I am there') concerning their position in the sequence and action formation. One general result of my analyses, despite some differences in number, is that the actions that speakers accomplish with direct reported thought in narrative present tense seem to generally converge with those in past tense. Compared to the same quotatives in past tense, the convergences are, however, less evident. I will first address some functional aspects before I move on to formal and multimodal convergences.

(1) There are fewer converging cases of direct reported thought allowing speakers to make publicly available their reasoning due to the rarity of the format in German. Whereas in past tense, *ich habe (mir) gedacht* is recurrently used by speakers for this action (see Section 6.3.6), the present-tense format is rarely used in this function. It is likely that speakers are also able to make a reasoning publicly available in German like French speakers do with [*je me dis* + direct reported thought]. This action just does not seem to be accomplished with direct reported thought—or, at least, not with direct reported thought introduced by *ich denke (mir)*. As is to be expected, however, this activity must be done in German, too. A hypothesis for further research may be that there are other formats (e.g., formulaic expressions) to accomplish this task. I therefore suggest analyzing reasoning or argumenting sequences to determine the social action formats that allow speakers to make their reasoning publicly available in German.

(2) Because *je suis là* ('I am there') is so rare in my data, there cannot be more than five converging cases between the French and German social action formats allowing speakers to take an affective stance. As I pointed out in Section 7.3.3, it seems that speakers routinely use the past tense form *j'étais là* to display their affective stance even if the surrounding telling is in narrative present tense. In German, speakers resort to *ich denke (mir)* to accomplish this action, thereby exploiting narrative present tense as a presentifying resource; there are cases where speakers switch to narrative present tense only for the direct reported thought, which enhances intersubjectivity and the dramatizing effect of the affective stance. This recreation of the original scene seems to shift response relevancies from the original telling toward the

ongoing telling, which may increase the pressure on the recipient to deliver a subsequent response to the strong affective stance.

(3) With *ich denke (mir)* and *je me dis*, speakers can stage their thinking process as an action emerging from prior (reported) events. In this case, the direct reported thought is positioned at the interface of the original, reported situation and the ongoing conversation. The narrative present tense of the social action format of the quotative enhances this ambiguity; speakers present their thought as if it were being assembled step-by-step while still thinking of the words to say (instead of recalling them). By doing so, speakers may blur the lines between the telling and ongoing talk in a way that the staged thinking process may be presented as valid in both speech situations, thus also altering the contingencies of ongoing talk.

There are also formal similarities, most importantly the syntactic independence of the quote from the quotative in both languages. The *direct* reporting format is generally defined by the syntactic non-linkage of quotative and quote, which is why, in none of both languages, does the quote have any syntactic characteristics that would point to syntactic linkage.

Even in German, where clause-combining formats between embeddedness and independence exist (i.e., the dependent main clause structure), it is rare that there is not at least a loose syntactic relation between quotative and quote when verbal material is quoted. However, my analysis has shown that, even if a dependent main clause occurs after *ich denke (mir)*, it is often preceded by a response cry or body movement that separates quotative and quote (*n* = 17/22), which makes a syntactic bind between the two parts of the bipartite structure questionable. There are only five cases where a complete clause follows *ich denke (mir)* without a response cry or body movement preceding it.

This tendency is less obvious in French, where an in-between format of a dependent main clause does not exist. Subordination is marked by a complementizer, juxtaposition by the absence of it. The ratio of quotes as main clauses with and without response cry is, however, inversed: there are 18 out of 28 cases where *je me dis* is followed by a main clause, and only three of them are preceded by a response cry. In the case of *je suis là*, the question of syntactic linkage is irrelevant, as I have explained in Section 7.3.2. It is, however, interesting that in all five occurrences of it, the two parts of the bipartite structure (quotative and quote) are separated by a response cry or a body movement.

What can be concluded from these results is that, in the specific case of direct reported thought, it may be necessary to describe the grammatical format [quotative + quote] with a larger concept than classical syntax may be able to provide. As with past tense quotatives, the term *bipartite structure* seems an appropriate description for a format that does not only combine clauses with other clauses but also with non-lexical vocalizations or body movements.

There are also some convergences on a multimodal level. My comparative analysis has brought to light that there are mainly two recurrent gaze patterns accompanying the present-tense quotatives in French and German. Speakers either shift their gaze to the recipient during the quote (at the latest at its end), which allows them to check the recipients' attention or availability to respond or to yield the turn (= pattern 1), or they do not gaze at the recipient at

all during the self-quotation (= pattern 2)[236]. The following table shows these two gaze patterns:[237]

Form	Frequency pattern 1	Frequency pattern 2
ich denke (mir)	58.4% (31/53)	18.9% (10/53)
je me dis	82.1% (23/28)	7.1% (2/28)
je suis là	80.0% (4/5)	20.0% (1/5)

Table 44: Gaze patterns co-occurring with narrative present-tense quotatives

As my qualitative analysis has shown, the occurrences where speakers continue gazing away and do not gaze back at the recipient occur in what can be called *delicate situations* (for a similar observation about 'and then I'm really like' occurring in the environment of gossiping, see Lamerichs & Te Molder 2009). Such situations are telling sequences where speakers report (1) embarrassing situations they were involved in, (2) dialogues in which they report their dispreferred response to a prior turn, or (3) a strong negative affective stance toward a reported issue, such as a third party's behavior. In German, nine of 10 occurrences of this gaze pattern with *ich denke (mir)* are part of tellings about delicate situations.[238] I am aware that these are small numbers that do not allow generalization, especially considering the variability of my data; however, the speakers' gaze conduct is so consistent in this context that I suggest treating the co-occurrence of direct reported thought with the speaker's averted gaze in delicate situations as a multimodal assembly.

The comparison of French and German social action formats in present tense projecting direct reported thought demonstrated that there are functional and formal parallels between *ich denke (mir)* and *je me dis / je suis là* despite their lexical difference. Starting from the action of reporting thought allowed me to identify parallels in action formation in the two languages: displaying their affective or epistemic stance. The multimodal analysis confirmed tendencies that I have already located for the same social action formats in past tense: when speakers take an affective stance, facial expressions are recurrently used to display this stance, often before its verbal formulation. In contrast, when an epistemic stance is displayed, headshakes and hand gestures occur more often than facial expressions. The distribution of these two actions is, however, different between the languages: speakers use *je me dis* to display a reasoning, and, much less often, *je suis là* to display their affective stance. In German, speakers use *ich denke (mir)* almost exclusively to take an affective stance. This result does not mean, however, that the actions that are less often accomplished with the respective quotatives in French or German are non-existent; they are simply not, or are less often, accomplished with direct reported thought introduced by *je me dis / je suis là* or *ich denke (mir)*. The investigation of other formats

---

[236] As mentioned before, these quantifications must be treated with caution as the data sets in German are diverse and do not all have the same contextual setting (see Chapter 5 for a detailed overview). Also, because the participant frameworks in both the German and the French corpora range between different numbers of participants, the gaze patterns will need to be investigated in more detail in future research.

[237] The numbers do not add up to the total number of occurrences because there are also in-between formats in which gaze cannot be distributed according to these two patterns.

[238] In French, during the one occurrence of *je suis là* in which the speaker continues gazing away even after her quote, she takes a strong negative affective stance toward a reported situation. In the one case of *je me dis*, the speaker delivers a dispreferred response to the interlocutor's prior suggestion.

than the pattern [quotative + direct reported thought] would reveal additional results that may account for this gap.

8.   Discussion and Conclusion

> *Man is least himself when he talks in his own person.*
> *Give him a mask and he will tell you the truth.*
>
> *- Oscar Wilde*

In the present work, I analyzed past and present-tense quotatives of direct reported thought in French and German talk-in-interaction. The multimodal, cross-linguistic analysis of the format [quotative + quote] brought to light new results concerning action formation that had not been addressed in prior research. One main result of this is that direct reported thought can be classified as "one type of demonstration" (Clark & Gerrig 1990: 802) that may share certain formal parallels with direct reported speech while accomplishing different actions in everyday talk. In the following section, I first summarize the main findings for each language individually before presenting the cross-linguistic results.

Summary of the main findings in German

All three quotatives introducing direct reported thought—*ich denke (mir)*, *ich dachte (mir)*, and *ich habe (mir) gedacht*—allow speakers to display their affective or epistemic stance in tellings, only with different quantitative distributions. Because *ich denke (mir)* in present and past tense is multi-functional (see Deppermann & Reineke 2017), I investigated how the complement-taking predicate becomes recognizable as a social action format for introducing a quote without complementizer. One clear feature in my data is that *denken* is used as a social action format projecting direct reported thought only when it co-occurs with additional lexical material like prepositioned *und* ('and') and *da/dann* ('there/then') and post-positioned *so* ('like') or as a verb-first-structure or coordination ellipsis. If none of these are the case, *ich denke (mir)* / *ich dachte (mir)* / *ich habe (mir) gedacht* are used for other types of actions, such as displaying a discrepant assumption (Jefferson 2004; Deppermann & Reineke 2017, 2020), instead of introducing reported thought.

I argue that these co-occurring lexical items may contribute meaningfully to the disambiguation of the CTP *denken* while also situating the upcoming quote with adverbs or conjunctions within the narration in which it usually occurs. Deictic items like *da* ('there') and *so* ('like') may heighten the projective force of the social action format by referring to the upcoming (potentially bodily) depiction that recipients are supposed to draw their attention to. As, according to Plank (1986), a direct quote does not include deictic shifts, it is probable that this is also valid for deictics like *da* ('there') that co-occur with the introducing CTP (see my analysis of *j'étais là* in Section 6.3 for a similar phenomenon). Also because of this systematic co-occurrence of *ich denke (mir)* in present and past tense with additional lexical items, the quotative is becoming a routinized pattern (bearing variation in the additional lexical material) assembled "on the fly" according to the necessities of the ongoing conversation.

Concerning action formation, my analyses suggest that there is a tendency for *ich dachte (mir)* to be used for affective stance-taking whereas *ich habe (mir) gedacht* is used for both affective and epistemic stance-taking to display an emergent reasoning. This distribution of

preterit and present perfect according to action formation has not been investigated in studies of German spoken language so far; it may need further exploration with a larger data set and other *verba sentiendi*, like *glauben* ('to believe') or *finden* ('to find'). In present tense, *ich denke (mir)* mostly projects affective stance-taking; only few occurrences are used by speakers to display their emergent reasoning. This result suggests that, unlike with *ich habe (mir) gedacht*, displaying an emergent reasoning is not systematically done with the present tense quotative *ich denke (mir)*.

The two kinds of actions that speakers accomplish with direct reported thought introduced by *ich denke (mir)* in present and past tense have also been detected thanks to differences in bodily conduct. When displaying an affective stance, speakers systematically depict their stance with multiple bodily resources—much more than when displaying an epistemic stance. Facial expressions are more likely to co-occur with affective stance-taking, while shrugs and palm-up open-hand gestures are more recurrent with displays of an emergent reasoning. Analyzing multiple resources allowed me to demonstrate tendencies of potential multimodal assemblies. Further research will need to (1) investigate the possible relation between action formation and the use of specific resources more thoroughly, (2) unravel differences in multimodal conduct depending on the conversational setting (stationary vs. non-stationary) and participant framework, and (3) investigate in more detail the recipients' multimodal conduct.

Syntactically, in most cases, there is either a loose or no syntactic link between the two parts of the bipartite structure [*denken*-quotative + quote]. In cases where the social action format is followed by a complete clause (59.3%), the syntactic configuration is exclusively [quotative + dependent main clause]. More embedded clauses like the independent subordinate clause do not occur in my data. In 40.7% of the cases (38.7% for both past tenses and 46.2% for the present tense), the quotative does not even introduce a complete clause; it rather introduces non-clausal structures, like body movements, response cries, and/or verbless structures. This result indicates that the grammatical description of the direct reported thought format as *bi-clausal structure* may need to be abandoned in favor of the larger conceptualization of *bipartite structure* (for a discussion, see *infra*). The clause-combining format of direct reported thought in German is thus, in line with Gast and Diessel (2012), the combination of a clause with "some other constituent" (*ibid*.: 4), whereby this other constituent includes verbal and bodily conduct.

Summary of the main findings in French

My qualitative analysis has shown that there is a strong tendency for the French quotatives *je me dis / je me suis dit* and *je suis là / j'étais là* to be distributed according to the actions that speakers carry out with direct reported thought. While *je me dis / je me suis dit* are used to display an emergent reasoning—and thus for more rational, non-affective post-hoc rationalizations—*je suis là / j'étais là* allow speakers to display their affective stance.

These two types of action formation come with different degrees of bodily reenactment: whereas *je suis là / j'étais là* introduces almost exclusively bodily and prosodic depictions that either accompany verbal utterances or do not, *je me dis / je me suis dit* projects quotes that contain fewer reenactments—in other words, speakers use fewer resources for action formation. Because the French data is consistent concerning the conversational setting, the results concerning the use of multimodal resources has greater informative value than the German data.

However, future research needs to address potential divergences concerning dialogical vs. multi-party conversations.

The general functional distribution between *se dire* and *être là* is more striking with the past forms than with the present forms. My analysis has shown that, compared to *j'étais là*, *je suis là* is rare. This quantitative result raised the question why such an evident difference in number exists between present and past tense with *être là* but not with *se dire*, which is similarly frequent in present and past tense. The quantitative distribution revealed that *j'étais là* seems to have routinized into a quotative that is still morphologically marked as *imparfait* but is on its way to losing its imperfective aspect. My analysis has demonstrated that there are cases in which speakers shift from narrative present tense throughout their whole telling to *j'étais là* to introduce their quote, which is also in narrative present tense. This result is a strong indicator of *j'étais là* routinizing into a chunk that projects an upcoming quote.

Concerning the distribution of tense more generally, in French, the quotatives *je me suis dit* and *j'étais là* complement each other in their quantitative distribution between synthetic and analytic past tense. This distribution suggests that "[t]he reenactment format may [...] be selected by reference to the type of action being accomplished in the talk" (Sidnell 2006: 406). My analyses have shown that *se dire* in past tense almost exclusively occurs in the analytic past form whereas *être là* exclusively occurs in the synthetic past form. This distribution may be due to a relation between aspect and action formation; as with *j'étais là,* speakers display their emotional state that has neither a specific beginning nor end, and the imperfective aspect is the corresponding format to use. In contrast, *je me suis dit* is mostly used to draw conclusions and report decisions as a consequence of prior events—in other words, the social action format projects punctual events that occur in temporal relation to what precedes them, often also marked as such through temporal adverbs like *pis* ('then') and *et* ('and'). It may be fruitful to pursue this line of research by investigating the distribution of imperfective and perfective for the same action in French talk-in-interaction, especially because prior research on French past tenses has mostly focused on the comparison of *imparfait* and *passé simple* in written texts.

Another important result of the present work is that *j'étais là* and *je suis là* are the only social action formats in the present work that allow speakers ambiguity, as the quotatives do not frame the quote specifically as thought or speech. I have shown in my analysis that this ambiguity is, however, not interactionally problematic in my data, as recipients do not topicalize it as unclear; on the contrary, Excerpt 6.15 clarifies that, probably through the contextual information from the reported situation and from shared social norms between participants, the recipients can guess that the quoted material has been thought.

One context where *j'étais là* projects direct reported speech instead of thought is in constructed dialogues where a turn-taking between two parties is reported. However, because speakers mostly report strong negative assessments with *je suis là* and *j'étais là*, sometimes about delicate matters like sexual orientation, relations, etc., this ambiguity may be an interactional tool allowing speakers to present strong opinions as reported thought, thus distancing them from the content. By doing so, speakers may be able to render their opinions more explicitly than when they said or thought them in the original situation.

Summary of the main cross-linguistic findings

Even though the social action formats projecting direct reported thought are lexically different in French and German, they are used to accomplish the same actions: affective and epistemic stance-taking to display an emergent reasoning. As my analysis has shown, the quotative itself plays a more important role in the process of action formation and action ascription than thus far expected (Fox & Heinemann 2017; Levinson 2013; Sidnell 2017). Whereas *ich denke (mir)* in present and past tense introduces direct reported thought that displays both kinds of stances, *je me dis* / *je me suis dit* and *je suis là* / *j'étais là*, two lexically different quotatives, are distributed according to their function.[239] Also, the parallelization between direct reported thought in French and German follows action formational principles more than it does lexical principles.

Cross-linguistically, the parallels in action formation regarding direct reported thought are between

- *je me suis dit* and *ich habe (mir) gedacht*, to make publicly available an emergent reasoning (=epistemic stance), and
- *ich dachte (mir)* and *j'étais là* as well as *ich denke (mir)* and *je suis là*, to display an affective stance.

Je me dis is the only social action format that does not find its equivalent in similar formats with *ich denke (mir),* as the German quotative is mostly used for affective stance-taking.

In both German and French, the actions of quoting and stance-taking seem to be distributed among different, laminated resources; When the quoting itself is done through language, speakers use (often simultaneously) bodily means—especially facial expressions—to display their affective stance. This observation is similar in both languages, suggesting that the bodily conduct is, for French and German direct reported thought, closely intertwined with action formation.

The comparison of the two languages has proven fruitful for this analysis in several ways: not only did it allow me to discover practices in both languages that I may not have detected if had I not started by looking at one language first, the cross-linguistic analysis also demonstrated that reporting thought functions similarly in these two languages and that speakers use the format to accomplish the same actions. Even though the quotatives are not lexically the same, action formation and ascription are largely comparable. Even the patterns of clause-combining are similar; one clause (the quotative) is combined with another constituent (the quote), which may be verbal and/or non-verbal. Even though German and French syntax are generally described as different, especially regarding clause-combining, the present work evidenced that direct reported thought is systematically formatted as a bipartite structure in which a quotative projects a syntactically independent quote. Despite the distinct grammatical status of the French quotatives *je suis là* and *j'étais là* compared to all other analyzed CTP-quotatives, they are formatted in a comparable bipartite structure.

---

[239] One could also argue that *j'étais là* formally resembles more *(und) ich so* ('(and) I like') because both of them have a demonstrative element: *so* and *là*. But that *(und) ich so* cannot be marked regarding tense and aspect makes comparison with *être là* (and *se dire* ('to say to oneself') for that matter) difficult.

