DOI 10.35662/unine-thesis-2834

‘Well, like, oh yeah…’ Non-dialogical functions of discourse markers in journalistic writing

Samuel Bourgeois
Faculté des lettres et science humaines
Institut de langue et littérature anglaises
Linguistique anglaise
Université de Neuchâtel

Thesis Supervisor:
Professor Martin Hilpert

Thesis Defense Jury Members:
Professor Anita Auer, Université de Lausanne
Professor Christian Mair, Albert-Ludwigs-Universität Freiburg
Professor Elena Smirnova, Université de Neuchâtel

Date of thesis defense:
18 August 2020

unine
UNIVERSITÉ DE
NEUCHÂTEL

Faculté des Lettres et Sciences Humaines
Espace Tilo-Frey 1
CH - 2000 Neuchâtel
Tél : +41 (0)32 718 17 00
Fax : +41 (0)32 718 17 71
secretariat.lettres@unine.ch

IMPRIMATUR

La Faculté des lettres et sciences humaines de l'Université de Neuchâtel, sur les rapports de M. Martin Hilpert, directeur de thèse, professeur, Université de Neuchâtel ; Mme Anita Auer, professeure, Université de Lausanne ; M. Christian Mair, professeur, Albert-Ludwigs-Universität Freiburg i. Br., Allemagne ; Mme Elena Smirnova, professeure, Université de Neuchâtel autorise l'impression de la thèse présentée par M. Samuel Bourgeois en laissant à l'auteur la responsabilité des opinions énoncées.

Neuchâtel, le 18 août 2020

Le doyen
Pierre Alain Mariaux

Acknowledgements

The completion of this thesis would not have been possible without the support of family, friends and co-workers. First and foremost, I must thank Kathrin and Logan for allowing me to be more absent than I wished to be in order to finish this work. I also wish to thank my supervisor Martin Hilpert for all of his advice and for reading over the various drafts that came before the final product. My big brother Jake also deserves much thanks for taking on the unofficial (and unpaid) job as my primary editor for the final version of this thesis. The final version of this thesis owes much to his detailed (and sometimes witty) contributions.

I would like to also thank everyone from the Neuchâtel English Institute for their constant support. Special thanks, of course, go to Anne-Claire for all of the fun chats over coffee breaks throughout my five years as a Doctoral Assistant. Without such moments, the completion of this thesis would not have been possible. I would also like to thank François for all of his help throughout my time here (MA-PhD) with all things administrational and otherwise.

Abstract

This thesis investigates the colloquial and highly (inter)subjective non-dialogical uses of discourse markers (DMs), especially *well*, *like* and *oh yeah*, in American journalistic texts based on corpus data from the Corpus of Historical American English (COHA) and the News on the Web (NOW). It demonstrates that such uses are a recent phenomenon, which have gained rapidly in frequency since the 1990s. In journalism, these DMs are developing genre-specific functions that resemble their textual oral functions, but are, at the same time, adapted to suit the needs of writing in (inter)subjective functions (Aijmer 2013). Moreover, these DMs are increasingly used to highlight lexical choices. These developments contribute further insights into what is already known about change occurring in journalistic writing. This thesis discusses how these colloquial DMs are adapted to journalistic writing in order to convey (inter)subjective meanings. Furthermore, the timing of these developments indicates a later wave of change to journalism that has not received sufficient attention in colloquialization studies.

Keywords: discourse markers, colloquialization, (inter)subjectification, journalistic texts, written English, oral English, language change, stance

Résumé

Cette thèse examine les usages colloquiaux et hautement (inter)subjectifs non-dialogiques des marqueurs discursifs, en particulier *well*, *like* et *oh yeah*, dans les textes journalistiques américains basés sur les données du Corpus of Historical American English (COHA) et celles du News on the Web (NOW) corpus. Elle démontre que ces utilisations sont un phénomène récent, qui a rapidement gagné en fréquence depuis les années 1990. En journalisme, ces marqueurs discursifs développent des fonctions spécifiques au genre qui ressemblent à leurs fonctions orales textuelles, mais qui sont en même temps adaptées aux besoins de l'écriture dans des fonctions (inter)subjectives (Aijmer 2013). Par ailleurs, ces marqueurs sont de plus en plus utilisés pour mettre en évidence des choix lexicaux. Ces développements apportent un éclairage supplémentaire sur ce que l'on sait déjà des changements qui se produisent dans l'écriture journalistique. Cette thèse examine comment ces marqueurs discursifs colloquiaux sont adaptés à l'écriture journalistique afin de transmettre des significations (inter)subjectives. En outre, le timing de ces développements indique une vague ultérieure de changement dans le journalisme qui n'a pas reçu suffisamment d'attention dans les études sur la colloquialisation.

Mots clés: marqueurs discursifs, colloquialisation, (inter)subjectification, textes journalistiques, anglais écrit, anglais oral, changement linguistique, posture linguistique

Table of Contents

Chapter: 1
Introduction 11

1.1 Introduction 11

1.2 Colloquialization 15

1.3 Speech and writing 16

1.4 Dialogical DMs in Writing 18

1.5 Research questions 21

1.6 Conclusions 23

Chapter 2:
Theoretical and methodological background 25

2.1 Introduction 25

2.2 Discourse Markers: Terminology and Functions 25

2.3 Theories that account for the development of DMs and their contextual relations 37

2.3.1 (Inter)subjectification 37

2.3.2 DMs and Genre-Specific Functions 44

2.4 Changes occurring in written English with a focus on journalistic writing 48

2.5 Selection of DMs for the case studies 57

2.6 Selection of Corpora 58

2.7 Methodology 59

2.8 Summary and conclusions 65

Chapter 3:
Non-dialogical *well*: Rhetorical marker of (inter)subjective word choice 67

3.1 Introduction 67

3.2 Previous work on *well* 69

3.3 Frequency trends of non-dialogical *well* in the COHA 76

3.4 Qualitative Analysis of non-dialogical *well* in the COHA data 89

3.4.1 Quote *Well* 90

3.4.2 Clause *Well* 92

3.4.3 Word-choosing marker *well* and predicative *well* 97

3.5 Discussion of written functions 108

3.6 Non-dialogical *well* in contemporary journalism data from the NOW Corpus 110

3.7 Functions of Non-dialogical *well* in the NOW Corpus 112

3.7.1 Quote *well* and Clause *well* 113

3.7.2 WCM and pred *well* in the NOW data 115

3.7.3 Meta-textual comment *well* 117

3.7.4 Discussion of changes from COHA to NOW 118

3.8 Conclusions 120

Chapter 4:
Like: from stigmatization to rhetorical marker of word choice 125

4.1 Introduction 125

4.2 Previous work on *like* 127

4.3 Trends of *like* in the COHA Data ... 137
4.4 Qualitative Analysis of non-dialogical *like* in the COHA (1990s-2000s) ... 141
4.4.1 Discussion of *like* in the COHA in the 1990s and 2000s ... 149
4.5 Trends of *like* in the News on the Web Corpus (NOW) ... 150
4.6 Qualitative Analysis of non-dialogical *like* in the NOW data ... 152
4.6.1 Focuser and hedge functions of *like* in the NOW data ... 156
4.6.2 Exemplifying and Approximative functions in the NOW data. ... 159
4.6.3 Hypothetical quotative *like* in the NOW data ... 162
4.6.4 Discussion of non-dialogical NOW data ... 163
4.7 Conclusions ... 167

Chapter 5:
Oh Yeah: Rhetorical marker of (inter)subjective information management ... 171
5.1 Introduction ... 171
5.2 Theoretical background of *oh yeah* ... 174
5.3 Trends of *oh yeah* in the COHA data ... 179
5.4 Qualitative analysis of non-dialogical *oh yeah* in the COHA data (1960s-2000s) ... 182
5.5 Trends of *oh yeah* in the News on the Web Corpus (NOW) ... 189
5.6 Qualitative analysis of non-dialogical *oh yeah* in the NOW data (2010-2018) ... 191
5.6.1 Faux spontaneous additive *oh yeah* ... 194
5.6.2 Faux utterance launcher *oh yeah* ... 196
5.6.3 Meta-textual comments ... 198
5.6.4 Discussion of non-dialogical NOW data ... 201
5.7 Conclusions ... 209

Chapter 6:
Discussion of common observations from the three case studies ... 213
6.1 What do these three DMs in writing have in common? ... 213
6.2 Comparing the functionalities of *well, like* and *oh yeah* in journalistic prose ... 214
6.3 Comparing the developments of the written functions of *well*, *like* and *oh yeah* ... 217
6.4 DM sequencing with *well*, *like* and *oh yeah* ... 221
6.5 DMs in relation to other colloquial phenomena ... 229
6.6 Why is there a surge of these DM in written functions starting in the 1990s? ... 235
6.7 Discussion ... 241
6.8 Final conclusions ... 242

References ... 245

Chapter 1:
Introduction

1.1 Introduction

The second half of the 20th century has been identified as a crucial period for changes in English writing across various genres (cf. Mair 2006; Leech *et al.* 2009). One specific area in this field of research that is starting to gain traction involves the increased use of dialogical inserts, interjections and discourse markers (DMs) in published texts, particularly in journalistic prose (cf. Rühlemann & Hilpert 2017; Tottie 2017). The study of such phenomenon in written genres has implications not only for genre analysis globally, but it also brings to light recent developments in the colloquialization of writing that past studies have largely overlooked. Through the analysis of how dialogical discourse markers (DMs) are used in journalistic texts, questions arise about our current understanding about the spoken – written language divide. For example, the changes that have occurred in 20th and early 21st century written English appear to be blurring the border between what we see as typically written and what is typically oral language. Furthermore, this ambiguity has only accelerated due in large part to the rapid shift from a print-based to digital based culture of writing (cf. Baron 2000; Crystal 2001; Thurlow & Mroczek 2011; Zappavigna, 2012; Barton & Lee 2013).

In addition to addressing changes in the overall writing style of English in the 20th and 21st centuries, this study also addresses issues that are pertinent to the study of DMs in general. For instance, this study investigates how DMs develop specifically written functions that are related to yet unique from their prototypical uses in spontaneous speech. Therefore this thesis analyzes not only the ability of DMs to adapt and evolve in specific text-types (cf. Aijmer 2013, 2015), but it also brings to light new ways in which DMs can develop new (inter)subjective functions that follow different paths of development than was previously observed in other studies with regard to their development along the (inter)subjective cline (cf. Traugott 2010).

Therefore, by bringing all of these interconnected disciplines together, this thesis charts new territory in the study of linguistic change and colloquialization. In particular, this thesis addresses how typically oral elements of language can be imported into writing at an increasing frequency. Furthermore, is also addresses how they can evolve and be re-interpreted and conventionalize new functions that are adapted to the new medium. In particular, this thesis investigates recent changes in how colloquial DMs (*well*, *like* and *oh yeah*) are used in journalistic writing. It demonstrates how these DMs are co-opted into writing, yet changed and adapted to the written modality. Furthermore, it will be shown that these DMs are used in journalistic prose in functions that are not present in spoken language, despite the fact that they share characteristics regarding how they are used in spontaneous speech.