My cross-linguistic analysis leads to the conclusion that the investigated complement-taking predicates in French and German as well as *je suis là / j'étais là* ('I am/was there') are in the process of routinizing as formats recurrently introducing direct reported thought. What is special about the constructions I investigate in this work is that, despite a tendency to routinize, there is one important feature that quotatives do not share with epistemic parentheticals (Thompson & Mulac 1991a) or formulaic stance markers (Thompson 2002) of the type *I think*: Quotatives are not positionally flexible. Hence, I argue that quotatives may have routinized into a specific format—namely, projector constructions (Pekarek Doehler 2011). Projector constructions—or *projector phrases* (Günthner 2011b)—are reanalyzed clauses used for "announcing upcoming talk by the same speaker, for holding the floor and for constructing complex turns" (Pekarek Doehler 2011: 106), thus "focusing recipient's attention on the following segment, which expresses the core message" (Günthner 2011b: 11). Günthner (2011b), investigating structures of the type *das Ding ist…* ('the thing/point is…'), shows that, in such structures, the complement clause undergoes a "functional upgrading […] accompanied by formal indications of prosodic and syntactic independence" (*ibid.*: 11), an observation that comes close to Thompson's (2002: 155) observation of an upgrade of the complement. Not all quotatives in my data fulfill the characteristics of such a projector-construction in the same way; There are several degrees of routinization depending on both the CTP (or the *je suis là / j'étais là* quotative) and the type of complement that follows. The most routinized social action format to quote direct reported thought is *j'étais là* as I have explained in detail in Section 7.3.3.

My comparative analytical chapters showed that there are not only parallels on a structural level—the kinds of body movements that accompany the two kinds of stance-taking show comparable tendencies too—but also in their distribution (e.g, facial expressions are more recurrent with affective stance-taking). However, because of my small data set, these tendencies should be checked with a larger corpus.

Discussion of the main findings

My multimodal analysis of direct reported thought demonstrated that bodily conduct is crucial for action formation. This finding, however, complicates the grammatical description of clause-combining patterns with quotatives containing complement-taking predicates (*ich denke (mir)* and *je me dis* in present and past tense), which have been traditionally described as bi-clausal structures with a loose syntactic bind between its two parts. After close investigation of the combining patterns in my data, it is more appropriate to describe the format [quotative + quote] as a bipartite structure, in which an almost exclusively prepositioned enquoting device in both languages projects an upcoming quote. My analysis demonstrated that quotes do not only consist of complete clauses; Instead, their description must include gaze, facial expressions, gesture, and posture—in other words, all the resources that speakers use to depict their purported thoughts from the past.

How does this larger definition of clause-combining patterns influence the grammatical status of the quotative? First, all quotatives that I investigated in the present work project the quote that follows them, regardless of whether the quote contains a complete clause. This projection is, first and foremost, an interactional projection due to the recurrent use of the social action formats that project a quote. However, I argue that *(sich) denken* and *se dire* being

complement-taking predicates maintains grammatical projection as at least somewhat relevant; as my analysis demonstrated, the principles of government and constituency still play a role when adopting a larger definition of *complement*—for the specific case of direct reported thought—including formats other than a complete clause.

To establish a possible grammatical description of direct reported thought, I suggest, in line with Couper-Kuhlen (2018), to conceptualize the open slot of the quote projected by the quotative as one syntagma, in which verbal and embodied elements either are interchangeable or co-occur. My analysis has shown that participants treat verbal and bodily quotes equally—what seems to be crucial is that the social action format functions as the device that situates (Brünner 1991) the introduced material (prospectively) as a quote, as it allows for an emergent continuation of the turn without restricting its shape. Depending on the action that speakers accomplish with the whole pattern [quotative + quote], more or fewer depictive resources may be necessary. In this sense, a grammar-in-use of direct reported thought may be best described with Couper-Kuhlen's suggestion that "[g]rammar must be able to account for the slot but not necessarily for what fills it" (*ibid.*: 24). If a grammar of direct reported thought accounted for what fills the slot, the grammar would need to be extended by bodily resources, as suggested by Keevallik (2018). Such a detailed description would necessitate a more fine-grained analysis of all individual resources that speakers deploy while quoting thought. Due to the multitude of features that I address in the present work, the multimodal analysis had to remain broad. However, bearing in mind the cross-linguistic consistency of certain tendencies, such as the co-occurrence of facial expressions with affective stance-taking, new research perspectives arose where additional work still needs to be done.

One of the initial questions of the present work was why people quote their thought at all. What is the added value of framing something as quoting one's own thought instead of as speech or simply describing past events? In French and German, direct reported thought is a possibility for speakers to comment on prior events while maintaining the guise of the narrative. By doing so, they can distance their self in the there-and-then, displaying a stance, from their self in the here-and-now of the ongoing conversation, thus diminishing their responsibility for what they say. Like Oscar Wilde's "mask" (see *supra*), the direct reported thought format is a means of framing a stance as merely reported, or as *not necessarily emerging in real time*. This configuration changes the response relevance in the participant framework of the ongoing conversation as well as the speaker's accountability for their utterances because of the lower face-threatening potential of strong stances.

Therein also lies a crucial difference between direct reported thought and direct reported speech. Presenting an utterance as reported speech means presenting an utterance as something that has originally been a social action—meaning, it has already been part of an *inter*action in which a recipient has been involved. Thoughts are *per definitionem* always private and do not need to be socially acceptable. The reported thought is, in this sense, the display of an originally a-social action because no interaction, or no possibility for a recipient to respond to the thought, has taken place before. Consequently, framing an utterance as direct reported thought means framing it as something that has not yet been socially approved. At the same time, formatting an utterance as direct reported thought may also be a way of hedging a strong opinion or reasoning by situating it in a past framework. This also leaves speakers the possibility to modify

their stance retrospectively if it is treated as problematic by the interlocutor of the ongoing conversation.

Being first and foremost a means to display a stance, direct reported thought may also establish or enhance intersubjectivity. As a state of "shared understanding between social actors" (Raymond 2019: 182), intersubjectivity does not only involve shared knowledge, but also a "shared understanding of actions and sequences of action and shared understanding of the expression of emotion in sequences of action" (Sorjonen *et al.* 2021: 2). As through direct reported thought speakers make publicly available their inner state, the format contributes to the sharedness of otherwise private processes on an emotional or rational level. The format makes explicit a speaker's stance toward a prior telling or reasoning and invites recipients to reexperience or comprehend this stance as it emerged in the reported situation; Elements of depiction and staging enhance this effect of directness. The social action format of the quotative importantly contributes to the organization of intersubjectivity by routinely projecting reported thought. The recipients' recurrent affiliation and the speakers' orientation towards the preferredness of an affiliative response demonstrate that direct reported thought seems to regularly achieve intersubjectivity concerning affectiveness and rationality in sequences of action between interlocutors. Since the investigated quotatives seem to have routinized into social action formats projecting specific actions—either an affective or epistemic stance—recipients may identify early on the type of shared understanding that is needed.

The investigation of direct reported thought as a multimodal phenomenon has evidenced its multi-layered character. In some cases, at least two actions seem to be distributed between different resources. While speakers verbally display an epistemic stance, their affective stance is simultaneously conveyed through bodily resources. In line with Goodwin's (2000, 2013) notion of lamination, using several resources, speakers can accomplish several actions simultaneously, and recipients subsequently respond to one or several of these actions, thus "contribut[ing] in consequential ways to the organization of a single action" (Goodwin 2013: 13) or several actions, as the present study has shown. Goodwin specifically points out that,

> though most analysis of how action can be built and understanding is displayed within talk-in-interaction has focused on its sequential organization […] the simultaneous, concurrent organization of action is equally important. By providing different kinds of semiotic materials actors in structurally different positions (e.g., speaker and hearer, story teller and principal character, etc.) can contribute in consequential ways to the organization of a single action […]. (*ibid.*: 13)

He clarifies that recipients and speakers are equally involved in talk-in-interaction in that both contribute to action formation. The present work has, especially on a multimodal level, focused on the production side of resources; how a recipient re-uses, "re-laminates" (Goodwin 2013: 12), or responds to these resources has not been investigated in detail even though it offers important insights into processes of action ascription and formation. Future research may investigate this side of interaction.

One clear limit of the present work are the results of the quantifying analyses of multimodal conduct. As I have already pointed out, due to different types of data that I have in German compared to the consistent corpus in French, it is problematic to quantify the individual resources comparatively; Not only do the different types of interactional settings (cooking,

driving, drinking coffee) or the variability of participant frameworks change the interaction and participant orientation, static and non-static interactional situations also influence the perceivability of certain bodily actions. Gaze conduct in particular may change when participants are moving as well as according to the number of recipients a speaker has. Future research should investigate individual resources in more consistent interactional environments to tackle their exact use during a specific social action.

A possibility to gain a more detailed picture of the phenomenon of direct reported thought is a re-examination of the same quotatives under a Construction Grammar approach (Fillmore 1988; Fillmore *et al.* 1988; Kay 1997; Goldberg 1995; Croft 2001)[240]. Different interactional phenomena, such as discourse markers (Imo 2005, 2011a, 2012a, 2012b) and verb-first structures (Günthner 2006) in spoken German, have already been investigated under this notion. Several researchers have specifically argued for combining the Construction Grammar and Interactional approach to build a grammar of spoken language (Deppermann 2006; Günthner 2009b; Imo 2012a) that considers corpus-based results *and* sequential features. Such an investigation would allow for the analysis of a large data set from a quantitative perspective. Specifically for German, the Construction Grammar approach would also enable a more detailed study of co-occurrences of *denken* with other lexical items. Such results may also uncover recurrent patterns that allow interlocutors to differentiate between *denken* in present and past tense as a quotative and *denken* when used for other practices (Deppermann & Reineke 2017). Because the present work focused on the emergent and online character of grammatical patterns, it was not compatible with a Construction Grammar approach, which regards units as "determined through form and content as a basis of communication" (Deppermann 2006: 58, my translation) and not as interactional, co-constructed achievements. Future research with a stronger focus on grammatical patterning of larger data sets may add further aspects to the study of quotatives introducing direct reported thought, including regarding the possibility of multimodal constructions, even though the existence of multimodal constructions is subject to debate (Ningelgen & Auer 2017; Schoonjans 2017; Steen & Turner 2013; Ziem 2017).

New research perspectives arising from my results

On a grammatical, multimodal, and cross-linguistic level, the results of my analyses open new research perspectives for future studies. One interesting starting point for further investigation is the diachronic evolution of *ich denke (mir)* and *je pense* to find out how *ich denke (mir)* in present and past tense developed the function of a quotative whereas *je pense* did not.

As I have pointed out, a detailed investigation of past tense distribution in spoken French and German are still lacking, albeit for different reasons; whereas French research on past tense has mainly focused on the differences between the written, literary past tense *passé simple* and the *imparfait*, the actual use of *passé compose* and *imparfait* in spoken language may reveal a potential relation between action formation and tense distribution. In German, whereas the distribution of past tense has been mainly linked to the speakers' region of origin, a verb-specific investigation may demonstrate, as I have shown with *ich dachte (mir)* and *ich habe*

---

[240] The non-alphabetical order is due to the grouping of references according to the different conceptualizations of Construction Grammar. For a distinction between these conceptualizations, see Deppermann (2006).

(mir) gedacht, that there is more to the use of the preterite and present perfect than initially assumed. Future research should investigate whether the grammar-in-use of specific verbs in past tenses may not differ according to preterit and present perfect. Initial investigations point to distinct patterning according to past tense, at least with *verba sentiendi* like *meinen* ('to mean') and *finden* ('to find') (see Fiedler forthc.).

In a more applied sense, my cross-linguistic findings may be beneficial for translation studies as well as for German and French as a foreign language. In both domains, the systematic translation of *denken* with *penser* and vice versa may need to be adapted according to the context of use of the respective verb.

Further research perspectives have been mentioned throughout the presentation of my language-specific and cross-linguistic results. The numerous new research ideas that emerged from the present work evidenced that a cross-linguistic, multimodal study of an interactional phenomenon like direct reported thought, while sometimes remaining necessarily superficial, is a good starting point for raising further questions and ideas for future in-depth studies.

The impact of a multimodal analysis on the grammatical description of the direct-reported-thought format in French and German: Returning to the seven principles shaping grammar

The multimodal analysis of direct reported thought has brought to light new aspects concerning the grammar of clause-combining. Not only may Gast and Diessel's (2012) definition of clause-combining need to be extended in the sense that "some other constituent" (*ibid*.: 4) can also be bodily conduct, I have also demonstrated how complex the lamination of actions becomes when including several multimodal resources.

In addition to other projectional means, such as prosody or embodiment, grammar is a crucial resource that speakers exploit to grammatically project their quote with a quotative (with the exception of the two *être là* quotatives). To discuss the place of grammar in the analysis of direct reported thought in talk-in-interaction, I would like to return to Fox's (2007) seven principles shaping grammar, now reanalyzing them specifically for the bipartite structure of [quotative + direct reported thought].

(1) **Collocations.** Quotatives introducing direct reported thought indeed form collocation-like structures. By this I mean that, even though it is possible to formulate other's purported thoughts or mental states (see, for instance, Broth *et al.* 2019), speakers use quotatives to report thought primarily in first person singular, present, and past tense. The clearest example in my collections for the formation of collocations is *j'étais là*, which seems to have routinized in the imperfective tense. As I have shown, it does not occur in perfective tense and only rarely in present tense. In German, there is a tendency for quotatives with *denken* to co-occur in patterns with additional items (see Chapter 5). These collocations also seem to have routinized through repeated use to accomplish the same social action, thus exploiting grammar as a resource with praxeological motivations.

(2) **Occurrence in turns.** The bipartite structure of [quotative + direct reported thought] usually forms one turn (or part of a multi-unit turn). In some cases, the quotative can project a multi-unit turn in itself, primarily when it introduces an emergent reasoning. My analysis has shown that recipients do orient to the bipartite structure regardless of the exact format of the quote; Neither turn-shape nor the kind of resources deployed influence this orientation toward

the completeness of the bipartite structure. What seems to be relevant is that a clause (i.e., the quotative) is followed by another constituent—namely, the purported reported thought (or speech in the case of *je suis là / j'étais là*) that contains the main information. In this sense, recipients orient to the grammatical unit of the clause-combining format [quotative + direct reported thought] as one turn that must be closed before a recipient has the right to take the turn themselves.

(3) **Occurrence in sequences.** My analysis has shown that specific quotatives are used for specific social actions—namely, affective stance-taking or epistemic stance-taking to display a reasoning. These two kinds of actions project different recipient responses; whereas speakers usually affiliate after an affective stance-taking through direct reported thought, they more often display understanding after emergent reasonings. What my multimodal analysis has demonstrated is that this functional difference, leading to different recipient responses, can often be made based on the bodily resources that speakers use. Affective stance-taking is often displayed through facial expression or prosody. Even if speakers also verbally display their affective stance, bodily resources recurrently make the speaker's stance available to others early on, before it is verbally uttered. Recipients thereby must rely on both verbal information and the speakers' bodily displays to deliver a preferred and sequentially fitted response.

(4) **Unidirectionality.** Regardless of the resources that speakers deploy while reporting thought, all these resources "move toward completion" (Fox 2007: 306). However, when analyzing formats that recurrently use bodily resources while talking, the dimension of simultaneity may need to be discussed as well. My analysis clarified that speakers layer grammar, facial expressions, gaze, gesture, and posture, sometimes for multiple simultaneous actions (e.g., reporting purported thought verbally while taking a negative stance through a facial expression while mobilizing the recipient's response through gaze). Although these actions move forward in time, recipients also must be able to perceive them simultaneously and respond to them, if necessary, in time. This high degree of lamination of actions (Goodwin 2013) would not have been carved out without a multimodal analysis.

(5) **Interactional construction.** My analysis demonstrated that speakers and recipients use grammatical projection to orient to the bipartite structure of direct reported thought. When the quotative consists of a complement-taking predicate, the principles of constituency and dependence are still relevant for the emergence of the clause-combining format. Excerpts that contain shifts in footing by the recipients to co-animate a telling especially evidence that direct reported thought also relies on interactional projection based on shared knowledge of interactional structures. That direct reported thought in the closing environment of a telling systematically projects affiliation or a display of understanding depending on the projected social action further demonstrates that interlocutors establish intersubjectivity. Excerpts where responses are mobilized through gaze after a noticeable absence of a response evidence that interlocutors work towards this shared understanding—a result that demonstrates, once again, the importance of analyzing multiple semiotic layers of an action.

(6) **Grammar as a public embodiment of action.** It has been specifically relevant for the direct reported thought format that bodily displays are often integral part of the action that speakers accomplish with it. When the social action format projects an affective stance, bodily resources often do the most explicit stance-taking compared to verbal resources. This means

that they are sometimes not simply dramatizing elements that reinforce a verbal message rather they are the main resource conveying the stance. As part of the clause-combining format, the projected bodily conduct may therefore indeed be a "public embodiment of action" that does not simply accompany a verbal utterance but also forms a constituent on its own.

(7) **Frequency.** Resorting to the notion of routinization and projector-construction to describe the recurrent use of the same linguistic formats projecting direct reported thought (or speech in the case of *je suis là / j'étais là*), I did not necessarily foreground frequency *per se* as usage-based analyses have (Bybee 2006, 2010; Bybee & Hopper 2001). Of greater importance for the present work is the result that one format is recurrently used to project the same social action—namely, quoting one's own thought for affective or epistemic stance-taking. Even though my collections of each quotative do not contain a large number of cases, I was able to show their systematic use in similar sequential environments and for introducing comparable actions, also cross-linguistically. Some of the quotatives—most strikingly *j'étais là*—show signs of recurrent use, such as through morpho-phonological reduction and the apparent loss of its aspectual dimension. This tendency is less clear in German and for *je me dis* and *je me suis dit*. Because of the small number of cases, no clear results concerning recurrence have been obtained regarding the speakers' bodily conduct; the only observation possible was that, when quotatives are used for affective stance-taking, they more often project facial expressions when used to display an emergent reasoning.

The cross-linguistic, multimodal investigation of direct reported thought in everyday talk evidenced the dynamic nature of grammar and the body *in use*. Because speakers adapt their talk—and, as the present work has shown, their bodily conduct—constantly to the changing contingencies and exigencies of talk-in-interaction, the analysis of grammatical phenomena like direct reporting formats confronts the analyst with a variety of cases. However, because grammar is a resource for interaction, interaction can also be a driving force for grammar to form and adapt to interlocutors' needs. Certain grammatical formats, like the quotative as a social action format, develop out of interactional motivations and frequent use.