Examples of these DMs in journalistic prose are provided in (1) - (4) below. In (1), *well* is used to highlight the author's alignment-seeking comment *you know* that coincidentally replaces *flowers*, which would have finished the idiomatic phrase *stop and smell the flowers*. This usage of *well* therefore highlights that the author intentionally addresses to the readership that he/she is about to make a joke involving this idiom, as if the decision to not say *flowers* was made at the last second:

(1) It's impossible not to feel cheery in the face of all the bright blooms, but I don't have time to stop and smell the, **well**, you know. I'm here for the food. (Hemispheres Magazine: Jan 2019[1])

Example (2) involves the DM *oh yeah*, which marks the end of a list of important people at the Cannes festival in 2019. In this example, the last person on the list is Quentin Tarantino, arguably the most famous movie director and producer named on this list. Though this usage

[1] This article in found in the January Issue of Hemispheres Magazine that is available online: <http://hemispheres.ink-live.com/html5/reader/production/default.aspx?pubname=&edid=52ffe258-7f1f-413b-b615-4a405c7572d1>

of *oh yeah* resembles a speaker using it to mark that he/she is just remembering something, it is obvious that Quentin Tarantino is not a person that the author almost forgot. This is even clearer when one reads the article in its entirety.

(2) Still to come are the Dardenne brothers, also in search of their third Palme, along with Ira Sachs, Bong Joon Ho, Xavier Dolan, Elia Suleiman — and, **oh yeah**, that Tarantino guy, who should be taking Cannes by storm on Tuesday. (TheWrap.com: May 2019[2])

Instead, in this case *oh yeah* is used rhetorically to end a list with the most important element appearing last. In fact, this is a highly conventionalized function of *oh yeah* that appears in journalistic writing at the very end of the 20th century, which will be discussed in detail in chapter 5. This function of *oh yeah* mimics the oral information management function of *oh yeah* that is used to mark the speaker remembering something. However, when used in writing, the function that it serves is the opposite. The author did not nearly forget something, instead *oh yeah* is used here to highlight what the author considers to be the most important piece of information.

Example (3) involves the DM *like.* In this example, *like* highlights the author elaborating on the *horror* of living in the fictional land of Westeros. In this example, the author repeats the usage of *horror* three times, with *like* marking the second and third repetitions. Normally in published texts, such explicit repetitions would not be considered good writing style. Therefore, such uses of *like* is reminiscent of the common oral hedging strategy of DMs which can be used to downplay dispreferreds and other face-saving functions.

[2] This article comes from an online entertainment news source:
< https://www.thewrap.com/cannes-halfway-point-how-did-elton-john-become-king-of-the-fest/>

(3) There's always the big picture to worry about, too. Considering after just 10 episodes they'll be back to needing more source material again, you can also think ahead by hiring an author that writes faster than any other novelist alive. Plus that guy can help capture the true horror of living in Westeros. **Like**, all the horror. **Like**, way too much horror. But remember, the only thing Game of Thrones seems to like more than death is its sexual content. So who knows how to write about sex for a large audience better than anyone else? (NOW: Nerdist 2016)

Besides marking the use of repetitions, *like* is also used in this extract to highlight the rhetorical importance of the use of the noun *horror,* which is clearly an aspect of the series and its setting that the author desires to elaborate upon.

As a final example, (4) illustrates that an author can use a series of DMs in sequences in order to express one's personal attitudes toward what they are discussing in their articles. In this particular case, the author is pointing out his/her cynical attitude towards the issue of how to measure success when it comes to medical research.

(4) WHICH LEAVES US WITH a final but very important question. When will the ESC promise pay off? When can we expect something more from them than arcane articles in medical journals (though repeatedly portrayed as miracle breakthroughs in the *New York Times*)? **You know, like, well**, actually making sick people better? (COHA Magazines: American Spectator: 2007)

All four of these examples demonstrate that these highly dialogical DMs are being used in journalism outside of dialogical contexts in ways that allow the author to display their subjective stance, or express intersubjective reader-oriented attitudes towards the information being highlighted. In examples (2) and (4), the authors are expressing their subjective attitude that relates to the importance or their personal stance towards what is being described. Example (1) and (3) also highlight the author's subjective stance, while also acknowledging the potential reactions of their readership. In (1) in particular, the author is making an alignment-seeking comment to the imagined readership in a way that looks similar to a dialogical function of *well* that marks a pause and rephrasing. In (3) the author highlights the fact the he/she is repeating

the same word *horror* multiple times in close proximity intentionally. Such uses of DMs in writing could be considered out of place, because they are used in contexts where there is no need to monitor what one is writing in real time. The editing process of monologue texts makes it possible to amend what one writes before publication. Furthermore, and most importantly, published texts do not allow the author to interact directly with the readership as one can in spontaneous oral discourse. Despite these differences between oral and written language, however, these DMs are used as rhetorical devices by the authors to inject into the text "a direct-speech-like immediacy" that imitates interaction between the author and reader (e.g. Rühlemann & Hilpert 2017: 131).

1.2 Colloquialization

One area of linguistic research that has already investigated changes toward a more colloquial-like writing style in 20th century English writing is the theory of colloquialization. When describing the colloquialization of written English since the second half of the 20th century, Mair (2006) and Leech *et al.* (2009) discuss changes occurring both at the macro-structural and micro-structural level of various genres of writing. How these features are used and how much they increase in usage varies by genre (Leech *et al.* 2009: 240). At the macro-structural level, for example, Mair (2006: 188) shows that the press progressively incorporates more quotations, real or fabricated, for the stylistic effect of making "the texts appear more dramatic, interesting, and accessible and, presumably, also to involve the reader emotionally". When discussing the changes at the micro-structural level, Mair demonstrates the growing preference for informal over formal options where both are available (*ibid.*: 188). Among the grammatical constructions spreading into writing Mair (2006: 188-9) and Leech *et al.* (2009: 239) show a rise in the use of verb and negation contractions, personal pronouns, zero relative clauses, progressives and questions. However, these changes are not occurring in all genres of written English equally. In fact, in their investigations of a variety of written genres, these changes are the most evident

in 'popular written registers' such as letters, fiction, and essays from the 19th and 20th centuries and in journalistic texts particularly from the second half of the 20th century (see also Westin & Geisler 2002).

These past investigations on the colloquialization of written English, share both similarities and differences with what is being observed with dialogical DMs in journalistic writing. For the similarities, all of these situations involve shifts in the writing style of journalism that include an increase in the usage of colloquial elements and constructions. However, these changes in how dialogical DMs are used in journalistic texts differ from previous studies of colloquialization because, unlike the other cases mentioned above, when DMs are used in non-dialogical contexts in journalistic writing, they also appear to be developing special functions that are particular to the needs of writing, as seen in (1) – (4). Therefore, the usage of these DMs in writing is not the same as increasingly using verbal or negative contractions or zero relative clauses like past studies have indicated. Instead, what (1) – (4) show is that colloquial elements can be adopted into writing in ways that are different from how they are used in speech. A major reason for this is likely the fact that monologue texts and spontaneous oral speech with participants can never become exactly the same. Therefore, when investigating how writing can become more speech-like, it is essential to keep in mind that though it is possible for writing to become more like speech, it will never be able to become the same as speech. This issue will be discussed in more detail in 1.3 as well as in the following chapters.

1.3 Speech and writing

As was evident in examples (1) – (4) above, the new written functions of the dialogical DMs in question are used in ways that resemble how they are used in oral speech, but perform rhetorical functions in writing that, in fact, are different from the oral functions that they resemble. For example, instead of performing the face-saving and/or word search functions that they

resemble, they are actually highlighting the author's intentional usage of special lexical items while also highlighting their rhetorical stance towards the use of said items. The usage of these DMs in non-dialogical contexts can perform a variety of functions such as indicating intentional word repetitions and listing the most important member of a list last. In example (4), the sequence of DMs highlights the author's subjective stance towards measuring success in medical research by using a common hedging strategy from oral speech that usually acts to downplay the dogmatic delivery of subjective stance (e.g. Beeching 2016: 132).

The fact that these DMs are developing new written functions that resemble their prototypical oral ones make it necessary to discuss what differentiates oral speech and written language. Crystal (2001, 2006) lists the key differences between speech and writing (see Table 1.1). He states that "(s)peech is typically time-bound, spontaneous, face-to-face, socially interactive, loosely structured, immediately revisable, and prosodically rich" while writing is "typically space-bound, contrived, visually decontextualized, factually communicative, elaborately structured, repeatedly revisable, and graphically rich" (2006: 20).

Speech	Writing
1. Time Bound	Space Bound
2. Spontaneous	Contrived
3. Face-to face	Visually decontextualized
4. Loosely structured	Elaborately structured
5. Socially interactive	Factually communicative
6. Immediately revisable	Repeatedly revisable
7. Prosodically rich	Graphically rich

Table 1.1 Differences between speech and writing (based on Crystal 2006: 291, 2001: 26-7)

Observing these supposed differences in the characteristics of spoken language and writing and comparing these dichotomies with the examples (1) – (4) above, it is clear that the modern journalistic texts from which the examples were found exist somewhere in-between the two extremes. Though these examples come from published journalistic texts, they obviously have some characteristics, particularly in how the DMs are being used, that are oral-like, yet adapted into the specific genre of writing. This is especially the case of characteristics 5 and 6 in Table 1.1. As demonstrated above, these DMs maintain the socially interactive characteristics of dialogical DMs in that they can signal attention given to the readership, while they also indicate their own opinion or subjective stance towards what they are discussing. Furthermore, they are also performing textual functions that resemble the immediately revisable characteristic (characteristic 6) of oral speech. Obviously, in writing, it is possible to revise a text without the need to signal it with a DM as it is in speech. Furthermore, it is also not possible for writers of journalistic texts to interact directly with their readership. Therefore, the fact that DMs are used in these newer and specially writing functions that mimic characteristics of oral speech indicates that there are developments going on in writing that are making it appear more interpersonal-like, while still maintaining characteristics that are uniquely written.

1.4 Dialogical DMs in Writing

As shown in examples (1) – (4), *well*, *like* and *oh yeah* are used in journalistic prose rhetorically in order to take up a subjective stance while also seeking (intersubjective) alignment with the readership. These three particular DMs are already proven to be usable in a variety of stance taking functions (Aijmer 2013: 15), though it is notable that their usage in journalistic prose is doing something different than they do in spontaneous speech since these texts are monologues as opposed to discourse with two or more participants.

The concept of *stance* is a theoretical element that is essential in this discussion of dialogical DMs in general and the use of these dialogical DMs in written texts in particular. Du Bois's influential paper defines stance as follows[3]:

> Stance is a public act by a social actor, achieved dialogically through overt communicative means, of simultaneously evaluating objects, positioning subjects (self and others), and aligning with other subjects, with respect to any salient dimension of the sociocultural field. (2007: 163)

Du Bois also defines stance from a first-person view with the following statement: "I evaluate something, and thereby position myself, and thereby align with you" (Du Bois 2007: 163). This is further exemplified by 'the stance triangle', which is found in Figure 1.1.