An emic approach to the seemingly chaotic "organization" of everyday talk allows for the detection of these routinizations. The investigation of the use of direct reported thought in mundane French and German conversation uncovered yet another piece of the puzzle of the actual orderliness of human sociality, which everyday talk-in-interaction reflects.

9. Bibliography

Abeillé, A. & Godard, D. (Eds.) (2022). *La Grande Grammaire du français*. Paris: Actes Sud.

Abouda, L. & Rendulić, N. (2017). Séquence d'introduction de discours représenté : *faire* ou *dire* ? *Discours [online]* 21.

Acuña Ferreira, V. (2021). Staging Mental Discursive Processes and Reactions: The Construction of Direct Reported Thought (DRT) in Conversational Storytelling. *Language in Society* 50(2), 235-257.

Adolf, H. (1944). Intonation and Word Order in German Narrative Style. *The Journal of English and Germanic Philology* 43(1), 71-79.

Aijmer, K. (2011). *I think* - an English Modal Particle. In Swan, T. & Westvik, O. J. (Eds.), *Modality in Germanic Languages: Historical and Comparative Perspectives*. Berlin/New York: de Gruyter, 1-48.

Anscombre, J.-C. (2010). Lexique et médiativité: Les marqueurs pour le dire. *Cahier Lexicol* 96, 5-33.

Antaki, C. (1994). Explaining and Arguing: The Social Organization of Accounts. London: Sage.

Auer, P. (1993). Zur Verbspitzenstellung im Deutschen. *Deutsche Sprache* 21, 193-222.

— (1996). On the Prosody and Syntax of Turn-Continuations. In Couper-Kuhlen, E. & Selting, M. (Eds.), *Prosody in Conversation*. Cambridge: Cambridge University Press, 57-100.

— (1998). Zwischen Parataxe und Hypotaxe: 'abhängige Hauptsätze' im gesprochenen und geschriebenen Deutsch. *InLiSt* 2, 1-32.

— (2002). Projection in Interaction and Projection in Grammar. *Text* 25(1), 7-36.

— (2005). Syntax als Prozess. *InLiSt* 41, 1-35.

— (2006). Construction Grammar meets Conversation: Einige Überlegungen am Beispiel von 'so'-Konstruktionen. In Günthner, S. & Imo, W. (Eds.), *Konstruktionen in der Interaktion*. Berlin/New York: de Gruyter, 290-314.

— (2007). Why Are Increments Such Elusive Objects? An Afterthought. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)* 17(4), 647-658.

Auer, P. & Maschler, Y. (Eds.) (2016). *NU/NÅ – A Family of Discourse Markers Across the Languages of Europe and Beyond*. Berlin/Boston: de Gruyter.

Auer, P. & Stukenbrock, A. (2018). When 'You' Means 'I': The German 2nd Ps.Sg. Pronoun *du* between Genericity and Subjectivity. *Open Linguistics* 4(1), 280-309.

Auer, P. (2021). Turn-Allocation and Gaze: A Multimodal Revision of the "Current-Speaker-Selects-Next" Rule of the Turn-Taking System of Conversation Analysis. *Discourse Studies* 23(2), 117-140.

Authier, J. & Meunier, A. (1977). Exercices de grammaire et discours rapporté. *Langue française* 33, 41-67.

Authier-Revuz, J. (1978). Les formes du discours rapporté. Remarques syntaxiques et sémantiques à partir des traitements proposés. *DRLAV. Documentation et Recherche en Linguistique Allemande Vincennes* 17(1), 1-87.

— (1979). Problèmes posés par le traitement du discours rapporté dans une grammaire de phrase. *Lingvisticæ Investigationes* 3(2), 211-228.

— (1984). Hétérogénéité(s) énonciative(s). *Langages* 73, 98-111.

— (1987). L'auto-représentation opacifiante du dire dans certaines formes de "couplage". *DRLAV. Documentation et Recherche en Linguistique Allemande Vincennes* 36(1), 55-103.

— (1992). Repères dans le champ du discours rapporté. *L'information grammaticale* 55(1), 38-42.

— (1993). Repères dans le champ du discours rapporté (suite). *L'information grammaticale* 56(1), 10-15.

— (2019). *La représentation du discours autre: principes pour une description*. Berlin/Boston: de Gruyter.

Avanzi, M. (2009). La prosodie des verbes parenthétiques en français parlé. *Linx* 61, 131-144.

Bakhtin, M. M. (1981). Discourse in the Novel. In Holquist, M. (Ed.), *The Dialogic Imagination*. Austin: University of Texas Press, 259-422.

Bamberg, M. (1997). Positioning Between Strucutre and Performance. *Journal of Narrative and Life History* 7(1-4), 335-342.

Banfield, A. (1982). *Unspeakable Sentences. Narration and Representation in the Language of Fiction*. Boston, MA: Routledge & Kegan Paul.

Bangerter, A., Mayor, E. & Pekarek Doehler, S. (2011). Reported Speech in Conversational Storytelling During Nursing Shift Handover Meetings. *Discourse Processes* 48(3), 183-214.

Barnes, R. & Moss, D. (2007). Communicating a Feeling: The Social Organization of 'Private Thoughts'. *Discourse Studies* 9(2), 123-148.

Barth-Weingarten, D. (2009). Contrasting and Turn Transition: Prosodic Projection with the Parallel-Opposition Construction. *Journal of Pragmatics* 41(11), 2271-2294.

Baudot, D. (Ed.) (2002). *Redewiedergabe, Redeerwähnung. Formen und Funktionen des Zitierens und Reformulierens im Text*. Tübingen: Stauffenburg.

Bavelas, J., Gerwing, J. & Healing, S. (2014). Effect of Dialogue on Demonstrations: Direct Quotations, Facial Portrayals, Hand Gestures, and Figurative References. *Discourse Processes* 51(8), 619-655.

Berger, E. & Pekarek Doehler, S. (2015). Direct Reported Speech in Storytellings: Enacting and Negotiating Epistemic Entitlements. *Text & Talk* 35(6), 789-813.

Bergmann, J. R. (1987). *Klatsch. Zur Sozialform der diskreten Indiskretion*. Berlin/New York: de Gruyter.

— (1993). *Discreet Indiscretions: The Social Organization of Gossip*. Berlin/New York: de Gruyter.

Betten, A. (1985). Formen fragmentarischer Gesprächsäußerungen in simulierter gesprochener Sprache. Versuch einer stilistischen Unterscheidung. In Meyer-Hermann, R. & Rieser, H. (Eds.), *Ellipsen und fragmentarische Ausdrücke. Band 2*. Tübingen: Niemeyer, 269-294.

Betz, E., Deppermann, A., Mondada, L. & Sorjonen, M.-L. (Eds.) (2021). *OKAY across Languages: Toward a Comparative Approach to its Use in Talk-in-Interaction*. Amsterdam/Philadelphia: Benjamins.

Bickel, B. (1991). Typologische Grundlagen der Satzverkettung: ein Beitrag zur allgemeinen Grammatik der Satzverbindung und des Fährtenlegens. *Zürich: Arbeiten des Seminars für Allgemeine Sprachwissenschaft der Universität Zürich* 9.

Blackwell, N. L., Perlman, M. & Fox Tree, J. E. (2015). Quotation as a Multimodal Construction. *Journal of Pragmatics* 81, 1-7.

Blanche-Benveniste, C. (1988). Constructions verbales "en incise" et rection faible des verbes. *Recherches sur le français parlée* 9, 53-73.

Blanche-Benveniste, C. & Willems, D. (2007). Un nouveau regard sur les verbes « faibles ». *Bulletin de la Société de linguistique de Paris* CII(I), 217-254.

Blanche-Benveniste, C. (2010). Les pseudo-clivées et l'effet deux points. In Béguelin, M.-J., Avanzi, M. & Corminbœuf, G. (Eds.), *La Parataxe. Structures, marquages et exploitations discursives. Tome 2*. Bern: Peter Lang, 185-217.

Blondeau, H. & Moreno, A. (2018). "On a fait comme 'c'est fou là'" ou l'émergence de *comme* et sa concurrence avec *genre* et d'autres formes d'introduction de discours direct dans le français de Montréal. In Barthelmebs-Raguin, H., Komur-Thilloy, G., Lopez-Muñoz, J. M., Marnette, S. & Rosier, L. (Eds.), *Le Discours rapporté. Temporalité, histoire, mémoire et patrimoine discursif*. Paris: Garnier, 41-58.

Blum-Kulka, S., House, J. & Kasper, G. (1989). *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood: Ablex Publishing Corporation.

Bohle, U. (2014). Body Posture and Movement in Interaction: Participation Management. In Müller, C., Cienki, A., Fricke, E., Ladewig, S., McNeill, D. & Bressem, J. (Eds.), *Body – Language – Communication*. Berlin/Boston: de Gruyter.

Brendel, E., Meibauer, J. & Steinbach, M. (2007). Aspekte einer Theorie des Zitierens. In Brendel, E., Meibauer, J. & Steinbach, M. (Eds.), *Zitat und Bedeutung*. Hamburg: Buske, 5-25.

Bressem, J. & Müller, C. (2014a). The Family of Away Gestures. In Müller, C., Cienki, A., Fricke, E., Ladewig, S., McNeill, D. & Bressem, J. (Eds.), *Body – Language – Communication*. Berlin/Boston: de Gruyter, 1592-1604.

— (2014b). A Repertoire of German Gestures. In Müller, C., Cienki, Alan, Fricke, Ellen, Ladewig, Silvia, McNeill, David, Bressem, Jana (Ed.), *Body – Language – Communication* Berlin/Boston: de Gruyter, 1575-1591.

Brøndal, V. (1943). Le problème de l'hypotaxe. Réflexions sur la théorie des propositions [1937]. In Brøndal, V. (Ed.), *Essaies de linguistique générale*. Copenhagen: Munksgaard, 72-80.

Broth, M., Cromda, J. & Levin, L. (2019). Telling the Other's Side: Formulating Other's Mental States in Driver Training. *Language & Communication* 65, 7-21.

Brunner, A. (2015). *Automatische Erkennung von Redewiedergabe in literarischen Texten. Ein Beitrag zur quantitativen Narratologie*. Berlin/Boston: de Gruyter.

Brünner, G. (1991). Redewiedergabe in Gesprächen. *Deutsche Sprache* 1, 1-15.

Buchstaller, I. & Van Alphen, I. (2012). Preface: Introductory Remarks on New and Old Quotatives. In Buchstaller, I. & Van Alphen, I. (Eds.), *Quotatives. Cross-Linguistic and Cross-Disciplinary Perspectives*. Amsterdam/Philadelphia: Benjamins, XI-XXX.

Buchstaller, I. (2014). *Quotatives: New Trends and Sociolinguistic Implications*. Malden/Oxford: Wiley.

Bühler, K. (1978 [1934]). *Sprachtheorie. Die Darstellungsfunktion der Sprache*. Frankfurt/Berlin/Wien: Ullstein.

Butters, R. R. (1980). Narrative Go 'Say'. *American Speech* 55(4), 304-307.

Butterworth, J. (2015). *Redewiedergabeverfahren in der Interaktion. Individuelle Variation bei der Verwendung einer kommunikativen Ressource*. Heidelberg: Universitätsverlag Winter.

Bybee, J. & Hopper, P. J. (2001). Introduction to Frequency and the Emergence of Linguistic Structure. In Bybee, J. L. & Hopper, P. J. (Eds.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam/Philadelphia: Benjamins, 1-26.

Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4), 711-733.

— (2010). *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Cantarutti, M. N. (2020). *The Multimodal and Sequential Design of Co-Animation as a Practice for Association in Interaction*. University of York.

Casartelli, D. E., Cruschina, S., Posio, P. & Spronck, S. (Eds.) (2023). *The Grammar of Thinking: From Reported Speech to Reported Thought in the Languages of the World*. Berlin/Boston: de Gruyter.

Casparis, C. P. (1975). *Tense Without Time. The Present Tense in Narration*. Bern: Francke.

Chafe, W. L. (1993). Prosodic and Functional Units of Language. In Edwards, J. A. & Lampert, M. D. (Eds.), *Talking Data. Transcription and Coding in Discourse Research*. Hillsdale: Lawrence Erlbaum, 33-43.

Cheshire, J., Kerswill, P., Fox, S. & Torgersen, E. (2011). Contact, the Feature Pool and the Speech Community: The Emergence of Multicultural London English. *Journal of Sociolinguistics* 15, 151-198.

Cheshire, J. & Secova, M. (2018). The Origins of New Quotative Expressions: The Case of Paris French. *Journal of French Language Studies* 28(2), 209-234.

Clark, H. H. & Gerrig, R. J. (1990). Quotations as Demonstrations. *Language* 66(4), 764-805.

Clark, H. H. (2016). Depicting as a Method of Communication. *Psychological Review* 123(3), 324-347.

Clayman, S. E. (2013). Turn-Constructional Units and the Transition-Relevance Place. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Malden: Blackwell, 150-167.

Clift, R. & Holt, E. (Eds.) (2007). *Reporting Talk. Reported Speech in Interaction*. Cambridge: Cambridge University Press.

Clift, R. (2016). Don't Make Me Laugh: Responsive Laughter in (Dis)affiliation. *Journal of Pragmatics* 100, 73-88.

Coulmas, F. (Ed.) (1986). *Direct and Indirect Speech*. Berlin/New York: de Gruyter.

Couper-Kuhlen, E. (1996). The Prosody of Repetition: On Quoting and Mimicry. In Couper-Kuhlen, E. & Selting, M. (Eds.), *Prosody in Conversation*. Cambridge: Cambridge University Press, 366-405.

Couper-Kuhlen, E. & Selting, M. (1996). Introduction. In Couper-Kuhlen, E. & Selting, M. (Eds.), *Prosody in Conversation*. Cambridge: Cambridge University Press, 1-10.

Couper-Kuhlen, E. (1999). Coherent Voicing: On Prosody in Conversational Reported Speech. *InLiSt* 1, 1-28.

Couper-Kuhlen, E. & Thompson, S. A. (2000). Concessive Patterns in Conversation. In Couper-Kuhlen, E. & Kortmann, B. (Eds.), *Cause, Condition, Contrast. Cognitive and Discursive Perspectives*. Berlin/New York: de Gruyter, 381-409.

Couper-Kuhlen, E. (2004). Analyzing Language in Interaction: The Practice of *Never Mind*. *English Language and Linguistics* 8(2), 207-237.

Couper-Kuhlen, E. & Thompson, S. A. (2005). A Linguistic Practice for Retracting Overstatements: 'Concessive Repair'. In Hakulinen, A. & Selting, M. (Eds.), *Syntax and Lexis in Conversation: Studies on the use of linguistic resources in talk-in-interaction*. Amsterdam/Philadelphia: Benjamins, 257-288.

Couper-Kuhlen, E. (2007). Assessing and Accounting. In Holt, E. & Clift, R. (Eds.), *Reporting Talk. Reported Speech in Interaction*. Cambridge: Cambridge University Press, 81-119.

Couper-Kuhlen, E. & Thompson, S. A. (2008). On Assessing Situations and Events in Conversation: 'Extraposition' and its Relatives. *Discourse Studies* 10(4), 443-467.

Couper-Kuhlen, E. (2009). A Sequential Approach to Affect: The Case of 'Disappointment'. In Haakana, M., Laasko, M. & Lindström, J. (Eds.), *Talk in Interaction: Comparative Dimensions*. Helsinki: Finnish Literature Society, 94-123.

— (2011). Affectivity in Cross-Linguistic and Cross-Cultural Perspective. *Mobilisierte Kulturen* 2, 231-257.

— (2018). Finding a Place for Body Movement in Grammar. *Research on Language and Social Interaction* 51(1), 22-25.

Couper-Kuhlen, E. & Selting, M. (2018). *Interactional Linguistics*. Cambridge: Cambridge University Press.

Couper-Kuhlen, E. (2021). Language Over Time: Some Old and New Uses of OKAY in American English. *Interactional Linguistics* 1(1), 33-63.

Croft, W. (2001). *Radical Construction Grammar*. Oxford: Oxford University Press.

D'Hertefelt, S. & Verstraete, J.-C. (2014). Independent Complement Constructions in Swedish and Danish: Insubordination or Cependency Shift? *Journal of Pragmatics* 60, 89-102.

De Brabanter, P. (2018). Pragmatic and Semantic Commitment when Using Quotative Markers, with Application to French *dire* and *genre*. *Journal of Pragmatics* 128, 137-147.

De Cornulier, B. (1978). L'incise, la classe des verbes parenthétiques et le signe mimique. *Cahier de linguistique* 8, 53-95.

De Ruiter, J. P. (2007). *Some Multimodal Signals in Humans,* Paper presented at the CTIT Proceedings of the Workshop on Multimodal Output Generation University of Aberdeen.

De Stefani, E. (2013). The Collaborative Organisation of Next Actions in a Semiotically Rich Environment: Shopping as a Couple. In Haddington, P., Mondada, L. & Nevile, M. (Eds.), *Interaction and Mobility: Language and the Body in Motion*. Berlin/Boston: de Gruyter, 123-151.

De Stefani, E. & Mondada, L. (2013). Reorganizing Mobile Formations: When 'Guided' Participants Initiate Reorientations in Guided Tours. *Space and Culture* 17(2), 157-175.

De Stefani, E. (2021). *If*-Clauses, Their Grammatical Consequents, and Their Embodied Consequence: Organizing Joint Attention in Guided Tours. *Frontiers in Communication* 6, 1-18.

De Stefani, E. & Mondada, L. (2021). Resources for Action Transition: OKAY and its Embodied and Material Habitat in Institutional Settings. In Deppermann, A., Betz, E., Mondada, L. & Sorjonen, M.-L. (Eds.), *OKAY Across Languages. Toward a Comparative Approach to Its Use in Talk-in-Interaction*. Amsterdam/Philadelphia: Benjamins, 295-330.

Debras, C. (2017). The Shrug: Forms and Meanings of a Compound Enactment. *Gesture* 16(1), 1-34.

Demers, M. (1998). Prosodie et discours. Le cas du discours rapporté en français québécois oral spontané. *Revue québécoise de linguistique* 26(1), 27-50.

Deppermann, A. (2006). *Construction Grammar* – Eine Grammatik für die Interaktion? In Deppermann, A., Fiehler, R. & Spranz-Fogasy, T. (Eds.), *Grammatik und Interaktion*. Radolfzell: Verlag für Gesprächsforschung, 43-65.

Deppermann, A. & Helmer, H. (2013). Zur Grammatik des Verstehens im Gespräch: Inferenzen anzeigen und Handlungskonsequenzen ziehen mit *also* und *dann*. *Zeitschrift für Sprachwissenschaft* 32(1), 1-39.

Deppermann, A. (2015). Positioning. In De Fina, A. & Georgakopoulou, A. (Eds.), *The Handbook of Narrative Analysis*. New York: Wiley, 369-387.

Deppermann, A. & Reineke, S. (2017). Epistemische Praktiken und ihre feinen Unterschiede: Verwendungen von *ich dachte* in gesprochener Sprache. In Deppermann, A., Proske, N. & Zeschel, A. (Eds.), *Verben im interaktiven Kontext. Bewegungsverben und mentale Verben im gesprochenen Deutsch*. Tübingen: Narr Franke Attempo, 337-375.

— (2020). Practices of Indexing Discrepant Assumptions with German *ich dachte* ('I thought') in Talk-in-Interaction. *Functions of Language* 27(2), 113-142.

Diewald, G. & Fischer, K. (1998). Zur diskursiven und modalen Funktion der Partikeln *aber*, *auch*, *doch* und *ja* in Instruktionsdialogen. *Linguistica* 38(1), 75-99.

Dingemanse, M., Torreira, F. & Enfield, N. J. (2013). Is "Huh?" a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items. *PloS one* 8(11), e78273.

Dingemanse, M. & Enfield, N. J. (2015). Other-Initiated Repair Across Languages: Towards a Typology of Conversational Structures. *Open Linguistics* 1(1), 96-118.

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., Gisladottir, R. S., Kendrick, K. H., Levinson, S. C. & Manrique, E. (2015). Universal Principles in the Repair of Communication Problems. *PloS one* 10(9), e0136100.

Discherl, F. & Pafel, J. (2015). Die vier Arten der Rede- und Gedankendarstellung. Zwischen Zitieren und Referieren. *Linguistische Berichte* 241, 3-47.

Dostie, G. (2020). Deixis à l'imaginaire et périphrases en *être* et *faire* préfaçant une scène recréée (*j'étais comme, t'es là, j'ai fait…*). In Diémoz, F., Dostie, G., Hadermann, P. & Lefeuvre, F. (Eds.), *Le français innovant*. Bern: Peter Lang, 83-106.

Drew, P. (1998). Complaints About Transgressions and Misconduct. *Research on Language & Social Interaction* 31(3-4), 295-325.

Drew, P. & Walker, T. (2009). Going Too Far: Complaining, Escalating and Disaffiliation. *Journal of Pragmatics* 41(12), 2400-2414.

Du Bois (2007). The Stance Triangle. In Englebretson, R. (Ed.), *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*. Amsterdam/Philadelphia: Benjamins.

Egbert, M. M. (1997). Some Interactional Achievements of Other-Initiated Repair in Multiperson Conversation. *Journal of Pragmatics* 27(5), 611-634.

Ehlich, K. (1986). *So*: Überlegungen zum Verhältnis sprachlicher Formen und sprachlichen Handelns, allgemein und an einem widerspenstigen Beispiel. *Ludar Germanistische Forschungen* 55, 279-298.

Ehmer, O. (2011). *Imagination und Animation: Die Herstellung mentaler Räume durch animierte Rede*. Berlin/New York: de Gruyter.

Eisenberg, P. (1994). *Grundriß der deutschen Grammatik*. Stuttgart/Weimar: Metzler.

Ekman, P. (1979). About Brows: Emotional and Conversational Signals. In Aschoff, J., von Carnach, M., Foppa, K., Lepenies, W. & Plog, D. (Eds.), *Human Ethology*. Cambridge: Cambridge University Press, 169-202.

Ekman, P., Friesen, W. V. & Hager, J. C. (2002). *Facial Action Coding System*. Salt Lake City, UT: Research Nexus.

Enfield, N. J. & Sidnell, J. (2017). On the Concept of Action in the Study of Interaction. *Discourse Studies* 19(5), 515-535.

Englebretson, R. (Ed.) (2007). *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction*. Amsterdam/Philadelphia: Benjamins.

Fabricius-Hansen, C. & Ramm, W. (2008). Editor's Introduction: Subordination and Coordination from Different Perspectives, *'Subordination' versus 'Coordination' in Sentence and Text*, 1-30.

Fabricius-Hansen, C. (2019). Redewiedergabe Reloaded? Speech and Thought Representation Reloaded? *Zeitschrift für germanistische Linguistik* 47, 1-27.

Fauré, L. (1997). Les interjections à l'oral : quelles valeurs pour les vocalisations? *Cahiers de praxématique* 28, 127-148.

Fauré, L. & Vérine, B. (2004). Authentifier un discours autre en y mettant du sien: les vocalisations *ah* et *oh* en frontière de discours rapporté direct à l'oral. In Lopez Muñoz, J. M., Marnette, S. & Rosier, L. (Eds.), *Le Discours rapporté dans tous ses états*. Paris: L'Harmattan, 317-327.

Fiedler, S. (2023). Thinking Out Loud? *Je me suis dit* 'I said to myself' and *j'étais là* 'I was there' in French Talk-in-Interaction. In Casartelli, D. E., Cruschina, S., Posio, P. & Spronck, S. (Eds.), *The Grammar of Thinking: From Reported Speech to Reported Thought in the Languages of the World*. Berlin/Boston: de Gruyter, 141-170.

— (forthc.). The Use of Past Tense Formats in German Talk-in-interaction. In Steensig, J., Jørgensen, M., Lindström, J., Mikkelsen, N., Suomalainen, K. & Sørensen Sandager, S. (Eds.), *Grammar in Action*. Amsterdam/Philadelphia: Benjamins.

Field, M. (2007). Increments in Navajo Conversation. *Pragmatics* 17(4), 637-646.

Filatkina, N., Stumpf, S. & Pfeiffer, C. (2020). Introduction: Formulaic Language and New Data. In Piirainen, E., Filatkina, N., Stumpf, S. & Pfeiffer, C. (Eds.), *Formulaic Language and New Data: Theoretical and Methodological Implications*. Berlin/Boston: de Gruyter, 1-16.

Fillmore, C. J. (1988). Grammatical Construction Theory and the Familiar Dichotomies. In Dietrich, R. & Graumann, C. F. (Eds.), *Language Processing in Social Context*. Amsterdam: Elsevier, 17-38.

Fillmore, C. J., Kay, P. & O'Conner, M. C. (1988). Regularity and Idiomaticity in Grammatical Constructions: The Case of *let alone*. *Language* 64(3), 501-538.

Fischer, H. (2018). Präteritumsschwund im Deutschen. Dokumentation und Erklärung eines Verdrängungsprozesses. Berlin/Boston: de Gruyter.

Fleischman, S. (1986). Evaluation in Narrative: The Present Tense in Medieval 'Performed Stories'. *Yale French Studies* 70, 199-251.

Fludernik, M. (2002). Tempus und Zeitbewusstsein. Erzähltheoretische Überlegungen zur englischen Literatur. In Middeke, M. (Ed.), *Zeit und Roman: Zeiterfahrung im historischen Wandel und ästhetischer Paradigmenwechsel vom sechzehnten Jahrhundert bis zur Postmoderne*. Würzburg: Königshausen & Neumann, 21-32.

Foley, W. A. & Van Valin Jr., R. D. (1984). *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Ford, C. E. (1993). *Grammar in Interaction*. Cambridge: Cambridge University Press.

Ford, C. E., Fox, B. & Thompson, S. A. (1996). Practices in the Construction of Turns: The 'TCU' Revisited. *Pragmatics* 6(3), 427-454.

Ford, C. E. & Fox, B. A. (2002). Constituency and the Grammar of Turn Increments. In Ford, C. E., Fox, B. A. & Thompson, S. A. (Eds.), *The Language of Turn and Sequence*. Oxford: Oxford University Press, 14-38.

Fox, B. A. (2007). Principles Shaping Grammatical Practices: An Exploration. *Discourse Studies* 9(3), 299-318.

Fox, B. A., Wouk, F., Hayashi, M., Fincke, S., Tao, L., Sorjonen, M.-L., Laasko, M. & Flores Hernandez, W. (2009). Cross-Linguistic Investigation of the Site of Initiation in Same-Turn Self-Repair. In Sidnell, J. (Ed.), *Conversation Analysis: Comparative Perspectives*. Cambridge: Cambridge University Press, 59-103.

Fox, B. A. & Robles, J. (2010). It's Like *mmm*: Enactments with *It's Like*. *Discourse Studies* 12(6), 715-738.

Fox, B. A. & Heinemann, T. (2016). Rethinking Format: An Examination of Requests. *Language in Society* 45(4), 499-531.

— (2017). Issues in Action Formation: Requests and the Problem with x. *Open Linguistics* 3(1).

Fox Tree, J. E. & Tomlinson, J. M. (2008). The Rise of *Like* in Spontaneous Quotations. *Discourse Processes* 45(1), 85-102.

Fries, N. (1988). Über das Null-Topik im Deutschen. *Forschungsprogramm Sprache und Pragmatik: Arbeitsbericht Nr. 3 (Lund)*, 19-49.

Fujii, S. (2006). Quoted Thought and Speech Using the *mitai-na* 'be like' Noun-Modifying Construction. In Suzuki, S. (Ed.), *Emotive Communication in Japanese*. Amsterdam/Philadelphia: Benjamins, 53-95.

Gachet, F. (2012). *Incises de discours rapporté et autres verbes parenthétiques: une étude grammaticale*. PhD Thesis. Université de Fribourg (Switzerland).

Garfinkel, H. (1967). *Studies in Ethnomethodology*. Cambridge: Polity Press.

Gast, V. & Diessel, H. (2012). The Typology of Clause Linkage: Status Quo, Challenges, Prospects. In Gast, V. & Diessel, H. (Eds.), *Clause Linkage in Cross-Linguistic Perspective: Data-Driven Approaches to Cross-Clausal Syntax*. Berlin/Boston: de Gruyter, 1-36.

Givón, T. (1980). The Binding Hierarchy and the Typology of Complements. *Studies in Language* 4, 333-377.

Glikman, J. (2013). L'absence de subordonnant en français. In Fesenmeier, L., Grutschus, A. & Patzelt, C. (Eds.), *L'absence au niveau syntagmatique*. Frankfurt a. Main: Vittorio Klostermann, 181-199.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. London: Penguin Books.

— (1974). *Frame Analysis: An Essay on the Organization of Experience*. Cambridge, Massachusetts: Harvard University Press.

— (1979). Footing. *Semiotica* 25(1/2), 1-29.

— (1981). *Forms of Talk*. Philadelphia: University of Pennsylvania Press.

Gohl, C. (2006). *Begründen im Gespräch. Eine Untersuchung sprachlicher Praktiken zur Realisierung von Beründungen im gesprochenen Deutsch*. Tübingen: Niemeyer.

Golato, A. (2000). An Innovative German Quotative for Reporting on Embodied Actions: *Und ich so/und er so* 'and I'm like/and he's like'. *Journal of Pragmatics* 32, 29-54.

— (2002a). Grammar and Interaction: Reported Discourse and Subjunctive in German. *Zeitschrift für Sprachwissenschaft* 21(1), 25-56.

— (2002b). Self-Quotation in German. Reporting on Past Decisions. In Güldemann, T. & von Roncador, M. (Eds.), *Reported Discourse: A Meeting Ground for Different Linguistic Domains*. Amsterdam/Philadelphia: Benjamins, 49-70.

— (2012). Impersonal Quotation and Hypothetical Discourse. In Buchstaller, I. & Van Alphen, I. (Eds.), *Quotatives. Cross-Linguistic and Cross-Disciplinary Perspectives*. Amsterdam/Philadelphia: Benjamins, 3-36.

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Gómez-Jordana Ferary, S. & Anscombre, J.-C. (2015). Introduction: *Dire* et ses marqueurs. *Langue française* 2(186), 5-12.

Goodwin, C. (1979). The Interactive Construction of a Sentence in Natural Conversation. In Psathas, G. (Ed.), *Everyday Language: Studies in Ethnomethodology*. New York: Irvington Academic Press, 101-121.

— (1981). *Conversational Organization. Interaction between Speakers and Hearers*. New York: Academic Press.

— (1984). Notes on Story Structure and the Organization of Participation. In Atkinson, J. M. & Heritage, J. (Eds.), *Structures of Social Action*. Cambridge: Cambridge University Press, 225-246.

— (1986). Between and Within: Alternative Sequential Treatments of Continuers and Assessments. *Human Studies* 9(2/3), 205-217.

— (1992). Assessments and the Construction of Context. In Duranti, A. & Goodwin, C. (Eds.), *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press, 147-189.

— (2000). Action and Embodiment within Situated Human Interaction. *Journal of Pragmatics* 32, 1489-1522.

— (2013). The Co-operative, Transformative Organization of Human Action and Knowledge. *Journal of Pragmatics* 46(1), 8-23.

— (2015). Narrative as Talk-in-Interaction. In De Fina, A. & Georgakopoulou, A. (Eds.), *The Handbook of Narrative Analysis*. New York: Wiley, 197.

Goodwin, C. & Goodwin, M. H. (1987). Concurrent Operations on Talk: Notes on the Interactive Organization of Assessments. *IPrA Papers in Pragmatics* 1(1), 1-54.

Goodwin, M. H. (1980). Directive-Response Speech Sequences in Girls' and Boys' Task Activities. In McConnell-Ginet, S., Borker, R. & Furman, N. (Eds.), *Women and Language in Literature and Society*. New York: Praeger, 157-173.

Goodwin, M. H. & Goodwin, C. (1986). Gesture and Coparticipation in the Activity of Searching for a Word. *Semiotica* 62(1/2), 51-75.

Goodwin, M. H. & Alim, H. S. (2010). "Whatever (Neck Roll, Eye Roll, Teeth Suck)": The Situated Coproduction of Social Categories and Identities Through Stancetaking and Transmodal Stylization. *Journal of Linguistic Anthropology* 20(1), 179-194.

Grosjean, M. & Lacoste, M. (1999). *Communication et intelligence collective. Le travail à l'hôpital*. Paris: Presses Universitaires de France.

Guardiola, M. & Bertrand, R. (2013). Interactional Convergence in Conversational Storytelling: When Reported Speech Is a Cue of Alignment and/or Affiliation. *Frontiers in Psychology* 4(705), 1-17.

Guérin, E. & Moreno, A. (2015). Présence/absence de particules d'amorce et de particules d'extension dans le discours rapporté : peut-on parler de variation? *Langage et société* 4(154), 67-82.

Güldemann, T., von Roncador, M. & van der Wurff, W. (2002). A Comprehensive Bibliography of Reported Discourse. In Güldemann, T. & von Roncador, M. (Eds.), *Reported Discourse: A Meeting Ground for Different Linguistic Domains*. Amsterdam/Philadelphia: Benjamins, 363-415.

Gumperz, J. J. (1977). Sociocultural Knowledge in Conversational Inference. In Saville-Troike, M. (Ed.), *Linguistics and Anthropology*. Washington: Georgetown University Press, 191-212.

Günthner, S. (1997). The Contextualization of Affect in Reported Dialogues. In Niemeier, S. & Dirven, R. (Eds.), *The Language of Emotions*. Amsterdam/Philadelphia: Benjamins, 247-276.

— (1999a). Polyphony and the 'Layering of Voices' in Reported Dialogues: An Analysis of the Use of Prosodic Devices in Everyday Reported Speech. *Journal of Pragmatics* 31(5), 685-708.

— (1999b). Direkte und indirekte Rede in Alltagsgesprächen. Zur Interaktion von Syntax und Prosodie in der Redewiedergabe. In Schlobinski, P. (Ed.), *Syntax des gesprochenen Deutsch*. Opladen: Westdeutscher Verlag, 227-262.

— (2000). *Vorwurfsaktivitäten in der Alltagsinteraktion: grammatische, prosodische, rhetorisch-stilistische und interaktive Verfahren bei der Konstitution kommunikativer Muster und Gattungen*. Tübingen: Max Niemeyer.

— (2002). Stimmvielfalt im Diskurs: Formen der Stilisierung und Ästhetisierung in der Redewiedergabe. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 3, 59-80.

Günthner, S. & Imo, W. (2003). Die Reanalyse von Matrixsätzen als Diskursmarker: *ich mein*-Konstruktionen im gesprochenen Deutsch. *InLiSt* 37, 1-31.

Günthner, S. (2006). Grammatische Analysen der kommunikativen Praxis – 'Dichte Konstruktionen' in der Interaktion. In Deppermann, A., Fiehler, R. & Spranz-Fogasy, T. (Eds.), *Grammatik und Interaktion*. Radolfzell: Verlag für Gesprächsforschung, 95-122.

— (2008a). "weil – es ist zu spät". Geht die Nebensatzstellung im Deutschen verloren? In Denkler, M., Günthner, S., Imo, W., Macha, J., Meer, D., Stoltenburg, B. & Topalović, E. (Eds.), *Frischwärts und unkaputtbar. Sprachverfall oder Sprachwandel im Deutschen*. Münster: Aschendorff, 103-128.

— (2008b). Projektorkonstruktionen im Gespräch: Pseudoclefts, *die Sache ist*-Konstruktionen und Extrapositionen mit *es*. *Online-Zeitschrift zur verbalen Interaktion* 9, 86-114.

— (2009a). Eine Grammatik der Theatralität. Grammatische und prosodische Inszenierungsverfahren in Alltagserzählungen. In Buss, M., Habscheid, S., Jautz, S., Liedtke, F. & Schneider, J. G. (Eds.), *Theatralität des sprachlichen Handelns*. München: Wilhelm Fink, 293-317.

— (2009b). Konstruktionen in der kommunikativen Praxis. Zur Notwendigkeit einer interaktionalen Anreicherung konstruktionsgrammatischer Ansätze. *Zeitschrift für Germanistische Linguistik* 37(3).

— (2011a). Between Emergence and Sedimentation: Projecting Constructions in German Interactions. In Auer, P. & Pfänder, S. (Eds.), *Constructions: Emerging and Emergent*. Berlin/New York: de Gruyter, 156-185.