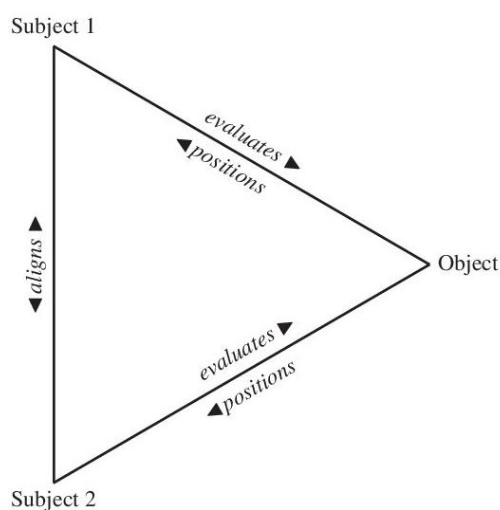


Figure 1.1 Du Bois's stance triangle (2007: 163)

In discussing the 'ultimate import of stance', Du Bois emphasizes that stance is comprised of "three key aspects of social life: act, responsibility and value" (2007: 173). The act is the public enactment of stance and its, "force can be measured by its effective penetration into virtually all domains of sociocultural life" (*ibid.*) Responsibility is the ownership of stance that is taken by the speaker. For example, according to Du Bois, "if you take it you own it," and the question of who took what stance is salient and remembered over time (*ibid.*). Finally, value "is what

[3] For a discussion of other similar approaches to stance (or modality) see Halliday 1994; Kress 1995 and Fairclough 2003.

stance is all about - literally", and it invokes directly or indirectly presupposed systems of sociocultural value, while simultaneously enacting and reproducing these systems (*ibid.*).

When looking specifically at how DMs are used to express stance, Aijmer (2013: 15) shows that DMs, "are capable of expressing a number of new stances in interaction". This includes a variety of cases that express affective and epistemic stances.[4] Moreover, they can have rhetorical functions that enable the speaker to align or disalign themselves to their collocutor and what they say. *Well*, for instance, is often used to hedge a negative or irrelevant response. DMs are also used to express their social identity/statuses (e.g. also authoritative or knowledgeable stance) (*ibid.*).[5] For a particular example, Aijmer demonstrates that a prosecutor can use *well* in court examinations to display authoritative stance, while the use of *well* by a witness is associated with deferential attitude (*ibid.*). Finally, she also states that DMs are used to express affective stance. For instance, she gives the example that *well* can be used to signal disappointment, resignation or reluctance (*ibid.*).

Taking these previous considerations of DMs and the concept of stance into account, this thesis will argue that the use of dialogical DMs in journalistic prose, as we have seen in examples (1) – (4), is a salient and 'interpersonal-like' tool used by authors to express their subjective and intersubjective stance toward specific information presented in their texts. How this is done precisely will be shown in the individual case study chapters 3-5. Moreover, the development of these DM functions points to the usage of colloquial elements of language in writing that are used differently than they are in typical speech. This could be seen as written language that is, "pulled some way in the direction of speech" (Crystal 2001: 47), which can be opposed to the

---

[4] In general, affective stances represent emotional states of a speaker, while epistemic stances convey the degrees of certainty about their propositions. Both affective stance and epistemic stance are culturally grounded. (Jaffe 2009:7).

[5] See Ochs (1996) for a discussion on the link between stance and the social identity, which is conveyed by cultural expectations

notion that certain genres of writing are simply turning into speech that is written down. Ironically, when Crystal talks about written language that is pulled in the direction of speech, he is, in fact, talking about the development of internet writing, which has developed rapidly since the 1990s, and not published texts. However, these parallels between changes occurring in journalistic texts and the development of online writing will be further explored throughout this thesis, but particularly in chapter 6. These observations also indicate that the distinction between the medium of speech and writing is considerably more complex than many theoretical frameworks claim. Furthermore, in terms of writing, it is important to consider that different genres of writing serve various purposes at different levels of (in)formality. This is certainly why previous studies on colloquialization have observed stronger changes in the popular written registers of writing as well as journalism and weaker pro-colloquialization shifts in others, such as academic writing.

1.5 Research questions

Based on the observations made above concerning examples (1) – (4) this thesis aims to answer the following questions.

a) Which dialogical DMs have entered the repertoire of journalistic prose?

Using corpus data that includes the journalism sub-corpora of the Corpus of Historical American English (COHA) and the News on the Web Corpus (NOW), this thesis identifies multiple DMs that have entered into the repertoire of journalistic prose. The reasoning for the selection of these corpora is provided in the next chapter in section 2.6. A key aspect beyond the mere identification of such DMs that will be investigated is what type of genre specific or ‘text-type’ functions (cf. Aijmer 2013, 2015) develop in journalistic prose. Furthermore, I will address how these written uses are similar to their prototypical spontaneous speech functions as well as how are they different.

b) What do these uses of dialogical DMs in journalistic prose tell us about the change and development of DMs in general?

One aspect of DMs that makes them so interesting as a research topic, and also so difficult to define, is their multifunctionality. Furthermore, Biber *et al.* (1999: 1086) point out that the, "(w)ords and phrases which are discourse markers are often ambiguous, sharing the discourse marker function with an adverbial function". This is certainly true with the three DMs discussed in (1) – (4) above, though it is notable that *well* and *like* can be more than DMs and adverbials. The particularities of each DM investigated in this thesis will be discussed in detail in each of their respective chapters. One process that is particularly important for the development of DMs from adverbials, however, is the process of (inter)subjectification. For instance, Traugott & Dasher (2002: 174) investigated how certain adverbials developed into DMs with textual functions and then, these textual DMs developed meanings that express the speaker's attitude or viewpoint (subjectivity). From this point, they were then shown to develop meanings that express the speaker's attention to the addressee's self-image (intersubjectivity) (*ibid.*)**.** As the example (1) – (4) were highly subjective as well as intersubjective in the sense that they were used in alignment seeking functions, it will be essential to address the process(es) in which these written functions developed. Furthermore, the fact that DMs have already been shown to develop specialized 'text-type' functions based on the type of discourse will also undoubtedly play a role in how these DMs transfer into journalistic prose.

c) Do these changes take place over a particular period of time?

Past investigations into colloquialization have found that many of the shifts toward a more colloquial writing style have occurred over the course of the second half of the 20th century in particular. The reasons for these developments during this period are due to a series of factors, though some of the major pressures of this change are due to, "a current of informalisation and (pseudo-)democratization affecting advanced industrial societies" which includes, naturally,

the United States and Great Britain (Mair 1997: 198). Though colloquialization has been addressed in a series of corpus studies and have identified a variety of phenomena that constitute pro colloquilization changes from second half of the 20th century (Hundt & Mair 1999; Mair 2006; Leech *et al*. 2009; Leech & Smith 2009; Rühlemann & Hilpert 2017; Tottie 2017), none to my knowledge have directly addressed whether or not there have been any further drastic shifts later on. Therefore, this thesis will address whether the shifts from a traditional print-based journalistic writing culture to a print and digital or purely digital writing culture has had any significant effect on the colloquialization of journalism from the end of the 20th century to today. Naturally, this question will be addressed through observing the changes that have occurred in terms of using dialogical DMs in journalistic writing.

d) What are the implications for the study of colloquialization as a whole?

In addressing the timing of the developments of the written DM functions, this thesis will further address other implications these changes have for the study of colloquialization. For example, how do the developments of dialogical DMs in written texts relate to other changes that have been observed in relation to colloquialization?

To respond to these questions, this study conducts a diachronic investigation of three dialogical DMs using a Usage-Based Corpus Pragmatics approach. The details pertaining to how these DMs will be investigated using corpus data will be discussed in detail in Chapter 2 sections 2.5 - 2.7.

1.6 Conclusions

In this introduction, I have discussed how the manner in which DMs are used in journalistic writing bring up many questions about how DMs develop genre or text-type specific functions. Such observations display yet another way in which journalistic writing is becoming

increasingly more colloquial-like, yet it also shows that this process of colloquialization may be more complicated than past studies have shown. As examples (1)-(4) indicate, DMs can be used for different purposes than would have occurred in speech. When these changes occurred and potential reasons for these changes will be investigated in this thesis. Chapter 2 addresses the theoretical issues regarding DMs, their ability to develop text-type specific functions as well as the shifts occurring in writing since the 20th century in much more detail. Furthermore, it discusses the corpora used to extract data, and it goes over the quantitative and qualitative methodologies adopted in this thesis to analyze them. Chapter 3 investigates *well*, the only DM of the three to my knowledge to have received any focused attention in the realm of journalistic writing thus far. This is followed by the case study of *like* in chapter 4. Chapter 5 analyzes *oh yeah*, a DM which has hitherto received little attention as a DM as such. Chapter 6 discusses what the results of the three case indicate about the changes to these DMs and how they contribute to further changes in journalistic writing. Furthermore, this chapter discusses why this new wave of change occurs when it does. The end of chapter 6 concludes this thesis.

Chapter 2:
Theoretical and methodological background

2.1 Introduction

In the previous chapter, the focus of this study was explained in general terms. This chapter focuses on the theoretical and methodological aspects of this work. It begins with section 2.2 which provides a definition for DMs and explains their major functions. Section 2.3 continues with a discussion of some of the most influential theoretical approaches to understanding DMs for the purposes of this study. Section 2.4 reviews past studies on the shifts occurring in certain genres of written texts. Section 2.5 explains which DMs were selected for the case studies that constitute the main body of this thesis. Section 2.6 discusses the two corpora that are used to extract data of DMs in journalistic texts. Section 2.7 explains some of the statistical tools used to analyze the diachronic frequency trends of these DMs in journalism and other genres of writing for the sake of comparison. Finally, section 2.8 offers a summary of the results and conclusions.

2.2 Discourse Markers: Terminology and Functions

Though there has been a considerable amount of work on DMs over the past few decades, the scope of inquiry has been primarily limited to oral data. Furthermore, much controversy remains over what constitutes a DM and how to describe their functions. For example, even though the term 'discourse marker' is not universally agreed upon by researchers of this field, the term has over the years gained substantial ground in being accepted as the "most popular of a host of competing terms used with partially overlapping reference" (Schourup 1999: 228). Other common terms used include 'discourse particles' (e.g. Schourup 1985, Aijmer 2002), 'connective' (e.g. Fraser 1988), and 'pragmatic marker' (e.g. Watts 1988, Brinton 1996, and Aijmer 2013). It is also necessary to point out that there is no complete agreement over what the properties of these DMs are. This issue of pinpointing what specifically designates a DM

as a DM is an unresolved debate that has persisted in this field of study since its inception. There are certain linguistic items that are generally agreed upon by most to constitute DMs, especially those that combine textual and interpersonal functions such as *well*, *just*, *like, oh (yeah)* and *you know*. On the other hand, matters become more contested when considering other items such as *or* and *but* that are considered to be more marginal members (see discussion in Biber *et al*. 1999: 1086 and Müller 2005: 3). This difficulty in distinguishing what fits into the category of DM outside of a certain group of prototypical ones demonstrates that as a conceptual category DMs remain fuzzy such that certain prototypes exist along with various marginal members.[6] When discussing what makes a DM a DM, however, researchers often give the following list of features, which can be used to identify and rate DMs on the basis of how well they fit into these criteria (Jucker & Ziv 1998: 4). Perhaps one key reason why there is such divergence over what the properties of DMs are is that few of these features are considered to be defining features of DMs. Therefore, non-compliance with one of them can rarely be the basis for excluding a linguistic item from the group of DMs (Müller 2005: 4). These features (a)-(f) are listed below.

a) No single word class

DMs are composed of a variety of words and word combinations from a variety of word classes. Not only is this one of the unique features about DMs, but this has also been the cause of much disagreement among researchers as to how to come up with a complete list of them (Müller 2005: 5). In their investigation of adverbials that develop discourse-marking functions, Traugott & Dasher (2002) list a collection of common DMs that developed from adverbials. This includes *in fact*, *indeed*, *actually* and *well*. Biber *et al.* (1999: 1086) also lists adverbials that developed DM functions. One of these adverbials is *well,* while they also show interactive uses of *right* and *now.* In other cases, interjections can develop textual and interpersonal DM

[6] For a discussion on Prototype theory, see Rosch (1978) and Taylor (2001).

functions, as is the case with *oh*. In addition to single word DMs, there are also finite verb formulae such as *I mean*, *you know* and *you see* that have developed into DMs as well (Biber *et al.* 1999: 1086). Another factor that further illustrates the difficulty of pinning down what types of word(s) can become DMs is the fact that many of the words that develop DM functions are ubiquitous, and therefore can operate in a variety of roles. *Well*, for example, can also be a noun (such as a water well) and adjective in addition to an adverb and DM. Perhaps the most ubiquitous of all is *like*, which can be a verb, noun, preposition, conjunction and adverb, in addition to serving a number of DM functions (Romaine & Lange 1991; D'Arcy 2006). Furthermore, in addition to the variety of word classes that give birth to DMs, things are further complicated by the fact that individual DMs often serve a number of varying functions. This will be further explained in (f).

b) Syntactic position

Brinton (1996: 33) states that, "it is often said that pragmatic markers are restricted to sentence-initial position, or may always occur sentence-initially", though she also points out that there are many studies that show them appearing sentence-medially and finally as well. Such examples can be seen in (1) and (2) below:

(1) The boss and the secretary work late all night, ***well***, not all night but late in the night (taken from Biber *et al.* 1999: 1087).