— (2011b). *N Be That*-Constructions in Everyday German Conversation. A Reanalysis of 'die Sache ist/das Ding ist' ('the Thing Is')-Clauses as *Projector Phrases*. In Laury, R. & Suzuki, R. (Eds.), *Subordination in Conversation: A Cross-Linguistic Perspective*. Amsterdam/Philadelphia: Benjamins, 11-36.

— (2016). Praktiken erhöhter Dialogizität: onymische Anredeformen als Gesten personifizierter Zuwendung. *Zeitschrift für germanistische Linguistik* 44(3), 406-436.

— (2018). Routinisierte Muster in der Interaktion. In Deppermann, A. & Reineke, S. (Eds.), *Sprache im kommunikativen, interaktiven und kulturellen Kontext*. Berlin/Boston: de Gruyter, 29-50.

Haakana, M. (2007). Reported Thought in Complaint Stories. In Clift, R. & Holt, E. (Eds.), *Reporting Talk. Reported Speech in Interaction*. Cambridge: Cambridge University Press, 150-178.

Haakana, M., Laakso, M. & Lindström, J. (2016). Introduction: Comparative Dimensions of Talk in Interaction. In Haakana, M., Laakso, M. & Lindström, J. (Eds.), *Talk in Interaction: Comparative Dimensions*. Helsinki: Finnish Literature Society, 15-47.

Haddington, P. (2006). The Organization of Gaze and Assessments as Resources for Stance Taking. *Text & Talk* 26(3), 281-328.

Haen, M. (2019). Reported Thought in Writing Centre Talk: A Resource for Doing Support and Socialization. *Studies in Applied Linguistics & TESOL* 19(1), 17-34.

Hailaselassie, A. (2015). Voilà*, an Orientation Shift Marker in Modern French Discourse: A Conversation Analytic Perspective*. PhD Thesis. University of Illinois at Urbana-Champaign.

Haiman, J. & Thompson, S. A. (1984). *"Subordination" in Universal Grammar*, Paper presented at the Annual Meeting of the Berkeley Linguistics Society.

Haiman, J. (1994). Ritualization and the Development of Language. In Pagliuca, W. (Ed.), *Perspectives on Grammaticalization*. Amsterdam/Philadelphia: Benjamins, 3-28.

Hausendorf, H. & Quasthoff, U. M. (1995). Discourse and Oral Contextualization: Vocal Cues. In Quasthoff, U. M. (Ed.), *Aspects of Oral Communication*. Berlin/New York: de Gruyter, 220-255.

Hayashi, M., Mori, J. & Takagi, T. (2002). Contingent Achievement of Co-Tellership in a Japanese Conversation: An Analysis of Talk, Gaze, and Gesture. In Ford, C. E., Fox, B. A. & Thompson, S. A. (Eds.), *The Language of Turn and Sequence*. Oxford: Oxford University Press, 81-122.

Heinemann, T. (2005). Where Grammar and Interaction Meet: The Preference for Matched Polarity in Responsive Turns in Danish, *Syntax and Lexis in Conversation: Studies on the Use of Linguistic Resources in Talk-in-Interaction*. Amsterdam/Philadelphia: Benjamins, 375-403.

Heinemann, T. & Traverso, V. (2009). Complaining in Interaction. *Journal of Pragmatics* 41(12), 2381-2384.

Heinold, S. (2015). *Tempus, Modus und Aspekt im Deutschen: Ein Studienbuch*. Tübingen: Narr Francke Attempto Verlag.

Helbig, G. & Buscha, J. (2001). *Deutsche Grammatik. Ein Handbuch für den Ausländerunterricht*. Berlin/München: Langenscheidt.

Heller, V. (2021). Embodied Displays of "Doing Thinking". Epistemic and Interactive Functions of Thinking Displays in Children's Argumentative Activities. *Frontiers in Psychology* 12, 1-21.

Henriksson, H. (2006). *Aspektualität ohne Aspekt? Progressivität und Imperfektivität im Deutschen und Schwedischen*. Stockholm: Almqvist och Wiksell International.

Hentschel, E. & Weydt, H. (2003). *Handbuch der deutschen Grammatik*. Berlin/New York: de Gruyter.

Heritage, J. (2002). *Oh*-Prefaced Responses to Assessments: A Method of Modifying Agreement/Disagreement. In Ford, C. E., Fox, B. A. & Thompson, S. A. (Eds.), *The Language of Turn and Sequence*. Oxford: Oxford University Press, 196-224.

— (2012). The Epistemic Engine: Sequence Organization and Territories of Knowledge. *Research on Language & Social Interaction* 45(1), 30-52.

— (2013). Turn-initial Position and Some of its Occupants. *Journal of Pragmatics* 57, 331-337.

Hoey, E. (2018). Drinking for Speaking: The Multimodal Organization of Drinking in Conversation. *Social Interaction. Video-Based Studies of Human Sociality* 1(1).

— (2020). *When Conversation Lapses: The Public Accountability of Silent Copresence*. Oxford: Oxford University Press.

Holt, E. (1996). Reporting on Talk: The Use of Direct Reported Speech in Conversation. *Research on Language & Social Interaction* 29(3), 219-245.

— (2000). Reporting and Reacting: Concurrent Responses to Reported Speech. *Research on Language & Social Interaction* 33(4), 425-454.

— (2007). 'I'm Eyeing Your Chop Up Mind': Reporting and Enacting. In Holt, E. & Clift, R. (Eds.), *Reporting Talk. Reported Speech in Interaction*. Cambridge: Cambridge University Press, 47-80.

Hopper, P. J. (1987). Emergent Grammar. In Aske, J., Beery, N., Michaelis, L. & Filip, H. (Eds.), *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 139-157.

Hopper, P. J. & Traugott, E. C. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.

Hopper, P. J. & Thompson, S. A. (2008). Projectability and Clause Combining in Interaction. In Laury, R. (Ed.), *Crosslinguistic Studies of Clause Combining. The Multifunctionality of Conjuctions*. Amsterdam/Philadelphia: Benjamins, 99-123.

Hopper, P. J. (2011). Emergent Grammar and Temporality in Interactional Linguistics. In Auer, P. & Pfänder, S. (Eds.), *Constructions: Emerging and Emergent*. Berlin/New York: de Gruyter, 22-44.

Horlacher, A.-S. (2015). *La dislocation à droite revisitée: une approche interactionniste*. Brussels: De Boeck-Duculot.

Horlacher, A.-S. & De Stefani, E. (2017). L'ancrage multimodal du déictique spatial *là* dans l'interaction. *Langue française* 193(1), 21-38.

Hustvedt, S. (2011). Three Emotional Stories: Reflections on Memory, the Imagination, Narrative, and the Self. *Neuropsychoanalysis* 13(2), 187-196.

Hutchby, I. (1992). The Pursuit of Controversy: Routine Skepticism in Talk on 'Talk Radio'. *Sociology* 26(4), 673-694.

Hymes, D. (1962). The Ethnography of Speaking. In Gladwin, T. & Sturtevant, W. C. (Eds.), *Anthropology and Human Behavior*. Washington: Anthropological Society of Washington, 13-53.

Imo, W. (2005). A Construction-Grammar Approach to the Phrase "I mean" in Spoken English. *InLiSt* 42.

— (2007a). *Construction Grammar und Gesprochene-Sprache-Forschung. Konstruktionen mit zehn matrixsatzfähigen Verben im gesprochenen Deutsch*. Tübingen: Niemeyer.

— (2007b). Zur Anwendung der *Construction Grammar* auf die gesprochene Sprache - der Fall "ich mein(e)". In Ágel, V. & Hennig, M. (Eds.), *Zugänge zur Grammatik der gesprochenen Sprache*. Tübingen: Max Niemeyer, 3-34.

— (2009). Inszenierungen eigener und fremder Rede durch Konstruktionen mit dem Verb *sagen*. In Buss, M., Habscheid, S., Jautz, S., Liedtke, F. & Schneider, J. G. (Eds.), *Theatralität des sprachlichen Handelns*. München: Wilhelm Fink, 319-335.

— (2011a). Clines of Subordination – Constructions with the German 'Complement-Taking Predicate' *glauben*. In Laury, R. & Suzuki, R. (Eds.), *Subordination in Conversation: A Cross-Linguistic Perspective*. Amsterdam/Philadelphia: Benjamins, 165-191.

— (2011b). Ad hoc-Produktion oder Konstruktion? – Verfestigungstendenzen bei Inkrement-Strukturen im gesprochenen Deutsch. In Lasch, A. & Ziem, A. (Eds.), *Konstruktionsgrammatik III*. Tübingen: Stauffenburg, 141-256.

— (Ed.) (2012a). *Konstruktionen in der Interaktion*. Berlin/Boston: de Gruyter.

— (2012b). "Da hat des kleine glaub irgendwas angestellt" – ein construct ohne construction? In Günthner, S. & Imo, W. (Eds.), *Konstruktionen in der Interaktion*. Berlin/Boston: de Gruyter, 263-290.

— (2015). Interactional Construction Grammar. *Linguistics Vanguard* 1(1), 69-77.

Imo, W., König, K., Wegner, L. & Weidner, B. (2020). Verfestigungen in der Interaktion – Die Einleitung als musterhaft sedimentierte Prä-Sequenz. In Beate, W., Katharina, K., Wolfgang, I. & Lars, W. (Eds.), *Verfestigungen in der Interaktion: Konstruktionen, sequenzielle Muster, kommunikative Gattungen*. Berlin/Boston: De Gruyter, 1-18.

Iwasaki, S. (2011). The Multimodal Mechanics of Collaborative Unit Construction in Japanese Conversation. In Streeck, J., Goodwin, C. & LeBaron, C. (Eds.), *Embodied Interaction: Language and Body in the Material World*. Cambridge: Cambridge University Press, 106-120.

Jefferson, G. (1978). Sequential Aspects of Storytelling in Conversation. In Schenkein, J. I. M. (Ed.), *Studies in the Organization of Conversational Interaction*. New York: Academic Press, 213-248.

— (2004). "At First I Thought": A Normalizing Device for Extraordinary Events. In Lerner, G. H. (Ed.), *Conversation Analysis. Studies from the First Generation*. Amsterdam/Philadelphia: Benjamins, 131-167.

Jehoul, A., Brône, G. & Feyaerts, K. (2017). The Shrug as Marker of Obviousness: Corpus Evidence from Dutch Face-to-Face Conversations. *Linguistics Vanguard* 3(1), 1-9.

Kallmeyer, W., Keim, I. & Nikitopoulos, P. (1994). Selbst- und Fremddarstellung im Gespräch und Regeln des Sprechens. Untersucht am Beispiel einer Stehcafé-Gruppe in Sandhofen. In Kallmeyer, W. (Ed.), *Kommunikation in der Stadt, Teil 1: Exemplarische Analysen des Sprachverhaltens in Mannheim*. Berlin/New York: de Gruyter, 39-140.

Kärkkäinen, E. (2003). *Epistemic Stance in English Conversation*. Amsterdam/Philadelphia: Benjamins.

— (2006). Stance Taking in Conversation: From Subjectivity to Intersubjectivity. *Text & Talk - An Interdisciplinary Journal of Language, Discourse Communication Studies* 26(6), 699-731.

— (2010). Position and Scope of Epistemic Phrases in Planned and Unplanned American English. In Kaltenböck, G., Mihatsch, W. & Schneider, S. (Eds.), *New Approaches to Hedging*. Bingley: Emerald, 203-236.

— (2012). *I thought It Was Very Interesting*. Conversational Formats for Taking a Stance. *Journal of Pragmatics* 44(15), 2194-2210.

Kärkkäinen, E. & Thompson, S. A. (2018). Language and Bodily Resources: 'Response Packages' in Response to Polar Questions in English. *Journal of Pragmatics* 123, 220-238.

Katelhön, P. & Moroni, M. C. (2018). Inszenierung direkter Rede in mündlichen Interaktionen. *studi germanici* 1, 179-208.

Kaufmann, G. (1976). *Die indirekte Rede und mit ihr konkurrierende Formen der Redeerwähnung*. München: Max Hueber Verlag.

Kaukomaa, T., Peräkylä, A. & Ruusuvuori, J. (2014). Foreshadowing a Problem: Turn-Opening Frowns in Conversation. *Journal of Pragmatics* 71, 132-147.

— (2015). How Listeners Use Facial Expression to Shift the Emotional Stance of the Speaker's Utterance. *Research on Language and Social Interaction* 48(3), 319-341.

Kay, P. (1997). *Words and the Grammar of Context*. Stanford CA: CSLI.

Kearns, K. (2007). Epistemic Verbs and Zero Complementizer. *English Language and Linguistics* 11(3), 475-505.

Keevallik, L. (2013). The Interdependence of Bodily Demonstrations and Clausal Syntax. *Research on Language and Social Interaction* 46(1), 1-21.

— (2015). Coordinating the Temporalities of Talk and Dance. In Deppermann, A. & Günthner, S. (Eds.), *Temporality in Interaction*. Amsterdam/Philadelphia: Benjamins, 309-336.

— (2018). What Does Embodied Interaction Tell Us About Grammar? *Research on Language and Social Interaction* 51(1), 1-21.

Kelly, J. & Local, J. K. (1989). On the Use of General Phonetic Techniques in Handling Conversational Material. In Roger, D. & Bull, P. (Eds.), *Conversation. An Interdisciplinary Perspective*. Clevedon: Multilingual Matters, 197-212.

Kendon, A. (1967). Some Functions of Gaze-Direction in Social Interaction. *Acta psychologica* 26, 22-63.

Kendon, A. & Cook, M. (1969). The Consistency of Gaze Patterns in Social Interaction. *British Journal of Psychology* 60(4), 481-494.

Kendon, A. (1973). The Role of Visible Behavior in the Organization of Social Interaction. In von Cranach, M. & Vine, I. (Eds.), *Social Communication and Movement: Studies of Interaction and Expression in Man and Chimpanzee*. New York: Academic Press, 29-74.

— (2002). Some Uses of the Head Shake. *Gesture* 2(2), 147-182.

— (2004). *Gesture. Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kendrick, K. H., Brown, P., Dingemanse, M., Floyd, S., Gipper, S., Hayano, K., Hoey, E. M., Hoymann, G., Manrique, E. & Rossi, G. (2014). *Sequence Organization: A Universal Infrastructure for Action*, Paper presented at the 4th International Conference on Conversation Analysis, University of California, Los Angeles, CA.

Kendrick, K. H. & Holler, J. (2017). Gaze Direction Signals Response Preference in Conversation. *Research on Language and Social Interaction* 50(1), 12-32.

Kendrick, K. H. (2019). Evidential Vindication in Next Turn, *Perception Metaphors*, 253-274.

Kim, M. S. (2014). Reported Thought as a Stance-Taking Device in Korean Conversation. *Discourse Processes* 51(3), 230-263.

Klewitz, G. & Couper-Kuhlen, E. (1999). Quote – Unquote? The Role of Prosody in the Contextualization of Reported Speech Sequences. *Pragmatics* 9(4), 459-485.

Koch, P. & Oesterreicher, W. (2011). *Gesprochene Sprache in der Romania*. Berlin/New York: de Gruyter.

Koester, A. & Handford, M. (2018). 'It's Not Good Saying "Well it Might Do That or it Might Not"': Hypothetical Reported Speech in Business Meetings. *Journal of Pragmatics* 130, 67-80.

König, K. & Oloff, F. (2018a). Ansätze zu einer multimodalen Erzählanalyse – Einführung in das Themenheft. *Gesprächsforschung. Online-Zeitschrift zur verbalen Interaktion* 19, 207-241.

— (2018b). Die Multimodalität alltagspraktischen Erzählens. *Zeitschrift für Literaturwissenschaft und Linguistik* 48(2), 277-307.

Kotthoff, H. (1998). Irony, Quotation, and Other Forms of Staged Intertextuality: Double or Contrastive Perspectivation in Conversation. *InLiSt* 5, 1-25.

Krahmer, E. & Swerts, M. (2004). More about Brows: A Cross-Linguistic Study via Analysis-by-Synthesis. In Ruttkay, Z. & Pelachaud, C. (Eds.), *From Brows to Trust. Human-Computer Interaction Series*. Berlin: Springer, 191-216.

Labov, W. & Waletzky, J. (1967). Narrative Analysis: Oral Versions of Personal Experience. In Helm, J. (Ed.), *Essays on the Verbal and Visual Arts*. Seattle/London: University of Washington Press, 12-44.

Labov, W. (1972). *Language in the Inner City*. Philadelphia: University of Pennsylvania Press.

Ladewig, S. & Hotze, L. (2020). Zur temporalen Entfaltung und multimodalen Orchestrierung von konzeptuellen Räumen am Beispiel einer Erzählung. *Linguistik online* 104(4), 109-136.

Ladewig, S. H. (2014). Creating Multimodal Utterances: The Linear Integration of Gestures into Speech. In Müller, C., Cienki, A., Fricke, E., Ladewig, S. H., McNeill, D. & Bressem, J. (Eds.), *Body – Language – Communication*. Berlin/Boston: de Gruyter, 1662-1677.

Lamerichs, J. & Te Molder, H. F. M. (2009). 'And Then I'm Really Like…': 'Preliminary' Self-Quotations in Adolescent Talk. *Discourse Studies* 11(4), 401-419.

Lanwer, J. P. (2015). Appositive Syntax oder appositive Prosodie? In Imo, W. & Lanwer, J. P. (Eds.), *Prosodie und Konstruktionsgrammatik*. Berlin/Boston: de Gruyter, 233-281.

Laury, R. & Okamoto, S. (2011a). *Teyuuka* and *I mean* as Pragmatic Parentheticals in Japanese and English. In Laury, R. & Suzuki, R. (Eds.), *Subordination in Conversation: A Cross-Linguistic Perspective*. Amsterdam/Philadelphia: Benjamins, 209-230.

— (Eds.) (2011b). *Subordination in Conversation: A Cross-Linguistic Perspective*. Amsterdam/Philadelphia: Benjamins.

Laury, R. & Ono, T. (2014). The Limits of Grammar: Clause Combining in Finnish and Japanese Conversation. *Pragmatics* 24(3), 561-592.

Laury, R., Ono, T. & Suzuki, R. (2019). Questioning the Clause as a Crosslinguistic Unit in Grammar and Interaction. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 43(2), 364-401.

Laury, R., Helasvuo, M.-L. & Rauma, J. (2020). When an Expression Becomes Fixed: *Mä ajattelin että* 'I thought that' in Spoken Finnish. In Laury, R. & Ono, T. (Eds.), *Fixed Expressions: Building Language Structure and Social Action*. Amsterdam/Philadelphia: Benjamins, 133-166.

Le Goffic, P. (2019). *Grammaire de la subordination*. Paris: Éditions Ophrys.

Leech, G. (1971). *Meaning and the English Verb*. London: Longman.

Lehmann, C. (1988). Towards a Typology of Clause Linkage. In Haiman, J. & Thompson, S. A. (Eds.), *Clause Combining in Grammar and Discourse*. Amsterdam/Philadelphia: Benjamins, 181-225.

— (2015). *Thoughts on Grammaticalization*. Berlin: Language Science Press.

Leistner, A. (2016). *Syntaktische Integration in Redewiedergabe. Eine Untersuchung der direkten und nicht-direkten Redewiedergabeformen in literarischen Texten*. Kassel: Kassel University Press.

Léon, J. (1988). Formes de discours direct dans des récits oraux. *Linx* 18(1), 107-124.

Leonhard, J. (2022). *Die Vergangenheitstempora im Alemannischen Deutschlands. Eine korpusbasierte quantitative und qualitative Untersuchung*. Berlin/Boston: de Gruyter.

Lerner, G. H. (1991). On the Syntax of Sentences-in-Progress. *Language in Society* 20(3), 441-458.

— (1996). On the "Semi-Permeable" Character of Grammatical Units in Conversation: Conditional Entry into the Turn Space of Another Speaker. In Ochs, E., Schegloff, E. A. & Thompson, S. A. (Eds.), *Interaction and Grammar*. Cambridge: Cambridge University Press, 238-276.

Lerner, G. H. (2003). Selecting Next Speaker: The Context-Sensitive Operation of a Context-Free Organization. *Language in Society* 32(2), 177-201.

Levey, S., Groulx, K. & Roy, J. (2013). A Variationist Perspective on Discourse-Pragmatic Change in a Contact Setting. *Language Variation and Change* 25(2), 225-251.

Levinson, S. C. (2013). Action Formation and Ascription. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Malden, MA: Blackwell, 103-132.

Levinson, S. C. & Holler, J. (2014). The Origin of Human Multi-Modal Communication. *Philosophical Transactions of the Royal Society B* 369(1651), 1-9.

Lexer, M. (Ed.) (1992). Mittelhochdeutsches Taschenwörterbuch. Stuttgart: Hirzel.

Li, X. (2014). Leaning and Recipient Intervening Questions in Mandarin Conversation. *Journal of Pragmatics* 67, 34-60.

Lindström, A. & Mondada, L. (2009). Assessments in Social Interaction: Introduction to the Special Issue. *Research on Language and Social Interaction* 42(4), 299-308.

Lindström, A. & Sorjonen, M.-L. (2013). Affiliation in Conversation. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Oxford: Blackwell, 250-369.

Lindström, J., Laury, R., Peräkylä, A. & Sorjonen, M.-L. (Eds.) (2021). *Intersubjectivity in Action: Studies in Language and Social Interaction*. Amsterdam/Philadelphia: Benjamins.

Linell, P. (2013). The Dynamics of Incrementation in Utterance-Building. Processes and Resources. In Szczepek Reed, B. & Raymond, G. (Eds.), *Units of Talk – Units of Action*. Amsterdam/Philadelphia: Benjamins.

Louagie, D. & Verstraete, J.-C. (2016). Noun Phrase Constituency in Australian Languages: A Typological Study. *Linguistic Typology* 20(1), 25-80.

Luke, K.-K. & Zhang, W. (2007). Retrospective Turn Continuations in Mandarin Chinese Conversation. *Pragmatics* 17(4), 605-635.

Luke, K.-K., Thompson, S. A. & Ono, T. (2012). Turns and Increments: A Comparative Perspective. *Discourse Processes* 49(3-4), 155-162.

Macaulay, R. S. (2001). You're Like 'Why Not?' The Quotative Expressions of Glasgow Adolescents. *Journal of Sociolinguistics* 5(1), 3-21.

Mandelbaum, J. (1991). Conversational Non-Cooperation: An Exploration of Disattended Complaints. *Research on Language and Social Interaction* 25(1-4), 97-138.

— (2013). Storytelling in Conversation. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Oxford: Blackwell, 492-508.

Marnette, S. (2005). *Speech and Thought Presentation in French*. Amsterdam/Philadelphia: Benjamins.

Marrese, O. H., Raymond, C. W., Fox, B. A., Ford, C. E. & Pielke, M. (2021). The Grammar of Obviousness: The Palm-Up Gesture in Argument Sequences. *Frontiers in Communication* 6, 1-16.

Maschler, Y., Pekarek Doehler, S., Lindström, J. & Keevallik, L. (Eds.) (2020). *Complex Syntax-in-Interaction: Emergent and Emerging Clause-Combining Patterns for Organizing Social Actions*. Amsterdam/Philadelphia: Benjamins.

Matthiessen, C. & Thompson, S. A. (1988). The Structure of Discourse and 'Subordination'. In Haiman, J. & Thompson, S. A. (Eds.), *Clause-Combining in Grammar and Discourse*. Amsterdam: Benjamins, 275-333.

Maynard, S. K. (1996). Multivoicedness in Speech and Thought Representation. *Journal of Pragmatics* 25, 207-226.

McGregor, W. B. (2019). Reported Speech as a Dedicated Grammatical Domain – and Why Defenestration Should Not Be Thrown Out the Window. *Linguistic Typology* 23(1), 207-219.

Mead, G. H. (1934). *Mind, Self and Society from the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press.

Meehan, T. (1991). It's Like, 'What's Happening in the Evolution of Like?': A Theory of Grammaticalization. *Kansas Working Papers in Linguistics* 16, 37-52.

Michel, G. (1966). Sprachliche Bedingungen der Wortwahl: Eine Untersuchung an Verben der Redeeinführung (Erster Teil). *STUF – Language Typology and Universals* 19(1-6), 103-130.

Moerman, M. (1977). The Preference for Self-Correction in a Thai Conversational Corpus. *Language* 53, 872-882.

Moerman, M. & Sacks, H. (1988). On 'Understanding' in the Analysis of Natural Conversation. In Moerman, M. (Ed.), *Talking Culture: Ethnography and Conversation Analysis*. Philadelphia: University of Pennsylvania Press.

Mondada, L. (2007). Multimodal Resources for Turn-Taking: Pointing and the Emergence of Possible Next Speakers. *Discourse Studies* 9(2), 194-225.

— (2011). Understanding as an Embodied, Situated and Sequential Achievement in Interaction. *Journal of Pragmatics* 43(2), 542-552.

— (2014a). Pointing, Talk, and the Bodies: Reference and Joint Attention as Embodied Interactional Achievements. In Seyfeddinipur, M., Gullberg, Marianne (Ed.), *From Gesture in Conversation to Visible Action As Utterance: Essays in Honor of Adam Kendon*. Amsterdam/Philadelphia: Benjamins, 95-125.

— (2014b). The Local Constitution of Multimodal Resources for Social Interaction. *Journal of Pragmatics* 65, 137-156.

— (2016). Challenges of Multimodality: Language and the Body in Social Interaction. *Journal of Sociolinguistics* 20(3), 336-366.

— (2018). Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. *Research on Language and Social Interaction* 51(1), 85-106.

Moreno, A. (2016). *Le discours rapporté dans les interactions ordinaires: l'effet de la proximité et des communautés de pratique sur sa construction à l'oral et à l'écrit.* Sciences du langage –Linguistique et phonétique générales. Paris, Université Paris Ouest Nanterre La Défense.

Mosegaard Hansen, M.-B. (1998). *The Function of Discourse Particles. A Study with Special Reference to Spoken Standard French.* Amsterdam/Philadelphia: Benjamins.

Mosegaard Hansen, M.-B. (2000). The Syntactic and Semiotic Status of Direct Quotes, with Reference to French. *Transactions of the Philological Society* 98(2), 281-322.

Müller, C. (2004). Forms and Uses of the Palm Up Open Hand: A Case of a Gesture Family? In Müller, C. & Posner, R. (Eds.), *The Semantics and Pragmatics of Everyday Gestures.* Berlin: Weidler Buchverlag, 253-266.

Mushin, I. & Pekarek Doehler, S. (2021). Linguistic Structures in Social Interaction: Moving Temporality to the Forefront of a Science of Language. *Interactional Linguistics* 1, 1-31.

Neveu, F., Moreno, A., Blumenthal, P., Hriba, L., Gerstenberg, A., Meinschaefer, J. & Prévost, S. (2014). Le discours rapporté dans l'interaction: proximité et variabilité. *SHS Web of Conferences* 8, 1685-1699.

Nguyen, H. T. (2015). Source Marking in Represented Talk and Thought in Vietnamese Narratives. *Text & Talk* 35(6), 731-757.

Ningelgen, J. & Auer, P. (2017). Is There a Multimodal Construction Based on Non-deictic *so* in German? *Linguistics Vanguard* 3(1), 1-5.

Nishiyama, H. (2019). Reenactment During Direct Reported Speech and Thought: Coordinated Information from Verbal and Non-verbal Devices. *Fora* 3, 26-48.

Ochs, E. (1996). Linguistic Resources for Socializing Humanity. In Gumperz, J. J. & Levinson, S. C. (Eds.), *Rethinking Linguistic Relativity.* Cambridge: Cambridge University Press, 407-437.

Ochs, E., Schegloff, E. A. & Thompson, S. A. (Eds.) (1996). *Interaction and Grammar.* Cambridge: Cambridge University Press.

Ogden, R. (2018). *The Actions of Peripheral Linguistic Objects: Clicks*, Paper presented at the Proceedings of Laughter Workshop 2018.

Ono, T. & Couper-Kuhlen, E. (2007). Increments in Cross-Linguistic Perspective. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)* 17(4), 505-512.

Park, I. (2018). Reported Thought as (Hypothetical) Assessment. *Journal of Pragmatics* 129, 1-12.

Park, Y. (2009). Interaction between Grammar and Multimodal Resources: Quoting Different Characters in Korean Multiparty Conversation. *Discourse Studies* 11(1), 79-104.

Pekarek Doehler, S., De Stefani, E., Horlacher, A.-S., Jullien, S. & Müller, G. (2010). Configurations paratactiques et grammaire-dans-l'interaction. In Béguelin, M.-J., Avanzi, M. & Corminbœuf, G. (Eds.), *La Parataxe – Structures, marquages et exploitations discursives.* Bern: Peter Lang, 387-409.

Pekarek Doehler, S. (2011). Clause-Combining and the Sequencing of Actions. Projector Constructions in French Talk-in-Interaction. In Laury, R. & Suzuki, R. (Eds.),

Subordination in Conversation: A Cross-Linguistic Perspective. Amsterdam/Philadelphia: Benjamins, 103-149.

— (2016). More Than an Epistemic Hedge: French *Je sais pas* 'I don't know' as a Resource for the Sequential Organization of Turns and Actions. *Journal of Pragmatics* 106, 148-162.

— (2019). At the Interface of Grammar and the Body: *Chais pas* ("dunno") as a Resource for Dealing with Lack of Recipient Response. *Research on Language and Social Interaction* 52(4), 365-387.

— (2021a). Word-Order Affects Response Latency: Action Projection and the Timing of Responses to Question-Word Questions. *Discourse Processes* 58, 328-352.

— (2021b). How Grammar Grows Out of Social Interaction: From Multi-Unit to Single-Unit Question. *Open Linguistics* 7(1), 837-864.

Pekarek Doehler, S., Polak-Yitzhaki, H., Li, X., Stoenica, I. M., Havlík, M. & Keevallik, L. (2021). Multimodal Assemblies for Prefacing a Dispreferred Response: A Cross-Linguistic Analysis. *Frontiers in Psychology* 12.

Pekarek Doehler, S., Keevallik, L. & Li, X. (Eds.) (2022). *The Grammar-Body Interface in Social Interaction*. Frontiers in Psychology.

Peräkylä, A. & Ruusuvuori, J. (2006). Facial Expression in an Assessment. In Knoblauch, H., Schnettler, B., Raab, J. & Soeffner, H.-G. (Eds.), *Video Analysis Methodology and Methods: Qualitative Audiovisual Data Analysis in Sociology*. Frankfurt a. Main: Peter Lang, 127–142.

— (2012). Facial Expressions and Interactional Regulation of Emotion. In Peräkylä, A. & Sorjonen, M.-L. (Eds.), *Emotion in Interaction*. Oxford: Oxford University Press, 64-91.

Persson, R. (2017). La prosodie comme ressource pour l'organisation de l'interaction: état des lieux et illustrations. *Revue française de linguistique appliquée* 22(2), 33-52.

Pfeiffer, M. & Weiß, C. (2022). Reenactments during Tellings: Using Gaze for Initiating Reenactments, Switching Roles and Representing Events. *Journal of Pragmatics* 189(3), 92-113.

Plank, F. (1986). Über den Personenwechsel und den anderer deiktischer Kategorien. *ZGL* 14, 284-308.

Price, G. (1984). *The French Language: Past and Present*. London: Grant and Cutler.

Proske, N. (2019). Emergent Pseudo-Coordination in Spoken German. A Corpus-Based Exploration. *Yearbook of the German Cognitive Linguistics Association* 7(1), 115-136.

Quasthoff, U. M. (1980). *Erzählen in Gesprächen: Linguistische Untersuchungen zu Strukturen und Funktionen am Beispiel einer Kommunikationsform des Alltags*. Tübingen: Narr.

— (2008). Erzählen als interaktive Gesprächsstruktur. In Brinker, K., Antos, G., Heinemann, W. & Sager, S. F. (Eds.), *Text- und Gesprächslinguistik (2. Halbband): Ein internationales Handbuch zeitgenössischer Forschung*. Berlin/New York: de Gruyter, 1293-1309.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London/New York: Longman.

Radtke, P. (2015). Tempus- und Modusgebrauch in der indirekten Rede: diachrone und synchrone Aspekte. In Liedtke, F. & Hundsnurscher, F. (Eds.), *Pragmatische Syntax*. Tübingen: Niemeyer, 97-114.

Raymond, C. W. (2016). *Intersubjectivity, Progressivity, and Accountability: Studies in Turn Design*. PhD Thesis. University of California, Los Angeles.

— (2019). Intersubjectivity, Normativity, and Grammar. *Social Psychology Quarterly* 82(2), 182-204.

Redder, A. (1994). "Bergungsunternehmen" – Prozeduren des Malfeldes beim Erzählen. In Brünner, G. & Graefen, G. (Eds.), *Texte und Diskurse*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Reichenbach, H. (1947). *Elements of Symbolic Logic*. New York: Macmillan.

Rendulić, N. (2015). *Le discours représenté dans les interactions orales. De l'étude des structures en contexte vers la construction de l'image des relations interlocutives*. PhD Thesis. Université d'Orléans.

Ricci Bitti, P. E., Bonfiglioli, L., Melani, P., Caterina, R. & Garotti, P. (2014). Expression and Communication of Doubt/Uncertainty Through Facial Expression. *Ricerche di Pedagogia e Didattica – Journal of Theories and Research in Education* 9(1), 159-178.

Ricœur, P. (1990). *Soi-même comme un autre*. Paris: Seuil.

Riegel, M., Pellat, J.-C. & Rioul, R. (2003). *Grammaire méthodique du français*. Paris: PUF.

Robert, L. P. (Ed.) (2013). Le Petit Robert. Paris: Le Robert.

Robinson, J. D. (2016a). Introduction: Accountability in Social Interaction. In Robinson, J. D. (Ed.), *Accountability in Social Interaction*. Oxford/New York: Oxford University Press, 1-44.

— (Ed.) (2016b). *Accountability in Social Interaction*. Oxford/New York: Oxford University Press.

Robles, J. S. (2015). Extreme Case (Re)formulation as a Practice for Making Hearably Racist Talk Repairable. *Journal of Language and Social Psychology* 34(4), 390-409.

Robles, J. S. & Kurylo, A. (2017). "Let's Have the Men Clean Up": Interpersonally Communicated Stereotypes as a Resource for Resisting Gender-Role Prescribed Activities. *Discourse Studies* 19(6), 673-693.

Rödel, M. (2007). *Doppelte Perfektbildungen und die Organisation von Tempus im Deutschen*. Tübingen: Stauffenburg.

Romaine, S. & Lange, D. (1991). The Use of *Like* as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress. *American Speech* 66(3), 227-279.

Roncador, M. v. (2012). *Zwischen direkter und indirekter Rede*. Tübingen: Niemeyer.

Rosier, L. (1999). *Le discours rapporté en français*. Paris: Ophrys.

Rossano, F. (2010). Questioning and Responding in Italian. *Journal of Pragmatics* 42(10), 2756-2771.

— (2012). *Gaze Behavior in Face-to-face Interaction*. Nijmegen, Radboud University.

Rossari, C. & Jayez, J. (2000). *Du coup* et les connecteurs de conséquence dans une perspective dynamique. *Lingvisticae investigationes* 23(2), 303-326.

Rossi, G. (2020). Other-Repetition in Conversation across Languages: Bringing Prosody into Pragmatic Typology. *Language in Society* 49(4), 495-520.

Rossi, G., Floyd, S. & Enfield, N. J. (2020). Recruitments and Pragmatic Typology. In Floyd, S., Rossi, G. & Enfield, N. J. (Eds.), *Getting Others to Do Things: A Pragmatic Typology of Recruitments*. Berlin: Language Science Press, 1-23.

Rühlemann, C. (2020). What Dialog Is Absent from Constructed Dialog? *English Text Construction* 13(2), 132-157.

Ruusuvuori, J. & Lindfors, P. (2009). Complaining about Previous Treatment in Health Care Settings. *Journal of Pragmatics* 41(12), 2415-2434.

Ruusuvuori, J. & Peräkylä, A. (2009). Facial and Verbal Expressions in Assessing Stories and Topics. *Research on Language and Social Interaction* 42(4), 377-394.

Ruusuvuori, J. (2013). Emotion, Affect and Conversation. In Sidnell, J. & Stivers, T. (Eds.), *The Handbook of Conversation Analysis*. Oxford: Blackwell, 330-350.

Sacks, H., Schegloff, E. A. & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 4, 696-735.