(2) A: Does Does shouting a lot and making lots of noise really have much effect uh in terms of intercession or

C: **Well** <,> I think it does actually
Yeah I think I think equally you can do it quietly (taken from Aijmer 2013: 115)

Furthermore, Schiffrin (1987), Fraser (1990) and Schourup (1999: 233) make the distinction that DMs tend to occur not necessarily at the beginning of a sentence, but rather the beginning of an utterance or turn. Schourup (1999: 233) explains this by stating that, "(t)he tendency of

DMs to appear initially is probably related to their 'superordinate' use to restrict contextual interpretation of an utterance". In this study in particular, it will be shown that this preference to view all DMs being characteristically sentence or utterance-initial is problematic. Furthermore, such a view of DMs overlooks many of the major developments going on with the three dialogical DMs in writing.

c) Syntactic independence and grammatical optionality

Fraser (1988:22) makes the claim that, "the absence of the discourse marker does not render a sentence ungrammatical and/or unintelligible". Furthermore, Schourup (1999: 231) also discusses how they are optional in the sense that they do not enlarge the semantic possibilities for the semantic relationship between the elements that they associate. In other words, according to Schourup (1999: 231), "if a DM is omitted, the relationship it signals is still available to the hearer, though no longer explicitly cued." However, while acknowledging that DMs are grammatically optional, Brinton (1996: 35-6) also specifies that:

> (w)hile pragmatic markers are grammatically optional and semantically empty, they are not pragmatically optional or superfluous: they serve a variety of pragmatic functions. If such markers are omitted the discourse is grammatically acceptable, but would be judged "unnatural", "awkward", "disjoined", "impolite", "unfriendly", or "dogmatic" within the communicative context. Creating such a discourse could be "incomprehensible" for the listener, and "mission impossible" for the speaker (Svartvik-Stenström 1985: 352), and there would be a greater chance of communicative breakdown (Fraser 1990: 390). Hence, pragmatic markers "help both encoder and decoder navigate alone a specific discourse" (Even-Zohar 1982:180).

Therefore, while DMs could typically be removed from a sentence without rendering it non-grammatical, they do indeed add much to discourse on the pragmatic level, especially in dealing with the interpersonal contexts of speech.

d) Lack of semantic content

When surveying the literature on DMs, it is very common to come across declarations stating that DMs have little or a vague meaning (e.g. Schiffrin 1987: 328) or containing at the very

least only “a residue of semantic meaning” (Ariel 1994: 3251). When one speaks of a ‘lack’ of semantic content in this case, however, it is not meant to say that there is a complete absence of any meaning at all. For example, Schiffrin investigates this very issue and demonstrates that some DMs are used in ways that are not clearly related to a semantic meaning (for instance *oh* and *well*), while others, like conjunctions, have pragmatic effects that are closely tied to their semantic content (1987: 317). Ariel (1994: 3251) expands upon this by distinguishing between transparent markers, intermediate markers, and opaque markers. The primary discourse functions of transparent markers are clearly tied to their semantic meaning – notable examples are *and*, *but*, *or*, *so*, and *because*. Intermediate markers – *now*, *then*, *you know* and *I mean* – are markers whose discourse functions can be related to their semantic meaning but at the same time are not necessarily primarily connected to them. The main discourse functions of opaque markers are (virtually) not semantic at all – examples given include *well*, *oh*, and *uh*. Though there is no agreement among researchers with regard to the extent of this continuum of DM types, few deny that there are indeed differences in the transparency/opaqueness of marker meaning (Müller 2005: 63).

e) Orality

It has often been claimed that DMs occur primarily in speech. For example, Watts (1989: 208) goes as far as to claim that DMs are “one of the most perceptually salient features of oral style”. Such ideas also find support in Schiffrin’s (1987: 326) definition of DMs as “contextual coordinates for utterances”. Though this definition does not explicitly list orality as the defining feature, the nature of Schiffrin’s approach clearly understands DMs in terms of spontaneous speech much more so than in terms of written language. Furthermore, this case for the association of DMs with orality has been claimed to cause them to be “stylistically stigmatized and negatively evaluated, especially in written or formal discourse” (Brinton 1996: 33). However, this claim that DMs are primarily associated with speech may have more to do with

the preference of past studies to do research on typically oral DMs and for researchers to investigate DMs using speech data (see Schourup 1999: 234)[7]. As this entire thesis discusses how dialogical DM are used in journalistic prose, the position taken in this thesis is that this idea that DMs are partial to orality is problematic, especially when one considers that there are DMs that are well established in writing such as *although, after* and *in addition* (cf. Hutchinson 2003). However, the fact that the most prototypical DMs combine both textual functions and interpersonal ones does provide evidence that many DMs are at least more oral-centric than written-centric since the combination of these roles would be more useful in oral discourse than written.

f) Multifunctionality

DMs have been shown to fulfill more than one function as well as having sub-functions. For example, Aijmer (2002: 3) states that, "(d)iscourse particles are different from ordinary words in the language because of a large number of pragmatic values that they can be associated with". However, despite this general observation, some DMs are less multifunctional than others. Furthermore, Jucker & Ziv (1998:4) point out that some DMs could be argued to be monofunctional, which makes it problematic to use multifunctionality as a defining criterion. Putting this issue to the side, however, most research on DMs in English has indeed found that they have more than one function or at least have sub-functions (Müller 2005: 8). Furthermore, for the purposes of this study, all of the DMs involved not only serve multiple functions in speech, but they also develop new functions in journalistic prose that are (inter)subjective adaptations of specific oral DM constructions.

This discussion of these six major properties of DMs display some of the criteria that linguists use to identify them. Though these criteria are useful for identifying some of the most typical

[7] For exceptions to this see Aijmer (2013), Rühlemann & Hilpert (2017) and Tottie (2017).

DMs, as already stated, none of them can be attributed to all DMs as a defining feature. Moreover, the above-mentioned features appear to be geared more towards identifying DMs that are stereotypical to informal spontaneous discourse as opposed to those that are geared towards formal speech or writing. Moreover, these properties say little of what DMs do. In light of this, this thesis investigates DMs from more of a functional approach. For example, in the literature on DMs, two major functions are often identified. Biber *et al.* (1999: 1086), for instance, describe DMs as:

> inserts which tend to occur at the beginning of a turn or utterance and combine two roles: (a) to signal a transition in the evolving progress of the conversation, and (b) to signal an interactive relationship between speaker, hearer, and message.

Well is given as an example of a 'versatile discourse marker' that has the general function of a 'deliberation signal' that indicates that the speaker needs to give a "(brief) thought or consideration to the point at issue" (*ibid.*: 1086). One example that they give involving *well* that combines roles (a) and (b) is listed below in (3). In this example *well* is used as a turn initiator (continuation marker), and it is also used in an interactive co-function that marks contrast (often a disagreement or prevaricating answers to questions).

(3) A: You never say what Stafford is, why is he going before the <unclear> parole board? You never say.
B: **Well**, I don't know why
A: **Well**, It's a badly written story.

Aijmer (2013), in a similar vein, defines DMs as manifesting both reflexivity and operating as contextualization cues. When discussing reflexivity, Aijmer (2013: 4) states that it is "manifested as the speaker's awareness of the linguistic choices made both with regard to what to say and how to say it" (see also Verschueren 1999: 187). Similarly, Redeker (2006) discusses how DMs mirror the speaker's mental processes, thus commenting on what is going on in the speaker's mind while interacting. When it comes to DMs having reflexive functions in

particular, they are argued to function as indicators of metapragmatic awareness along with other features such as prosody, hesitation and pausing (Aijmer 2013*:* 5). This can be observed in the reflexive uses of *well* that 'remark on' the planning taking place in the speaker's mind, as well as when it is used to accompany reformulations or revisions in ongoing discourse. See (4) below:

(4) A: One's about the human brain and language. And the other's about uh this guy called Chomsky who's uh <,,> **well** one of the world's most important human beings if you happen to be interested in linguistics (taken from Aijmer 2013:32).

As contextualization cues, on the other hand, DMs make explicit different segments of the discourse. This, therefore, helps the hearer to understand how the stream of talk is organized (Aijmer 2013: 6). For example, Aijmer gives the following example (5) from a broadcast discussion where *well* has the function of signaling the introduction of a new speaker to a debate:

(5) A: Melvyn Bragg you're President of the National Campaign for the Arts the lead signatory in the letter part of which I quoted a few moments ago. What do you think's gone wrong <,>
B: **Well** before we start to talk about finances which'll occupy a lot of this programme and blame which'll occupy a lot of this programme the reason why I'm here and people are watching is because most people think that the arts add something to their lives that nothing else will give them (taken from Aijmer 2013: 7)

Gumperz (1996: 379) defines a contextualization cue as, "one of a cluster of indexical signs […] produced in the act of speaking that jointly index, that is, invoke a frame of interpretation for the rest of the linguistic content of the utterance". Though Gumperz restricts his discussion of contextualization cues to elements such as prosody, gestures or shifts (code switches or style shifts); other linguists have already expanded the definition to include DMs. One example is De Fina (1997), who investigates Spanish *bien* (the equivalent of English *well*). This inclusion of DMs into what Gumperz calls contextualization cues is relevant because DMs, like the other

elements discussed by Gumperz, also "represent speakers' ways of signaling and providing information to interlocutors and audience about how language is being used at any point in the ongoing stream of talk" (1996: 366). When looking at DMs as contextualization cues, they have been found to serve a variety of functions such as being used to signal the transition to a new topic, activity, argument, stage in a narrative, a new speaker in a debate and so on. Moreover, studies of DMs in different genres of discourse have found that the contextualization cues of particular DMs can differ depending on the "expectations associated with the particular speech activity where it is used" (Aijmer 2013: 7). For example, in De Fina's (1997) analysis of *bien* in classroom dialogue, she finds that the teacher, who has the privileged role in the 'teaching frame', uses *bien* in specialized 'slots' that mark the transition from one classroom activity to another. Furthermore, when *well* is used in oral narratives, Norrick (2001: 857) finds that it is used as a contextualization cue that acts as a specialized narrative DM where it is keyed on expectations about the organization of stories. In this case, *well* is used to lead to beginning and endings of narratives as well as mark a return to the main theme of the story.