Sacks, H. (1978). Some Technical Considerations of a Dirty Joke. In Schenkein, J. I. M. (Ed.), *Studies in the Organization of Conversational Interaction*. New York: Academic Press, 249-269.

— (1986). Some Considerations of a Story Told in Ordinary Conversations. *Poetics* 15(1), 127-138.

— (1989). An Analysis of the Course of a Joke's Telling in Conversation. In Sherzer, J. & Bauman, R. (Eds.), *Explorations in the Ethnography of Speaking*. Cambridge: Cambridge University Press, 337-353.

— (1992). *Lectures on Conversation (Vols. 1-2)*. Oxford: Blackwell.

Sams, J. (2008). *A Syntactical Analysis of Written English Quotatives*. PhD Thesis. University of Colorado.

— (2010). Quoting the Unspoken: An Analysis of Quotations in Spoken Discourse. *Journal of Pragmatics* 42(11), 3147-3160.

Sandager Sørensen, S. (2021). Affiliating in Second Position: Response Tokens with Rising Pitch in Danish. *Research on Language and Social Interaction* 54(1), 101-125.

Sandlund, E. (2014). Prescribing Conduct: Enactments of Talk or Thought in Advice-Giving Sequences. *Discourse Studies* 16(5), 645-666.

Saville-Troike, M. (2003). *The Ethnography of Communication*. Malden: Blackwell.

Schegloff, E. A. & Sacks, H. (1973). Opening up Closings. *Semiotica* 8, 289-327.

Schegloff, E. A. (1980). Preliminaries to Preliminaries: "Can I Ask You a Question?". *Sociological Inquiry* 50(3-4), 104-152.

— (1982). Discourse as an Interactional Achievement: Some Uses of 'Uh huh' and Other Things That Come between Sentences. In Tannen, D. (Ed.), *Analyzing Discourse: Text and Talk*. Washington, D.C: Georgetown University Press, 71-93.

— (1996). Turn Organization: One Intersection of Grammar and Interaction. In Ochs, E., Schegloff, E. A. & Thompson, S. A. (Eds.), *Interaction and Grammar*. Cambridge: Cambridge University Press, 52-133.

Schegloff, E. A., Ochs, E. & Thompson, S. A. (1996). Introduction. In Ochs, E., Schegloff, E. A. & Thompson, S. A. (Eds.), *Interaction and Grammar*. Cambridge: Cambridge University Press, 1-51.

Schegloff, E. A. (1997). "Narrative Analysis" Thirty Years Later. *Journal of Narrative and Life History* 7(1-4), 97-106.

— (1998). Body Torque. *Social Research* 65(3), 535-596.

— (2000). On Granularity. *Annual Review of Sociology* 26, 715-720.

— (2006). Interaction: The Infrastructure for Social Institutions, the Natural Ecological Niche for Language, and the Arena in Which Culture is Enacted. In Enfield, N. J. & Levinson, S. C. (Eds.), *Roots of Human Sociality: Culture, Cognition and Interaction*. Oxford: Berg, 70-96.

— (2007). *Sequence Organization in Interaction: A Primer in Conversation Analysis*. Cambridge: Cambridge University Press.

Schmidtke-Bode, K. (2014). *Complement Clauses and Complementation Systems: A Cross-Linguistic Study of Grammatical Organization*. PhD Thesis. Friedrich-Schiller-Universität Jena.

Schmidtke-Bode, K. & Diessel, H. (2017). Cross-Linguistic Patterns in the Structure, Function, and Position of (Object) Complement Clauses. *Linguistics* 55(1), 1-38.

Schneider, S. (2007). *Reduced Parenthetical Clauses*. Amsterdam/Philadelphia: Benjamins.

Schoonjans, S. (2013). Gestische Modalpartikeln oder Modalpartikelgesten? Zur Kookkurrenz von Modalpartikeln und Gestikmustern im Deutschen. In Trautmann, C. & Noel Aziz

Hanna, P. S., Barbara (Eds.), *Interaktionen*. Online Publication: Ludwig-Maximilians-Universität München, 91-106.

— (2017). *Multimodal Construction Grammar Issues are Construction Grammar Issues*. Berlin/Boston: de Gruyter.

— (2018). *Modalpartikeln als multimodale Konstruktionen: Eine korpusbasierte Kookkurrenzanalyse von Modalpartikeln und Gestik im Deutschen*. Berlin/Boston: de Gruyter.

Schourup, L. (1982). Quoting with *Go* 'Say'. *American Speech* 57, 149.

Schwitalla, J. (1986). Jugendliche hetzen über Passanten. Drei Thesen zur ethnographischen Gesprächsanalyse. *Linguistische Berichte* 149, 248-261.

Secova, M. (2015). Discours direct chez les jeunes: nouvelles structures, nouvelles fonctions. *Langage et société* 151(1), 131-151.

Secova, M., Gardner-Chloros, P. & Cheshire, J. (2020). *Deux ou trois choses que je sais d'elle* : les variantes émergentes en français multiculturel de la région parisienne. In Diémoz, F., Dostie, G., Hadermann, P. & Lefeuvre, F. (Eds.), *Le français innovant*. Bern: Peter Lang, 107-136.

Selkirk, E. (2011). The Syntax-Phonology Interface. In Goldsmith, J., Riggle, J. & Yu, A. (Eds.), *The Handbook of Phonological Theory*. Oxford: Blackwell Publishing, 435-483.

Selting, M. (1996a). On the Interplay of Syntax and Prosody in the Constitution of Turn-Constructional Units and Turns in Conversation. *Pragmatics* 6(3), 357-388.

— (1996b). Prosody As an Activity-Type Distinctive Cue in Conversation: The Case of So-Called 'Astonished' Questions in Repair Initiation. In Couper-Kuhlen, E. & Selting, M. (Eds.), *Prosody in Conversation*. Cambridge: Cambridge University Press, 231-270.

— (1998). TCUs and TRPs: The Construction of Units in Conversational Talk. *InLiSt* 4.

— (2000). The Construction of Units in Conversational Talk. *Language in Society* 29, 477-517.

Selting, M. & Couper-Kuhlen, E. (Eds.) (2001). *Studies in Interactional Linguistics*. Amsterdam/Philadelphia: Benjamins.

Selting, M. (2007). Lists as Embedded Structures and the Prosody of List Constructions as an Interactional Resource. *Journal of Pragmatics* 39(3), 483-526.

Selting, M., Auer, P., Barth-Weingarten, D., Bermann, J., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U. M., Schütte, W., Stukenbrock, A. & Uhmann, S. (2009). Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* 10, 353-402.

Selting, M. (2010). Affectivity in Conversational Storytelling: An Analysis of Displays of Anger or Indignation in Complaint Stories. *Pragmatics* 20(2), 229-277.

Selting, M., Auer, P., Barth-Weingarten, D., Bermann, J., Bergmann, P., Birkner, K., Couper-Kuhlen, E., Deppermann, A., Gilles, P., Günthner, S., Hartung, M., Kern, F., Mertzlufft, C., Meyer, C., Morek, M., Oberzaucher, F., Peters, J., Quasthoff, U. M., Schütte, W., Stukenbrock, A. & Uhmann, S. (2011). A System for Transcribing Talk-in-Interaction: GAT 2 (Translated and Adapted for English by Elizabeth Couper-Kuhlen and Dagmar Barth-Weingarten). *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* 12, 1-51.

Selting, M. (2012). Complaint Stories and Subsequent Complaint Stories with Affect Displays. *Journal of Pragmatics* 44(4), 387-415.

— (2017). The Display and Management of Affectivity in Climaxes of Amusing Stories. *Journal of Pragmatics* 111, 1-32.

Sidnell, J. (2006). Coordinating Gesture, Talk, and Gaze in Reenactments. *Research on Language and Social Interaction* 39(4), 377-409.

— (2009). Comparative Perspectives in Conversation Analysis. In Sidnell, J. (Ed.), *Conversation Analysis. Comparative Perspectives*. Cambridge: Cambridge University Press, 3-28.

— (2017). Action in Interaction is Conduct under a Description. *Language in Society* 46(3), 313-337.

Skogmyr Marian, K. (2021). Assessing Without Words: Verbally Incomplete Utterances in Complaints. *Frontiers in Psychology* 12, 1-21.

— (2023). *The Development of L2 Interactional Competence: A Multimodal Study of Complaining in French Interactions*. New York/London: Routledge.

Sorjonen, M.-L., Peräkylä, A., Laury, R. & Lindström, J. (2021). Intersubjectivity in Action: An Introduction. In Lindström, J., Laury, R., Peräkylä, A. & Sorjonen, M.-L. (Eds.), *Intersubjectivity in Action: Studies in Language and Social Interaction*. Amsterdam/Philadelphia: Benjamins, 1-22.

Spronck, S. & Nikitina, T. (2019). Reported Speech Forms a Dedicated Syntactic Domain. *Linguistic Typology* 23(1), 119-159.

Stec, K., Huiskes, M. & Redeker, G. (2015). Multimodal Analysis of Quotation in Oral Narratives. *Open Linguistics* 1(1).

Steen, F. & Turner, M. B. (2013). Multimodal Construction Grammar. In Borkent, M., Dancygier, B. & Hinnell, J. (Eds.), *Language and the Creative Mind*. Stanford CA: CSLI Publications, 255-274.

Steensig, J. & Asmuß, B. (2005). Notes on Disaligning 'Yes but' Initiated Utterances in German and Danish Conversations. Two Construction Types for Dispreferred Responses. In Hakulinen, A. & Selting, M. (Eds.), *Syntax and Lexis in Conversation: Studies on the Use of Linguistic Resources in Talk-in-Interaction*. Amsterdam/Philadelphia: Benjamins, 349-375.

Steensig, J. & Drew, P. (2008). Introduction: Questioning and Affiliation/Disaffiliation in Interaction. *Discourse Studies* 10(1), 5-15.

Steensig, J. (2020). Conversation Analysis and Affiliation and Alignment. In Chapelle, C. A. (Ed.), *The Concise Encyclopedia of Applied Linguistics*. Hoboken: Wiley Blackwell, 249-252.

Stein, S. (1995). *Formelhafte Sprache: Untersuchungen zu ihren pragmatischen und kognitiven Funktionen im gegenwärtigen Deutsch*. Bern: Peter Lang.

Stevanovic, M. (2013). Constructing a Proposal As a Thought. *Pragmatics* 23(3), 519-544.

Steyer, K. (1997). *Reformulierungen. Sprachliche Relationen zwischen Äußerungen und Texten im öffentlichen Diskurs*. Tübingen: Narr.

Stivers, T. (2008). Stance, Alignment, and Affiliation During Storytelling: When Nodding Is a Token of Affiliation. *Research on Language and Social Interaction* 41(1), 31-57.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P. & Yoon, K.-E. (2009). Universals and Cultural Variation in Turn-Taking in Conversation. *Proceedings of the National Academy of Sciences* 106(26), 10587-10592.

Stivers, T. & Rossano, F. (2010). Mobilizing Response. *Research on Language and Social Interaction* 43(1), 3-31.

Stivers, T., Mondada, L. & Steensig, J. (2011). Knowledge, Morality and Affiliation in Social Interaction. In Stivers, T., Mondada, L. & Steensig, J. (Eds.), *The Morality of Knowledge in Conversation*. Cambridge: Cambridge University Press, 3-24.

Stoenica, I.-M. & Pekarek Doehler, S. (2020). Relative-Clause Increments and the Management of Reference. In Maschler, Y., Pekarek Doehler, S., Lindström, J. & Keevallik, L. (Eds.), *Emergent Syntax for Conversation: Clausal Patterns and the Organization of Action*. Amsterdam/Philadelphia: Benjamins, 303-330.

Stoenica, I.-M. & Fiedler, S. (2021). Multimodal Practice for Mobilizing Response: The Case of Turn-Final *Tu vois* 'You see' in French Talk-in-Interaction. *Frontiers in Psychology* 12, 1-21.

Stoenica, I. M. (2020). *Actions et conduites mimo-gestuelles dans l'usage conversationnel des relatives en français*. Bern: Peter Lang.

Stokoe, E. (2012). Moving Forward with Membership Categorization Analysis: Methods for Systematic Analysis. *Discourse Studies* 14(3), 277-303.

Streeck, J. (2002). Grammars, Words, and Embodied Meanings: On the Uses and Evolution of *So* and *Like*. *Journal of Communication* 52(3), 581-596.

— (2009). *Gesturecraft: The Manu-Facture of Meaning*. Amsterdam/Philadelphia: Benjamins.

— (2014). Mutual Gaze and Recognition: Revisiting Kendon's 'Gaze Cirection in Two-Person Interaction'. In Seyfeddinipur, M. & Gullberg, M. (Eds.), *From Gesture in Conversation to Gesture as Visible Utterance: Essays in Honor of Adam Kendon*. Amsterdam/Philadelphia: Benjamins, 35-55.

Stukenbrock, A. (2011). Überlegungen zu einem multimodalen Verständnis der gesprochenen Sprache am Beispiel deiktischer Verwendungsweisen des Ausdrucks 'so'. In Dittmar, N. & Bahlo, N. (Eds.), *Beschreibungen für gesprochenes Deutsch auf dem Prüfstand. Analysen und Perspektiven*. Frankfurt a. Main: Lang, 165-193.

— (2012). Imagined Spaces as a Resource in Interaction. *Bulletin Suisse de Linguistique Appliquée* 96, 141-161.

— (2014). Pointing to an 'Empty' Space: *Deixis am Phantasma* in Face-to-Face Interaction. *Journal of Pragmatics* 74, 70-93.

— (2016). Deiktische Praktiken: Zwischen Interaktion und Grammatik. In Deppermann, A., Feilke, H. & Linke, A. (Eds.), *Sprachliche und kommunikative Praktiken*. Berlin/Boston: de Gruyter, 81-126.

— (2018). Blickpraktiken von Sprechern und Hörern bei der Lokaldeixis: Mobile Eye Tracking-Analysen. *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion*, 132-168.

— (2021). Mit Blick auf die Geste – multimodale Verfestigungen in der Interaktion. In Weidner, B., König, K., Imo, W. & Wegner, L. (Eds.), *Verfestigungen in der Interaktion*. Berlin/Boston: de Gruyter, 231-262.

Svennevig, J. (2004). Other-Repetition as Display of Hearing, Understanding and Emotional Stance. *Discourse Studies* 6(4), 489-516.

— (2013). Reformulation of Questions with Candidate Answers. *International Journal of Bilingualism* 17(2), 189-204.

Tannen, D. (1986). Introducing Constructed Dialogue in Greek and American Conversational Literary Narrative. In Coulmas, F. (Ed.), *Direct and Indirect Speech*. Berlin/New York: de Gruyter, 311-332.

Thompson, S. A. & Mulac, A. (1991a). A Quantitative Perspective on the Grammaticization of Epistemic Parentheticals in English. In Traugott, E. C. & Heine, B. (Eds.), *Approaches to Grammaticalization*. Amsterdam/Philadelphia: Benjamins, 313-329.

— (1991b). The Discourse Conditions for the Use of the Complementizer *that* in Conversational English. *Journal of Pragmatics* 15(3), 237-251.

Thompson, S. A. (2002). "Object Complements" and Conversation. Towards a Realistic Account. *Studies in Language* 26(1), 125-164.

Thompson, S. A. & Couper-Kuhlen, E. (2005). The Clause as a Locus of Grammar and Interaction. *Discourse Studies* 7(4-5), 481-505.

Thompson, S. A. & Suzuki, R. (2014). Reenactments in Conversation: Gaze and Recipiency. *Discourse Studies* 16(6), 816-846.

Thompson, S. A. (2019). Understanding 'Clause' as an Emergent 'Unit' in Everyday Conversation. *Studies in Language* 43(2), 254-280.

Thurmair, M. (1989). *Modalpartikeln und ihre Kombinationen*. Tübingen: Niemeyer.

Traugott, E. C. & Heine, B. (Eds.) (1991a). *Approaches to Grammaticalization (Volume 1). Theoretical and Methodological Issues*. Amsterdam/Philadelphia: Benjamins.

— (Eds.) (1991b). *Approaches to Grammaticalization (Volume 2). Types of Grammatical Markers*. Amsterdam/Philadelphia: Benjamins.

Traugott, E. C. (1995). Subjectification in Grammaticalization. In Stein, D. & Wright, S. (Eds.), *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge: Cambridge University Press, 31-54.

Traverso, V. (2009). The Dilemmas of Third-Party Complaints in Conversation between Friends. *Journal of Pragmatics* 41(12), 2385-2399.

Van Bogaert, J. (2010). A Constructional Taxonomy of *I think* and Related Expressions: Accounting for the Variability of Complement-Taking Mental Predicates. *English Language & Linguistics* 14(3), 399-427.

Vandelanotte, L. (2009). *Speech and Thought Representation in English: A Cognitive-Functional Approach*. Berlin/New York: de Gruyter.

Vásquez, C. & Urzúa, A. (2009). Reported Speech and Reported Mental States in Mentoring Meetings: Exploring Novice Teacher Identities. *Research on Language and Social Interaction* 42(1), 1-19.

Vendler, Z. (1967). Linguistics in Philosophy. Ithaka/New York: Cornell University Press.

Verhagen, A. (2001). Subordination and Discourse Segmentation Revisited, or: Why Matrix Clauses May Be More Dependent Than Complements. In Sanders, T., Schilperoord, J. & Spooren, W. (Eds.), *Text Representation: Linguistic and Psycholinguistic Aspects*. Amsterdam/Philadelphia: Benjamins, 337-357.

Verstraete, J.-C. (2008). The Status of Purpose, Reason, and Intended Endpoint in the Typology of Complex Sentences: Implications for Layered Models of Clause Structure. *Linguistics* 46(4), 757-788.

Vincent, D. & Dubois, S. (1997). *Le discours rapporté au quotidien*. Québec: Nuit Blanche éditeur.

Vliegen, M. (2010). Verbbezogene Redewiedergabe: Subjektivität, Verknüpfung und Verbbedeutung. *Deutsche Sprache* 38, 210-233.

Vološinov, V. N. (1973). *Marxism and the Philosophy of Language*. New York/London: Seminar Press.

Weerd, v. d. (2019). "Those Foreigners Ruin Everything Here": Interactional Functions of Ethnic Labelling among Pupils in the Netherlands. *Journal of Sociolinguistics* 23(3), 244-262.

Weinrich, H. (1993). *Textgrammatik der deutschen Sprache*. Mannheim: Dudenverlag.

Wessels, E. (2019). Gedankenwiedergabe mit "PERS+*denk*-Aff+so+*x* / *denk*-Aff+PERS+so+*x*" in Alltagserzählungen von Jugendlichen und jungen Erwachsenen – Eine erste Form- und Funktionsbestimmung. In Engelberg, S., Fortmann, C. & Rapp, I. (Eds.), *Rede- und Gedankenwiedergabe in narrativen Strukturen – Ambiguitäten und Varianz*. Hamburg: Buske, 55-74.

Whitehead, K. A. (2015). Everyday Antiracism in Action: Preference Organization in Response to Racism. *Journal of Language and Social Psychology* 34(4), 374-389.

Whitehead, K. A. & Stokoe, E. (2015). Producing and Responding to -isms in Interaction. *Journal of Language and Social Psychology* 34(4), 368-373.

Whitehead, K. A. (2019). Using Ethnomethodology and Conversation Analysis to Study Social Categories: The Case of Racial Categories in South African Radio Talk. In Flynn, A. & Kramer, S. (Eds.), *Transforming Research Methods in the Social Sciences: Case Studies from South Africa*. Johannisburg: Wits University Press.

Wierzbicka, A. (1974). The Semantics of Direct and Indirect Discourse. *Papers in Linguistics* 7, 267-308.

Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus*. London: Routledge & Kegal Paul.

— (2014 [1921]). *Logisch-philosphische Abhandlung*. Frankfurt a. M.: Suhrkamp.

Wolff, G. (2004). *Deutsche Sprachgeschichte von den Anfängen bis zur Gegenwart*. Tübingen: Francke.

Wolfson, N. (1978). A Feature of Performed Narrative: The Conversational Historical Present. *Language in Society* 7(2), 215-237.

Yule, G. & Mathis, T. (1992). The Role of Staging and Constructed Dialogue in Establishing Speaker's Topic. *Linguistics* 30, 199-215.

Zeman, S. (2010). *Tempus und "Mündlichkeit" im Mittelhochdeutschen: Zur Interdependenz grammatischer Perspektivensetzung und "Historischer Mündlichkeit" im mittelhochdeutschen Tempussystem*. Berlin/New York: de Gruyter.

— (2013). Vergangenheit als Gegenwart? Zur Diachronie des "Historischen Präsens". *Jahrbuch für Germanistische Sprachgeschichte* 4(1), 236-256.

Ziem, A. (2017). Do We Really Need a Multimodal Construction Grammar? *Linguistics Vanguard* 3(1).

Zifonun, G., Hoffmann, L. & Strecker, B. (1997). *Grammatik der deutschen Sprache, Bd. 3*. Berlin/New York: de Gruyter.

Zima, E. & Bergs, A. (2017). Multimodality and Construction Grammar. *Linguistics Vanguard* 3(1).

Zima, E. & Weiß, C. (2020). Erzählen als multimodale Aktivität – zur Einführung in das Themenheft. *Linguistik online* 104(4), 3-14.

Zinken, J. & Ogiermann, E. (2013). Responsibility and Action: Invariants and Diversity in Requests for Objects in British English and Polish Interaction. *Research on Language and Social Interaction* 46(3), 256-276.

10.   Appendix

10.1.   Tables of all occurrences of German *denken* (present and past tense)

Form	Number of occurrences
wo ich gedacht habe (where I thought have)	1
und dann habe ich gedacht (and then have I thought)	3
und dann habe ich so gedacht (and then have I like thought)	1
dann habe ich gedacht (then have I thought)	3
da habe ich gedacht (there have I thought)	2
da habe ich schon gedacht (there have I PART thought)	1
und dann habe ich halt gedacht so (and then have I PART thought like)	1
und ich habe gedacht (and I have thought)	1
und da habe ich schon gedacht so (and there have I PART thought like)	1
und da habe ich halt gedacht (and there have I PART thought)	1
da habe ich immer gedacht (there have I always thought)	1
da habe ich auch gedacht (there have I also thought)	1
und ich habe erst gedacht (and I have first thought)	1
weil ich gedacht habe (because I thought have)	1
adverbial(phrase) habe ich (nur/schon) gedacht (adverbial(phrase) have I (PART) thought)	3
ich habe gedacht (I have thought)	17
Verb-first syntax: habe (ich) gedacht (have (I) thought)	11
coordination ellipsis	5
Pivot	1

ich habe immer gedacht (I have always thought)		3
ich habe noch gedacht (I have PART thought)	289	1
ich habe jetzt gedacht (I have PART thought)		1
ich habe echt gedacht (I have really thought)		4
ich habe aber auch gedacht dass (I have but also thought that)		1
ich habe nur als gedacht (I have just sometimes thought)		1
ich habe halt gedacht (I have PART thought)		2
ich habe aber gedacht (I have but thought)		1
ich habe schon gedacht (I have PART thought)		2
ich habe ja gedacht (I have PART thought)		1
zwischendurch habe ich schon gedacht so (sometimes have I PART thought like)		1
TOTAL		74

Table 45: Remaining occurrences of *ich habe gedacht* (all actions)

Form	Number of occurrences
wo ich mir gedacht habe (where I to myself thought have)	3
wo ich mir halt auch so gedacht habe (where I to myself also like thought have)	1
wo ich selber mir so gedacht habe (where I myself to myself like thought have)	1
und da habe ich mir nur so gedacht (and there have I to myself just like thought)	1
dann habe ich mir halt gedacht (then have I to myself PART thought)	1
aber dann habe ich mir aber gedacht (but then have I to myself but thought)	1

da habe ich mir gedacht (there have I to myself thought)		2
da habe ich mir erst mal gedacht (there have I to myself first of all thought)		1
und dann habe ich mir gedacht (and then have I to myself thought)		2
dann habe ich mir aber gedacht (then have I to myself but thought)		1
weil ich mir gedacht habe (because I have to myself thought)		3
ich habe mir gedacht (I have to myself thought)		2
Verb-first syntax: hab (ich) mir gedacht (have (I) to myself thought)		8
ich habe mir jetz auch so gedacht so (I have to myself just also like thought like)		1
ich habe mir auch gedacht (I have to myself also thought)		1
ich habe mir einfach gedacht (I have to myself just thought)		1
ich habe mir extra gedacht (I have to myself specifically thought)		1
ich habe mir gestern auch gedacht (I have to myself yesterday also thought)		1
ich habe mir schon gedacht (I have to myself already thought)		1
ich habe mir auch schon gedacht (I have to myself also already thought)		1
ich habe mir schon die ganze zeit gedacht (I have to myself already the whole time thought)		1
TOTAL		35

Table 46: Remaining occurrences of *ich habe mir gedacht* (reported thought)

Form	Number of occurrences
NP dachte ich auch (NP thought I also)	1
NP dachte ich nämlich (NP thought I namely)	1
was ich kurz dachte is (what I briefly thought is)	1
was ich auch schon dachte ist (what I also already thought is)	1
ich dachte nicht dass (I thought not that)	1
ich dachte jetzt auch nicht dass (I thought now also not that)	1
TOTAL	6

Table 47: Excluded occurrences of *ich dachte (mir)* due to grammatical features

Form	Number of occurrences
und ich dachte (and I thought)	2
und ich dachte irgendwann (and I thought at some point)	1
und da dacht ich so (and there thought I like)	1
und dann dacht ich so (and then thought I like)	1
und dann dachte ich (and then thought I)	2
und dann dacht ich auch so (and then thought I also like)	1
dass ich so dachte (that I like thought)	1
und ich dachte so (and I thought like)	4
dann dachte ich (then thought I)	2
ich dachte so (I thought like)	1
aber dann dachte ich so (but then thought I like)	1
dann dachte ich so (then thought I like)	1
und dann dachte ich aber so (and then thought I but like)	1
und ich dachte schon so	2

(and I thought PART like)	
und ich dachte halt so (and I thought PART like)	1
dacht ich (clause-final syntax) (thought I)	2
ich dachte (I thought)	50
Verb-first syntax: dacht (ich) thought (I)	4
coordination ellipsis	8
ich dachte immer (I thought always)	7
ich dachte jetzt einfach (I thought now just)	1
ich dachte irgendwie (I thought somehow)	1
ich dacht grad (I thought now)	1
ich dachte halt (I thought PART)	1
ich dachte jetzt gerade (I thought just now)	1
ich dachte auch (I thought also)	1
ich dachte auch immer (I thought also always)	1
ich dachte ja immer (I thought PART always)	1
ich dachte auch echt (I thought also really)	1
ich dachte schon (I thought PART)	1
ich dachte eigentlich (I usually thought)	1
ich dachte noch (I thought PART)	2
jetzt dachte ich noch (now thought I PART)	1
ich dachte halt so (I thought PART like)	1
ich dachte grad so (I thought just like)	1
ich dachte eher (I thought more)	1

TOTAL	110

Table 48: Total occurrences of *ich dachte* (all actions)

Form	Number of occurrences
und ich dachte mir halt (and I thought to myself PART)	1
und dann dachte ich mir (and then thought I to myself)	1
und dann dachte ich mir so (and then thought I to myself like)	2
und ich dachte mir grad (and I thought to myself now)	1
und ich dachte mir (and I thought to myself)	1
und ich dachte mir nur (and I thought to myself only)	1
aber ich dachte mir auch (but I thought to myself also)	1
verb-first syntax: dachte (ich) mir (thought (I) to myself)	3
coordination ellipsis	3
ich dachte mir nur (I thought to myself only)	1
ich dachte mir auch (I thought to myself also)	1
ich dachte mir schon dass (I thought to myself already that)	1
TOTAL	17

Table 49: All occurrences of *ich dachte mir* introducing an open slot

Form	Number of occurrences
ich denke da auch immer nicht dran (I think there also always not of it)	1
also ich denke auch immer nicht dran (well I think also always not of it)	1
und dann denke ich jetz dran (and then think I now of it)	1
da denke ich immer an x (there think I always of x)	1
so etwas denke ich (something like that think I)	1
das denke ich (that think I)	1
jetzt denke ich es auch nicht mehr (now think I it also not anymore)	1
wenn ich daran denke (dass) (when I think of it (that))	2
ich denke es (I think it)	1
so denke ich mir das (so think I to myself it)	1
ich denke es mir auch immer (I think it to myself also always)	1

Table 50: *Ich denke (mir)* filtered due to grammatical features

Form	Number of occurrences
aber ich denke mal (but I think PART)	2
aber ich denke mal schon (but I think PART PART)	1
aber ich denke (but I think)	2
aber ich denke dass (but I think that)	1
aber irgendwie denke ich halt auch (but somehow think I PART also)	1
also ich denke mal (well I think PART)	1
also ich denke (well I think)	2
und ich denke (and I think)	3
und ich denke so (and I think like)	1
und ich denke immer so (and I think always like)	1

und jedes mal denke ich (and every time think I)	1
und deshalb denke ich so (and therefore think I like)	1
und ich denke dass (and I think that)	1
und ich denke (and I think)	1
und wo ich denke (and where I think)	1
wo ich denke dass (where I think that)	1
wo ich denke (where I think)	3
wo ich so denke so (where I like think like)	2
da denke ich (there think I)	1
da denke ich so (there think I like)	1
dann denke ich (then think I)	1
ADV denke ich so (ADV think I like)	2
wenn x denke ich halt (when x think I PART)	1
Parenthesis	6
Verb-first syntax	5
Coordination ellipsis	8
ich denke dass (I think that)	7
ich denke (I think)	21
turn-final denke ich (I think)	9
ich denke mal (I think PART)	8
denke ich mal (turn-final) (I think PART)	4
ich denke mal dass (I think PART that)	1
ich denke schon (I think PART)	6
ich denke auch (I think also)	2

ich denke noch (I think PART)	1
ich denke halt (I think PART)	1
ich denke schon dass (I think PART that)	2
ich denke immer (I think always)	1
immer wenn x denk ich da dran (always when x think I there of it)	1
aber wenn ich daran denke (but when I of it think)	1
würde ich denken (would I think)	1
deswegen denke ich mal dass (therefore think I PART that)	1
vielleicht denke ich auch nur (maybe think I also only)	1
TOTAL	119

Table 51: All occurrences of *ich denke* as clause introducing an open slot

Form	Number of occurrences
aber dann DENKE ich mir (but then THINK I to myself)	1
weil ich mir halt denke (because I to myself PART think)	1
und dann denke ich mir das (and then think I to myself that)	1
und da denke ich mir (and there think I to myself)	1
und dann denke ich mir so (and then think I to myself like)	1
und dann denke ich mir (and then think I to myself)	1
und deswegen denke ich mir halt so manchmal (and therefore think I to myself PARt like sometimes)	1
und deswegen denke ich mir (and therefore think I to myself)	1
und ich denke mir so (and I think to myself like)	1
und ich denke mir (and I think to myself)	1
wo ich mir denke	2

(where I to myself think)	
wo ich mir (echt) denke (where I to myself (really) think)	1
dann denke ich mir (then think I to myself)	1
da denke ich mir auch so (there think I to myself also like)	1
Verb-first syntax	6
ich denke mir so (I think to myself like)	1
TOTAL	22

Table 52: All occurrences of *ich denke mir* as clause introducing an open slot

10.2.   Glossing Abbreviations

1 first person
3 third person
ADJ adjective
ADV adverb(ial)
AUX auxiliary
DEIK deictic
DET determiner
IPFV imperfective
KONJ conjunction
PART particle
PP personal pronoun
PL plural
PREF prefix
PREP preposition
PRF perfect
PRO pronoun
PRS present
PRT preterit
PST past
PTCP participle
QUOT quotative
REFL reflexive
QUOTE purported thought introduced by quotative
SG singular

10.3.   Transcription Conventions GAT2 for Talk

The conventions for speech and its linguistic and suprasegmental features follow Selting *et al.* (2009). They have been slightly adapted to the present work. All translations are taken from the official adaptation of GAT2 for English (Selting *et al.* 2011).

Sequential structure

`[ ] []`	overlap and simultaneous talk
`=`	Fast, immediate continuation with a new turn or segment (latching)

In- and outbreaths

`°h / h°`	in-/outbreaths of appr. 0.2-0.5 sec. duration
`°hh / hh°`	in-/outbreaths of appr. 0.5-0.8 sec. duration
`°hhh / hhh°`	in-/outbreaths of appr. 0.8-1.0 sec. duration

Pauses

`(.)`	micro pause, estimated, up to appr. 0.2 sec. duration
`(-)`	short estimated pause of appr. 0.2-0.5 sec. duration
`(--)`	intermediary estimated pause of appr. 0.5-0.8 sec. duration
`(---)`	longer estimated pause of appr. 0.8-1.0 sec. duration
`(0.5)/(2.0)`	measured pause of appr. 0.5 / 2.0 sec. duration (to tenth of a second)

Other segmental conventions

`and_uh`	cliticizations within units
`uh, uhm, etc.`	hesitation markers, so-called "filled pauses"
`:`	lengthening, by about 0.2-0.5 sec.
`::`	lengthening, by about 0.5-0.8 sec.
`:::`	lengthening, by about 0.8-1.0 sec.
`/`	cut-off

Accentuation

`SYLlable`	focus accent
`sYllable`	secondary accent
`!SYL!lable`	extra strong accent

Final pitch movements of intonation phrases

?	rising to high
,	rising to mid
–	level
;	falling to mid
.	falling to low

Pitch jumps

↑	smaller pitch upstep
↓	smaller pitch downstep
↑↑	larger pitch upstep
↓↓	larger pitch downstep

Loudness and tempo changes, with scope

`<<f>`	`>`	forte, loud
`<<ff>`	`>`	fortissimo, very loud
`<<p>`	`>`	piano, soft
`<<p>`	`>`	pianissimo, very soft
`<<cresc>`	`>`	crescendo, increasingly louder
`<<dim>`	`>`	decrescendo, increasingly softer
`>>word<<`		allegro, fast
`<<word>>`		lento, slow

Laughter and crying

`haha`	
`hehe`	syllabic laughter
`hihi`	
`((laughs))`	description of laughter and crying
`((cries))`	
`<<laughing>  >`	laughter particles accompanying speech with indication of scope
`<<:-)>     >`	smile voice

Continuers

`hm, yes, no, yeah`	monosyllabic tokens
`hm_hm, ye_es, no_o`	bi-syllabic tokens
`ʔhmʔhm`	with glottal closure, often negating

Changes in voice quality and articulation, with scope

`<<creaky>     >`	glottalized
`<<whispery>   >`	change in voice quality as stated

Other conventions

`<<surprised>  >`	interpretive comment with indication of scope
`((coughs)) <<coughing>`	non-verbal vocal actions and events with indication of scope
`( )`	unintelligible passage
`(xxx), (xxx xxx) (may i)`	one or two unintelligible syllables assumed wording
`(may i say/let us say)`	possible alternatives
`((unintelligible, appr. 3 sec))`	unintelligible passage with indication of duration
`((...))`	omission in transcript

10.4. Transcription Conventions for Multimodal Conduct

The conventions for bodily conduct are based on Mondada (2018). For more information, see https://franzoesistik.philhist.unibas.ch/fileadmin/user_upload/franzoesistik/home/Personen/Mondada/Unterordner/Mondada_conv_multimodality.pdf.

+	+	Indicates start and ending of gaze of speaker A.
%	%	Indicates start and ending of a gesture of speaker A.
@	@	Indicates start and ending of another embodied conduct (e.g., nodding) of speaker A.
$	$	Indicates start and ending of gaze of speaker B.
§	§	Indicates start and ending of a gesture of speaker B.
Ω	Ω	Indicates start and ending of another embodied conduct (e.g., nodding) of speaker B.
¢	¢	Indicates start and ending of gaze of speaker C.
∫	∫	Indicates start and ending of a gesture of speaker C.
¥	¥	Indicates start and ending of gaze of speaker D.
△	△	Indicates start and ending of a gesture of speaker D.
+--->l.12		Continuation of the described embodied conduct until line 12 of transcript
+------->		Continuation of the described embodied conduct until the following line (not necessarily coinciding with the next intonation unit)
------->+		End of the described embodied conduct.
+------>>		Continuation of the described embodied conduct until end of excerpt.
------>>+		Described embodied conduct has been going on since the beginning of the excerpt
........+		Preparation phase of an embodied conduct
,,,,,,,,+		Retraction phase of an embodied conduct