When looking at the individual functions of specific DMs, it is important to make the distinction between textual and interpersonal functions. Müller (2005: 30-31) states that the textual DM functions "remain focused on lexical expressions and propositional content in units of various length, from single words or phrases to a sequence of utterances describing a particular scene". When discussing the textual uses of DMs broadly, she states that a DM can be used to "indicate a search for a phrase which expresses what the speaker has in mind" and it is used to mark false starts or repairs (*ibid.*: 30-31). In her analysis of *well*, for instance, Müller (2005: 107) lists textual level functions that include searching for the right phrase, rephrasing/correcting, quotative *well*, etc.

Interpersonal functions of DMs on the other hand “interactionally focus on the relationship between speaker and hearer. They are typically used to mark “a speech act, a response, an opinion, or an evaluation” (Müller 2005: 31). Coming back to Müller’s investigation of *well*, she found that the interpersonal functions include well being used to mark indirect answers, direct answers, response to self-raised questions, contributing an opining, etc. This clear distinction between the textual and interpersonal functions of DMs, however, gives a somewhat oversimplified representation of how DMs operate in action. For example, Beeching (2016: 6) states that:

For example, Beeching (2016: 6) states that:

> Pragmatic markers are notoriously both polysemous and multifunctional. In other words, not only does each pragmatic marker fulfil a variety of functions, it can do so simultaneously. This poses problems of interpretation.

Furthermore, this inter-relation between the textual and interpersonal functions is something that will prove to be even more relevant to the discussion of DMs in journalistic written texts, despite the fact that written texts have little need of many of the prototypical DM textual functions such as word-searching, rephrasing or self-correcting. What appears to be occurring in journalistic texts is the usage of DMs in functions that use the conventions of the textual functions (thus adding elements of spontaneity to the texts) in a fashion that serves the co-function of marking the author’s subjective and/or intersubjective stance.

Another property that is essential to explaining the role of DMs in communication is indexicality. Much discussion has already been made about the textual and interpersonal functions of DMs, but it is also important to stress how DMs are used to index the speaker’s “social or professional identity, social relations and activities” (Aijmer 2015: 199). Ochs (1996), for example, identifies three socio-cultural dimensions of the communicative situation besides time or place:

> *Social identity* encompasses all dimensions of social personae, including roles (e.g. speaker, overhearer, master of ceremonies, doctor, teacher), relationships (e.g. kinship, occupational, friendship, recreational relations), group identity (gender, generation, class, ethnic, religious, educational group membership), and rank (titled and untitled persons, employer, employee), among other properties, and other properties;
> *Social act* refers to a socially recognized goal-directed behavior (e.g. request, an offer, a complement:
> *Activity* refers to a sequence of at least two social acts, e.g. disputing, storytelling interviewing, giving advice;
> *Affective stance* refers to a mood, attitude, feeling, and disposition, as well as degrees of emotion intensity vis-à-vis some focus of concern;
> *Epistemic stance* refers to knowledge or belief vis-à-vis some focus of concern, including degrees of certainty of knowledge, degrees of commitment to truth or propositions, and sources of knowledge, among other epistemic qualities (Ochs 1996: 410)

A more recent look at indexicality is Silverstein (2003), which identifies different layers of *indexical order*. The first order indexical order is identified as the first level of pragmatic meaning that is drawn from an utterance. In such cases, in the words of Beeching (2016: 15), "a demonstrable correlation is revealed between a linguistic usage and a sector of the population (which is, however, not overtly comment upon)". Second order indexicality, is concerned with the awareness that a particular linguistic phenomenon is associated with a particular group(s) (Silverstein 2003). In other words, it is concerned with the relation between linguistic variables and the metapragmatic means that are encoded in them. To explain third order (or higher order) idexicality, Silverstein (2003) discusses what he calls 'wine talk', which is used by people of a certain social group (yuppies). This wine talk has it origin from professional wine critics who use 'technical vocabulary' that is "used in prestige realms of traditional English gentlemanly horticulture and especially animal husbandry" (2003: 225). Therefore, a certain 'lingo' is created for wine that indexically entails notions of prestigious social classes or subjects. Therefore, according to Silverstein, when the yuppies use this lingo in the context of drinking wine, they are indexing their social standing as "the well-bred, characterologically interesting (subtle, balanced, intriguing, winning, etc.) person iconically corresponding to the metaphorical "fashion of speaking" of the 'higher-order' classes that those in yuppiedom aspire to be (2003: 226). In simpler terms, Beeching explains third order indexicality as a situation "where a

creative use is made of the association to evoke a particular effect" (2016: 15). Table 2.1, which is adapted from Beeching's (2016: 16) table that attempts to apply Silverstein's orders of indexicality to DMs, demonstrates how DMs develop and come to take on social meanings that index the speaker and the hearer, social identities and the speech event (or social activity) itself.

Silverstein (2003) Indexical order	**1st order Indexicality**	**2nd order Indexicality**	**3rd order Indexicality**
Overview	DM in general are used (only) in informal discourse	They come to be heard as speaking casually, informally, between friends, where meaning is implicit rather than explicit	Which comes to index, solidarity, positive politeness, 'buddyness'
Like	*Like* used in informal situations by young people	This comes to be heard as speaking casually	Which comes to index 'young (female) people'
Sort of	*sort of* is used to hesitate, express tentativeness, approximation	This comes to be heard as speaking casually, downplaying expertism	Which to index 'self-deprecating/nuancing' personality
You know	*you know* is used amongst intimates	This comes to be heard as 'sharing common knowledge'	Which comes to index 'we are intimates who share background information'

Table 2.1 Layering of Indexicality (adapted from Beeching 2016: 16; see also Johnstone 2010: 32)

By indexing these different levels of the context of speech, DMs are therefore able, in Aijmer's words, "to express a number of new stances in interaction (2013: 15). This is particularly relevant to the discussion of DMs in this study, because it also enables DMs to be used rhetorically thus "allowing the speaker to take up a stance of alignment or disalignment to the hearer or to what is said" (Aijmer 2013: 15). In the case of the four examples discussed in the introduction chapter where dialogical DM were used in journalistic prose, the DMs in question were utilized in ways that resembled how they are used in spontaneous speech. However, the DMs were used in a rhetorical manner in order to indicate their subjective stance (reservation or irony) while also using said strategy to get the reader to align with this stance. Taking advantage of these oral textual functions of DMs to mark one's subjective stance is operating

at the 2nd order of indexicality, while using these DMs as a strategy to attempt to achieve alignment with his/her readers operates at the 3rd order of indexicality.

Taking all of these theoretical aspects of DMs into account the big questions that must be accounted for in this study in terms of the use of dialogical DMs in writing are:

a.) In what ways are the use of dialogical DMs in journalistic prose similar to and different from how they are used in spontaneous speech?

b.) How are they able to reconcile their typical speech dialogue oriented textual and interpersonal functions with a medium that is both monological and contrived?

c.) When used in journalistic writing, in what ways are these DMs expressing stance? Are they taking up alignment or disalignment stances with their readership? Are they using them to express authority over their readership or are they expressing 'closeness'?

2.3 Theories that account for the development of DMs and their contextual relations

2.3.1 (Inter)subjectification

In the introduction examples, (1) – (4) displayed how some DMs such as *well*, *like*, *oh yeah* and even *you know* can be used in writing (inter)subjectively. While subjectivity has been recognized in semantic study as early as Bréal ([1900] 1964), it was not brought to the forefront of linguistic study until Benveniste's paper ([1958] 1971) distinguished subjectivity from intersubjectivity. (Inter)subjectification refers to the diachronic process of semanticization (Traugott 2010): 30). Subjectivity, "roughly covers the fact that a particular element or construction requires reference to the speaker" (De Smet & Verstraete 2006: 365). Intersubjectivity, on the other hand, is the "explicit expression on the SP/W's attention to the 'self' of addressee/reader in both an epistemic sense (paying attention to their presumed attitudes to the content of what is said), and in a more social sense (paying attention to their 'face' or 'image needs' associated with social stance and identity" (Traugott 2003: 128). In the past few decades, research on (inter)subjectivity and (inter)subjectification has been dominated

by the frameworks developed by Traugott or Langacker (Cuyckens *et al.* 2010; López-Couso 2010). The Traugottian approach is based on pragmatics and the historical analysis of semantic change in context (Narrog 2010: 386). In this approach, subjectification and intersubjectification are the mechanisms by which:

a. Meanings are recruited by the speaker to encode and regulate attitudes and beliefs (subjectification), and,
b. Once subjectified, may be recruited to encode meanings centred on the addressee (intersubjectification).

The Langackerian approach, on the other hand, "aims at being maximally precise within a specific modal of grammatical description, viz. Cognitive Grammar", (Narrog 2010: 386) and studies subjectivity from a synchronic perspective. Langacker characterized subjectification initially as "the realignment of some relationship from the objective to the subjective axis (Langacker 1990: 17). In later work, however, he redefines it in terms of semantic bleaching or attenuation where the "subjective component is there all along, being *immanent* in the objective conception, and simply remains behind when the latter fades away" (1999: 151, original emphasis). Therefore, according to Langacker (2008: 77), maximum subjectivity is construed when the speaker functions as a tacit conceptualization presence that is not made explicit. "At the opposite extreme, construed with maximal objectivity, is the focused object of attention: the entity an expression puts onstage and profiles" (*ibid.*). When comparing these two different approaches to subjectivity, Breban (2006:272) states that "for Traugott, the activity of the speaker is to 'express' (and in the second place to organize what he/she expresses), whereas for Langacker, the subject 'construes'". Therefore, the issue at hand for Langacker is not whether the word or construction is speaker related or not, but how explicitly the reference to the speaker is in the utterance. De Smet & Verstraete (2006: 370) show the relation between the two approaches in Figure 2.1 below where the two axes represent the Traugottian and Langackerian views of subjectivity. For the Traugottian axis, subjectivity requires a reference to the speaker.

For the Langackerian axis, subjectivity is based on whether the speaker explicitly figures in the form of the word or construction.

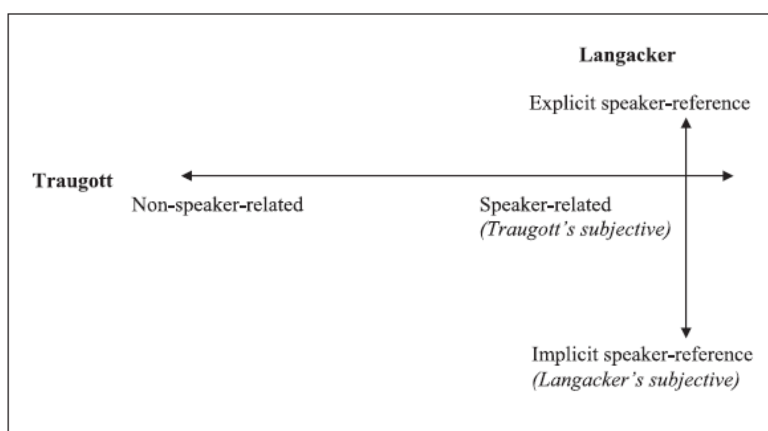


Figure 2.1 Langacker's vs. Traugott's definition of subjectivity (taken from De Smet & Verstraete 2006: 370)

The approach adopted in this thesis will be Traugottian for a number of reasons. First, Traugott's theory is based on empirical evidence, such as historical facts and diachronic language data. Langacker's approach, on the other hand, is a purely conceptual perspective on subjectification (cf. Narrog 2010: 392). Furthermore, this work, much like those that adhere to the Traugottian approach to (inter)subjectification, analyzes elements of language diachronically using historical corpus data. Another reason is that while Langacker's work discusses subjectivity at length, it does not sufficiently address intersubjectivity, which is an important aspect of this thesis. The Traugottian approach, on the other hand, investigates intersubjectification at length, which, according to this framework, follows and arises out of subjectification (Traugott 2010). Finally, there already exists a vast range of studies on the (inter)subjective developments of DMs that follow the Traugottian framework and that already shed much light on what has been happening with the development of a variety of DMs from a variety of languages (cf. López-Couso 2010: 136).

When looking at the diachronic development of (inter)subjectivity, the general trend is that the direction of semantic change proceeds from non-/less-subjective to (inter)subjective. The elements of language that go through this change are what De Smet & Verstraete (2006: 385) would call 'semantic subjectivity'. With semantic subjectivity, the speaker's relatedness is part "of the expression's inherent meaning" (2006:385). Semantic subjectivity is distinguished from the general notion of subjectivity, or 'pragmatic subjectivity', which "is inherent in language use and is independent of the semantics of a particular expression" (2006: 384) According to De Smet & Verstraete, semantic subjectivity can be broken down into two parts. One is 'ideational subjectivity', which "involves the description of a content internal to the speaker (2006: 385). Typical elements of ideational subjectivity are evaluative adjectives and nouns such as *stupid* or *pig*. The other type is 'interpersonal subjectivity', which refers to expressions that enact speaker positions with respect to a particular content. (2006: 387). Examples of interpersonal subjectivity include deontic and epistemic auxiliaries, discourse markers, illocutionary speech acts, modal adverbs and intensifying adjectives. As this particular study focuses on the use of dialogical DMs in recently developed written functions, the remainder of this discussion will focus on (inter)subjectivity in this area.

Traugott & Dasher (2002: 174-6) discuss the (inter)subjectification of DMs including *in fact*, *actually* and *well*, which evolved from adverbials only to DMs with textual functions. These textual DMs then developed meanings that express the speaker's attitude or viewpoint (subjectivity). From there, they also developed meanings that express the speaker's attention to the addressee's self-image (intersubjectivity) as hedge functions that "soften or mitigate what is said with the purpose of acknowledging the addressee's actual or possible objections" (2002:174). Therefore, this process moves along a cline that ranges from non/less subjective at the left to intersubjective to the right, as shown in (5) below.

(5) non-/less subjective > subjective > intersubjective

In the case of *well*, for example, Traugott & Dasher (2002) demonstrate that DM *well* first appears in limited contexts before it is progressively generalized to more contexts (2002: 176). Precisely, *well* as a DM was first anchored in the speech of others. It then branched out into the speech of the narrator. Finally, it developed meanings with strong orientation to the addressee's/reader's face, where it acts as a response hedge (*ibid.*). Such response hedges, as explained by Traugott & Dasher (2002: 176), tend to be used in cases of conflict or disagreement. Example (6) below gives an example where *well* hedges a response.

(6) *Tom*: Yes, you must keep a Maid, but it is not fit she should know of her Master's privacies. I say you must do these things yourself.

Ione: *Well* if it must be so, it must.
(Traugott & Dasher 2006: 176 [Taken from Jucker 1997: 102])

This change across a cline from non-/less subjective to intersubjective shows how over time the functions of DMs can generalize and even evolve into more contexts and increase in subjectivity and after that intersubjectivity. This idea of semantic change increasing in the direction of increased (inter)subjectivity is part of the hypothesis of the unidirectionality that has been present in the Traugottian approach to (inter)subjectification from the start (cf. Traugott 1982, 1995: 45; Traugott and Dasher 2002: 87 etc.). Though there have been cases that run counter to this widely recognized shift from non-subjective > subjective > intersubjective, they do not necessarily invalidate the unidirectionality hypothesis due to the fact that they are relatively uncommon (cf. López-Couso 2010: 143; Narrog 2010). Furthermore, even in cases where modifications are suggested to the unidirectionality hypothesis, they seek to generalize, or expand it, not disprove it altogether[8].

[8] See Narrog's (2010) suggestion to modify the unidirectional hypothesis to a more general increased speaker-orientation instead.

In the case of English DMs, however, things develop along the cline expected by the unidirectionality hypothesis. Diachronically, Traugott argues that DMs, particularly those that develop from adverbials to DMs, develop in "a manner that is consistent with grammaticalization in its early stages" (1997: 15). Therefore, in her investigation of the *indeed*, *in fact* and *besides*, she found that they developed diachronically from specific constructions to sentential adverbials and ultimately to DMs. Furthermore, though they do not follow the grammaticalization to the secondary stage, they do develop along the primary grammaticalization stage (lexical to grammatical) (Traugott 1997:15-16). Moreover, in a later study Traugott points out that "(s)ubjectification is more likely to occur in primary grammaticalization (the shift from lexical/constructional to grammatical) than in secondary grammaticalization (the development of already grammatical material into more grammatical material) (Traugott 2010: 40-41).

Having discussed in what direction linguistic items such as DMs develop, it is now time to discuss how this change occurs. Traugott (2018: 43) discusses the multifunctionality of another DM, *after all* (that was initially used in clause-initial temporal prepositional phrases only) in a diachronic study that argues that its multifunctionality results from "speakers using it to instantiate extant subtypes of DM[9]". Traugott goes on to argue that this occurs because "constructionalizations are conceptualized as the outcome of speakers partially matching the expressions to extant abstract constructional schemas and subschemas" (2018: 43). Furthermore, this is "enabled by interlocutors drawing on several elements of discourse production and understanding, most especially i) the pragmatic implicatures that arise from engagement in coherent discourse, and ii) the unconscious knowledge of both general

[9] In this case, DM means 'discourse marking'.

conceptual domains and of the nodes in the extant networks of DMs allowing for analogical matching” (43).

This process of developing subjective and intersubjective meanings occurs along this cline according to Traugott (2010: 32) as follows:

> At issue is the development of semantic (coded) polysemies that have to be learned with subjective or intersubjective meanings, and how these come into being. These polysemies may later be reinterpreted as homonymies (e.g. *fairly* ‘in a fair manner’ and ‘somewhat’), or one or more of the polysemies may cease to be used (e.g. *villain* ‘peasant’ and ‘evil person’), but by hypothesis most new semantic developments emerge as polysemies, pragmatic to begin with, then semantic.

Therefore, according to Traugott (*ibid.*) subjectified polysemies may:

> index evaluation of others (*silly* ‘blessed, innocent’ > ‘stupid’), of relative position on a scale (adverbs like *pretty* ‘cleverly’ > ‘attractively’ > ‘rather’), of attitude toward the truth of a proposition (epistemics like *probably* ‘provably’ > ‘in all likelihood’); they may index information structure (e.g. the topicalizer *as far as*), connectivity of clauses to each other (*anyway*), the speech act being undertaken (*promise* in its illocutio- nary uses), or the relationship of chunks/episodes of speech to each other (*then* in its discourse marker use).

Following suit, Traugott (2010: 32-3) adds that intersubjectified polysemies may, “index euphemisms *(the Lord* ‘god’, *pass* ‘die’, etc., see Allen & Burridge 1991), politeness (*please* < formulae like *If you please*, where the surrounding context has been absorbed into the meaning of *please*)”. For the purposes of this thesis, all established functions that semantically and pragmatically encode (inter)subjectivity will be referred to as (inter)subjective functions, while uses that are only pragmatically enriched by the context will be referred to as (inter)subjectified.

In the discussion thus far, a great deal has been said about how DMs can develop new meanings and how these new meanings develop across the (inter)subjective cline from non-/less subjective>subjective>intersubjective. Considering the highly (inter)subjective nature of

examples (1)-(4) in the introduction and the fact that these DMs were being used in novel ways, it will be imperative to verify how these novel uses developed (assuming that these are indeed conventionalized functions) and if these new functions indicate any recent shifts of the DMs under study toward further intersubjectification.

2.3.2 DMs and Genre-Specific Functions

Beyond the changes along the (inter)subjectification cline, DMs have also been proven to develop genre-specific functions. In recent years, there has been a number of studies that look at how DMs and other inserts develop specific text-type functions in both oral and written discourse (Aijmer 2013, 2015; Rühlemann & Hilpert 2017; Tottie 2017). Heading this research is Aijmer (2013, 2015), who studies specifically how DMs are used across a variety of different text-types. In these two studies, she investigates how a variety of DMs including *well*, *in fact*, *actually* and *I think* are used in different oral and written discourse types. This approach builds on many of the past 'integrative' or 'coherence-based' approaches to DMs from the past, starting with Schiffrin's (1987) pioneering study, 'Discourse markers', that defines DMs operationally as "sequentially dependent elements which bracket units of talk" (1987: 31). Moreover, Schiffrin also proposes a discourse model in which these DMs occurred that includes different planes: a participation framework, information state, ideational structure, action structure and exchange structure. Her analysis showed that DMs could work at different levels of discourse to connect utterances either on a single plane, or across different planes.[10] In a general sense, she defines the contribution of DMs to coherence as follows, "discourse markers provide contextual coordinates for utterances; they index an utterance to the local contexts in which utterances are produced and in which they are to be interpreted" (*ibid.*: 326).

[10] See also Schiffrin (2001).

However, in addition to being influenced by the integrative theories, Aijmer (2013, 2015) is most strongly influenced by the meaning potentials theory as developed by Norén & Linell, who state the following:

> A definition of meaning potential might run as follows: The meaning potential of a lexical item or a grammatical construction is the set of properties which together with contextual factors, including features of the linguistic co-text as well as various situational conditions, make possible all the usages and interpretations of the word or construction that language users find reasonably correct, or plainly reasonable in the actual situations of use. Speakers and listeners use the potentials of words and other expressions to mean and understand specific things in context (2007: 389).

Aijmer (2015: 202) argues that this theory of meaning potentials can account for why DMs have a number of easily identifiable (or conventionalized) meanings that can also be "adjusted to the demands of a new situation". This is possible because, as stated by Norén and Linell (2007: 390; italics in original), "parts of a word's meaning are evoked, activated or materialized, foregrounded or backgrounded, *in different ways in the different types of contexts, in which it is exploited*". Furthermore, Aijmer (2013) also argues that her approach can be compared to the description of DMs in a construction grammar framework by citing work by Fried & Östman (2005) in particular. Fried & Östman (2005) investigated DMs in Czech and Solv (a genetically Swedish dialect spoken in Finland), and underlined the fact that the constructional meaning of these DMs consists of a complex co-occurrence of different factors (syntactic, semantic and pragmatic).

This emphasis on accounting for how DMs are able to adapt to new situations is essential to understanding how these three colloquial DMs are able to transfer into journalistic prose with new specific functions, yet still maintain their colloquial character. Furthermore, this hypothesis that DMs can be "adjusted to the demands of a new situation" give insights into why certain DMs are able to be used in journalistic prose in ways that resemble how they are used in speech while also adapting to the constraints of the written medium.

Aijmer's approach to DMs contrasts with this, for example, in the Relevance theoretical approach[11]. In this approach to DMs, they are argued to act as a signal, or a signpost, to the co-locutor directing the way in which the following utterance should be processed (e.g. Jucker 1993). Furthermore, Blakemore (2002: 47) argues that *well*, for example, functions as a signal in the sense that "it provides a green light for the hearer, a sign to go ahead with the inferential process involved in the derivation of cognitive effects". In another Relevance theoretical approach to *well*, De Klerk (2005) finds that *well* is often used to signal that the speaker needs time to contemplate or to signal that the hearer needs to reconsider an assumption. In simpler terms, this framework sees DMs as being semantically underspecified, and that they are pragmatically enriched specifically in the communication situation. The Relevance theoretical attempt to find a common principle for all uses of *well*, and all DMs for that matter, is done in order to account for its highly multifunctional nature. However, in doing this, it is less able to explain how DMs like *well* are interpreted in relation to the text type, the role of the speaker, and the cultural context (cf. Talbot 1994; Lam 2009; Aijmer 2013, 2015).

In this discussion of how DMs get their variable meanings due to their context-bound character, Aijmer (2013) argues that context needs to be taken in a broad sense and has to go beyond the linguistic context, in order to include cognitive and social aspects. In the case of DMs, the social context is especially important in understanding how they are used. When discussing the social context of discourse, Aijmer (2013: 14) bases her approach to the social context on the social dimensions identified by Ochs (1996), which include social identity (e.g. speaker roles, relations to hearer), social act (e.g. request, offer), activity (e.g. story-telling, interview), affective stance (attitude, feeling) and epistemic stance (degree of certainty). As discussed in the previous chapter, stance is generally evaluated as "the smallest unit of social action" (Du

[11] See Sperber & Wilson (1986)

Bois 2007: 173) in which "social actors simultaneously evaluated objects, position subjects (themselves and others), and align with other subjects" (Du Bois 2007: 163). In terms of how DMs do this in particular, Fischer (2006: 445) observes that "discourse particles make the human interlocutors (with their hopes, fears, desires, and imperfections) part of the situation." Furthermore, Aijmer (2013) points out that DMs can be used with different stances and exhibit different speaker statuses depending on the social activity and speaker role.

This inclusion of social factors can explain a lot in terms of why DMs change and develop text-specific functions while also accounting for why they move along the (inter)subjective cline. Furthermore, it also gives us a theoretical explanation for the increasing use of DMs in journalistic writing in functions that look oral-like. Furthermore, the fact that these DMs are used in these functions most frequently in informal journalistic texts also indicates that the social atmosphere of which these articles are a part also play a role in the sudden usage of these DMs in specialized written functions that resemble orality.

Another framework that proves beneficial when it comes to analyzing these DMs in journalistic prose is the thetical framework (Kaltenböck *et al.* (2011). DMs constitute what are called formulaic theticals. Formulaic theticals are non-compositional information units that tend to be positionally flexible. Moreover, they express functions that are mostly procedural. In addition to this, they relate to the situation of discourse rather than to sentence syntax (Kaltenböck *et al.* 2011: 871). In terms of how theticals develop their functions, Kaltenböck *et al*. (2011: 875) argue in favor of the process of 'cooptation' whereby a clause, a phrase, a word, or any other unit is taken from sentence grammar and is coopted (or re-defined) for use as a thetical which operates under the domain of thetical grammar. Thetical grammar is shaped by the situation of discourse that is made up of six components. These components are listed in (7) below.

(7) Components determining the situation of discourse
Text organization (TO)
Source of information (SI)
Attitudes of the speaker (AS)
Speaker-hearer interaction (SHI)
Discourse settings (DS)
World Knowledge (WK) (taken from Kaltenböck *et al*. 2011: 861)

In thetical grammar it has been found that in most cases, rather than just one component being foregrounded in the usage of a thetical, it is most likely the case that a combination of components is foregrounded at once. In the particular cases of this investigation of DMs that cross over into journalistic prose, it will be worth noting if the components of the DMs under investigation shift in terms of their priority. For example, should the DMs under investigation in this thesis be shown to go through re-interpretation and conventionalize increasingly (inter)subjective changes insofar their written functions are concerned, then a shift in which components are more strongly activated in the situation of discourse is possible when DM uses cross over from spontaneous speech into writing.

The previous two sub-sections have explained how DMs evolve diachronically and how DMs expand into new conversation or text-types. One factor that has come up time and again in these two sub-sections is that the social context in which DMs are used is vital in how these DMs develop across the (inter)subjective cline and into new text-type specific functions. However, the issues taken up so far do not sufficiently explain why the changes should be occurring in journalistic writing at the time when they do. To get more insight into this timing, the hypothesis of colloquialization will be taken up in 2.4 below.

2.4 Changes occurring in written English with a focus on journalistic writing

Before discussing colloquialization as such, it is first necessary to discuss what distinguishes written and oral language. In chapter 1, Crystal's (2006, 2001) list of features differentiating

oral speech and writing was presented in Table 1.1. It was argued in the introduction that the usage of the colloquial DMs displayed how features 5 (socially interactive vs factually communicative) and 6 (immediately revisable vs. repeatedly revisable) are arguably less concretely discernable as belonging to either speech or writing than was previously believed. In a similar approach to the differences between oral speech and writing, Koch & Oesterreicher (1985, 1990) depict the difference between the two as a dichotomy between the phonetic and graphic mediums on the one hand and between the language of immediacy (typically oral) and distance (typically written) on the other (see Figure 2.2). Among the characteristics of the language of immediacy are features such as physical immediacy, privacy, familiarity (intimacy) of partners, high emotionality, deictic immediacy, dialogue, communicative cooperation and spontaneity. The opposite features would characterize communication distance (see also discussion in Kytö 2010: 49; Jucker 2000: 19-24). This continuum between language of immediacy and distance gives a global depiction of the major differences between written and oral speech. However, Jucker (2000: 24) points out that "the language of immediacy also occurs in the graphic code, and the language of distance also occurs in the phonic code", which indicates that there is a fuzzy border between these two language styles. This, therefore, indicates that there are specimens of speech or writing that exist "between the two extremes of the 'communicative continuum'" (Kytö 2010: 49).

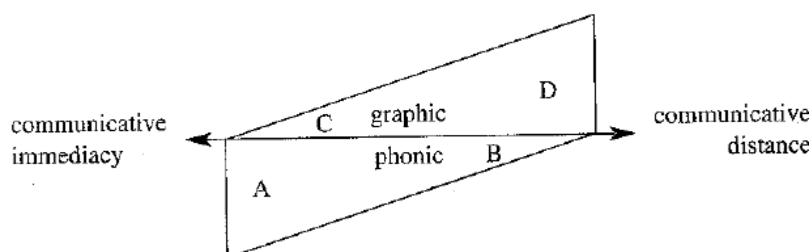


Figure 2.2. Continuum of text types (taken from Jucker 2000: 20)

The argument here is taken from the model developed by Koch & Oesterreicher (1985, 1990) which, much like the one created by Crystal, considers the two most extreme cases of speech

and writing. Though this model certainly would characterize the difference between formal academic writing and spontaneous speech in informal contexts, it hardly accounts for differentiating informal forms of writing, such as e-mail or personal journals from formal forms of speech such as public speeches or university lectures at least in terms of distinguishing communicative immediacy and distance.

However, over the past few decades there has been a variety of studies that do acknowledge these differences in writing styles and their relation to speech. For example, Biber (1988, 1995), and Biber & Finegan (1989) analyze the variation of speech and writing in terms of a continuum that comprises three categories (the element to the left being that most associated with writing, the one to the right with orality): "Informational vs. Involved Production", "Elaborated vs. Situation-Dependent Reference" and "Abstract vs. Non-Abstract Style". Biber and Biber & Finegan's works have established that three genres in particular – essays, letters, and fiction – have moved closer to the oral dimensions over the past two centuries. With these shifts towards writing styles that are involved, situation-dependent and non-abstract also come opportunities for the usage of dialogical elements of language in writing.

Furthermore, in the following years, other studies have also found evidence to suggest that the trends toward colloquialization in the 'popular written registers' mentioned above have begun to affect other more formal genres of writing as well.[12] When looking specifically at changes occurring in the 20th century, Mair (2006:188) as well as Leech *et al.* (2009: 240) expand on this by discussing how certain genres and particularly the press and fiction have shown more signs of colloquialization than others. For example, one reason that the colloquialization effect has been observed more in journalistic and fictional prose is probably due to the growing social

[12] See Westin (2002), Westin & Geisler (2002) and Mair (2006: 185) for a discussion on the expansion of colloquialization into formal writing.

and cultural pressures felt by the authors in these genres to appeal to a more informal audience (Mair 2006: 188). Using Biber's (1988, 1995) multi-dimensional framework mentioned above, Westin & Geisler (2002: 150) in their diachronic investigation of British newspaper editorials of three major newspapers found the following:

> The results of the dimension score analyses show that, during the 20th century, the language of British up-market editorials became less narrative (Dimension 2) but more persuasive and argumentative (Dimension 4). It also became less abstract (Dimension 5) and less dependent on referential elaboration (Dimension 3), which resulted in more informal language. As regards Dimension 1, which distinguishes between involved and informational discourse, no regular pattern was observed.
>
> The analyses also indicate that it was mainly during the latter part of the 20th century that these changes took place, since on three of the dimensions, the last time period (representing the years 1960 through 1993) stands out as different from the preceding two periods.

Furthermore, in addition to these overall changes in the British newspaper editorials, they also found that one newspaper (The Guardian) showed stronger changes in its linguistic and stylistic shifts toward informality than the other two (The Daily Telegraph and The Times). In particular, The Guardian was found to be "more involved and less explicit and abstract, in other words, more informal" (Westin & Geisler 2002: 150). Their hypothesis is that this stronger shift in the Guardian is probably due to the audiences that are addressed.

However, despite these shifts towards more oral-like writing styles, there are still limitations to how oral-like writing can become. Such limitations include the lack of being able to use prosody, facial-expressions, gestures and conventions of body posture etc. (Crystal 2001: 47). Moreover, the most obvious difference between writing and speech is the phonic and graphic mediums of producing language. Therefore, even in cases where genres of writing are becoming more like speech, the fact of the matter is that the two can never become identical. Therefore, one open question that this thesis takes up is how dialogical elements of language can be adopted in writing and also innovated in ways that make them useful in the written medium.

Moreover, this thesis will also explore in what types of articles the usage of dialogical DMs are used. For example, will these DMs be used in all areas of news equally, or are they used more particularly in certain sub-sections of journalism. Furthermore, based on the findings of the types or articles that include dialogical DMs in journalistic prose, this thesis will determine whether or not authors of such articles constitute a specialized 'discourse community' amongst other journalism authors. Swales (1990, 2016), who primarily analyzes (academic) written texts, describes discourse communities as groups that have goals or purposes, and use communication to achieve these goals. In terms of defining what constitutes a discourse community, Swales lists six defining characteristics whereby a discourse community:

1) Has a broadly agreed set of common public goals;
2) Has mechanisms of intercommunication among its members;
3) Uses its participatory mechanisms to provide information and feedback;
4) Utilizes and possesses one or more genre in the communicative furtherance of its aims;
5) In addition to owning genres, it has acquired some specific lexis;
6) Has a threshold level of members with a suitable degree of relevant content and discoursal expertise

Another key element to Swales' interpretation of a discourse community is the notion of 'genre', which is the organizational patterns of written language (Swales 1990: 9). Swales states that in genres these particular patterns of written language both belong to the discourse communities in question and help to define these same discourse communities (*ibid*.). Nine years later when discussing the discourse of 18th century grammar writers, Watts (1999: 43) defined a discourse community as "a set of individuals who can be interpreted as constituting a community on the basis of the ways in which their oral or written discourse practices reveal common interests, goals and beliefs, i.e. on the degree of institutionalization that their discourse displays". He goes on to explain that the members of these discourse communities may or may not be conscious of sharing discourse practices and therefore, the membership affiliation may be strong or weak depending on the circumstances and the "community itself may only become 'visible' through the course of time" (*ibid*.). A discourse community is differentiated from other

similar concepts such as a 'community of practice which 'Eckert & McConnell-Ginet (1992: 8) define as follows:

> A community of practice is an aggregate of people who come together around mutual engagement in some common endeavor. Ways of doing things, ways of talking, beliefs, values, power relations - in short, practices - emerge in the course of their joint activity around that endeavor. A community of practice is different as a social construct from the traditional notion of community, primarily because it is defined simultaneously by its membership and by the practice in which that membership engages. Indeed, it is the practices of the community and members' differentiated participation in them that structures the community socially.

A community of practice, which is applied often to sociolinguistic studies involving speech data is differentiated to a discourse community by Pogner (2005: 9) who states that

> If the focus is on the aspect of a common discourse developed in order to be effective in the domain in question, the community can primarily be analyzed as a discourse community. If the focus is on the aspect of common practice developed in order to be effective in the domain, the community can primarily be examined as a community of practice.

Ultimately, the investigation in this study involves articles written by professional journalistic writers in newspapers and magazines. Journalism in and of itself constitutes a written genre of writing. However, this thesis will also take the types of articles from which the dialogical DMs are used in journalistic prose into account. Therefore, it will be analyzing not just the journalism genre as a whole, but also the individual sections that encompass any given newspaper or magazine. Therefore, this thesis will not only set out to explain trends occurring in journalism writing as a whole, but it will also set out to analyze specific trends in any of the various sub-genres of journalism in which dialogical DMs are used. Specific patterns of DM usage and what makes them unique will be addressed in the upcoming case study chapters 3-5.

In addition to the studies on the changes occurring in certain areas of written English, Mair (2006) and Leech *et al.* (2009) have discussed the specific changes that occur due to the

colloquialization of 20th century English. The term 'colloquialization' was first introduced in Mair (1997) as a framework that attempted to highlight and explain stylistic shifts in 20th century written English:

> -away from a written norm which is elaborated to maximal distance from speech and towards a written norm that is closer to spoken usage, and
> -away from a written norm which cultivates formality towards a norm which is tolerant of informality and even allows for anti-formality as a rhetorical strategy (Mair 2006: 187).

As already mentioned in the introduction, these changes have been observed primarily from the second half of the 20th century, and it involves changes at two levels of writing. This includes changes at the macro-structural level (such as the increase of quotations) and at the micro-structural level (where a more informal construction is selected over a formal one, such as the use of verb and negation contractions). Furthermore, the changes also include an increase in the informalization of written language, which must be distinguished from colloquialization. Leech *et al.* (2009: 239) discuss this distinction in the following passage:

> There is, however, a distinction (in principle) between informalization and colloquialization. Informality of style, which is a matter of the absence of distance between addresser and addressee, is strongly associated with speech, but can also have characteristics specific to the written language. Consider the article-free apposition of Supermum Sue (as contrasted with Susan, a perfect mother) – a decidedly informal construction which is nevertheless absent from speech. On the other hand, colloquialization (as seen, for example, in the increasing use of contractions) is more evidently a matter of adopting speech-like habits.

Moreover, Leech *et al.* (2009: 247) also specify that with an informal style, in which the writer attempts to create a personal rapport with the reader, comes an increase of colloquial elements in language. Therefore, though informalization is not the same as colloquialization, an informal style of writing often comes with an increase in the use of colloquial linguistic elements as well.

However, what I will be investigating in this thesis is different from what has been observed in previous studies on the colloquialization of journalistic texts. The main aim of this thesis is to analyze the usage of colloquial DMs in journalistic writing in newly evolved and specialized

written functions. Until now, there is only one study to my knowledge that has shed any light on this question, and that is Rühlemann & Hilpert (2017). This paper investigates the use of the DM *well* in the TIME corpus (Davies 2007), which covers 1923-2006. Not only do they find that *well* enjoys a marked rise in frequency of usage in the 1990s, they also find that coincidentally around the same time a new major function of *well* appears. Furthermore, this function, labelled by them as 'predicative *well*', was not only a new function in their corpus data from the end of the 20th century, but it was also a function not yet discussed in any study of *well* involving written or oral data. When discussing predicative *well*, Rühlemann & Hilpert (2017: 129) state that:

> It appears that the common functional denominator to both lexico-stylistic variants is as a marker of word choice. As a word-choice marker, predicative-*well* prepares the reader for upcoming wording which is, in one way or another, peculiar (or "marked"), either as repetition or as word play. Repetition is peculiar in that it violates the principle of "elegant variation" that authors of news magazines, as other written text types, are normally held to observe. Word play is peculiar in that, in order to be enjoyed by the reader, it needs to be recognized as word play. Subtle word play may go unnoticed, so *well* may help to make the reader become aware of it.

Despite the fact that this function of *well* is relatively recent as far as conventionalized functions of *well* are concerned, they also find that predicative *well* is reminiscent of other prototypically oral functions of *well* as they discuss below:

> Nonetheless it seems possible to argue that the word-search function is the model after which the word-choice function is crafted. That is, *well* is used as if the writer were searching for the appropriate wording. The effect is carefully calculated: just as conversational word-search *well* indexes the speaker's planning difficulties drawing the interlocutor's attention to the searched-for wording so too does predicative-well focus the reader's attention on the expression to follow […] Bluntly re-using lexical material in close vicinity to its first use is generally considered bad style and therefore avoided. Seen through this lens, *well* in the context of repetition is reminiscent of *well*'s function in conversation, where it "standardly prefaces and marks dispreferreds" (Levinson 1983: 334) – that is, adjacency second-pair parts which are in some way contrary to the expectations raised by the preceding adjacency first-pair part (e.g., refusing an invitation) (130).

The results of this study by Rühlemann & Hilpert demonstrate how *well,* a common dialogical DM, is capable of being used in writing in a way that is reminiscent of how it is used in oral

discourse, yet is used in a way that is specifically useful to the authors. In this case, it is useful to Time Magazine authors in such a way that it lets them make their articles feel more interpersonal. It is precisely this specific development of a written function of *well* that differentiates it from the changes observed in the past colloquialization studies, because those changes simply involved the usage of the more colloquial options over the more formal ones or the inclusion of more quotations. Some open questions that remain from this study, however, is how many of these uses of *well* in their data were from true journalistic prose, and in what types of articles were these new functions most used.

There are similar findings in Tottie's (2017) investigation on *uh*, *um* and *er* (EHM in written language UHM in speech). In this study, EHM was found to be used the most in journalistic prose, specifically in the sentence-medial position (Tottie 2017). When used in writing, EHM resembles being used in a word-searching function. However, in journalistic writing EHM is used in an innovated way that actually highlights the author's attitude. EHM was used the most in areas of news such as reviews, comment columns and sports pages in magazines and newspapers (*ibid.*) According to Tottie:

> Although EHM gives a colloquial flavor to the written message, there is more at stake than merely using a spoken expression for the same function as it has in typical speech. EHM serves very different functions in the two media. Spoken UHM is un unintentionally used (along with pauses) for online planning of upcoming speech…, but written EHM is used intentionally by writers to highlight their attitude either to a preceding or following proposition or expression. (2017:128)

In this study, I will be looking at the overall frequencies of each DM in question, but also specifically at their usage in journalistic prose. By investigating both how these DMs are evolving overall in writing and in journalistic prose specifically, I will be able to address many questions that remain open in terms of how colloquial DMs are used in writing. For example, is this development of new written functions unique to *well* or EHM or are similar developments happening with other colloquial DMs too? Furthermore, are the new functions like predicative

well used across the board in journalistic texts, or are they niched into certain sub-genres of journalism as was the case the EHM? Also, are these new functions truly unique to journalistic texts, or have they also been creeping into other genres of writing as well?

2.5 Selection of DMs for the case studies

Three DMs were ultimately selected and will be investigated in individual case studies, which will make up the empirical parts of this thesis These three DMs are *well*, *like* and *oh yeah*. I selected these three DMs on both theoretical and practical grounds. As the main aim of this study is to observe change in how dialogical DMs are used in journalistic texts, it was necessary to find prototypically dialogical DMs that are also actually used in the written medium.[13] Another major criterion for the selection of the final list was the development of new functions that were specific to writing. Based on these criteria and the results of a series of exploratory studies a variety of DMs[14], *well*, *like* and *oh yeah* were selected as the best candidates for the case studies. Besides being one of the most widely studied and recognizable DMs in the English language, *well* is also the only DM in the list that has received any attention in journalistic writing, as discussed above. However, considering that Rühlemann & Hilpert's (2017) study used data that comes from only one magazine, it is important to investigate how *well* is used in other news sources. Furthermore, as stated in the previous section, special attention will also be given to the type of articles where the written variants of *well* are used. *Like* is the most stigmatized of the three DMs investigated in this study. This is due primarily to its association with youth culture and Valley Girls in particular from the 1980s to the present (e.g. Dailey-O'Cain 2000; D'Arcy 2007). Finally, *oh yeah* is the least studied of the three DMs under investigation, but it was selected as the final case study for this thesis because of the tremendous

[13] For the purposes of this study, the usage of DMs under investigation in quotations are considered transcribed speech and not written variants of the DMs. For a discussion on the increased usage of quotations in modern English writing, see Leech *et al.* (2009: 248-9).

[14] Other DMs that were investigated were *actually*, *in fact* and *you know*. Though there was evidence of usage of these DMs in writing, similar surges of newly developed functions from the 1990s were not confirmed.

amount of growth and change that it has taken place in journalistic writing since the 1990s. Though only these three DMs were selected for the case studies of this thesis, comparisons with other DMs from the same corpora will be made in order to show that the changes occurring with these three DMs in writing are special and different from what is happening with the others.

2.6 Selection of Corpora

The corpora for this study were selected on practical grounds. A corpus is a collection of texts used for linguistic analyses, usually stored in an electronic database so that the data can be accessed easily by means of a computer. The aim of this study is to observe diachronic change in how the three selected DMs are used in American journalistic writing. To achieve this diachronic aim, the journalistic sub-corpora – which comprises the newspaper and magazine corpora – of the COHA were selected. The COHA has data ranging from the 1810-2009. This makes it a great resource to observe diachronic trends in American writing. In addition to providing data from the past two centuries, it also provides a well-balanced representation of different newspapers and magazines. There are over 100 different magazines sampled in the COHA with, according the website, "a good mix (overall and by year) between specific domains (news, health, home and gardening, women, financial, religion, sports, etc.) (Davies 2010)"[15]. A balanced representation of newspapers and article types are also well represented in the COHA. Overall, ten newspapers from across the U.S., including USA Today, New York Times, Atlanta Journal Constitution and San Francisco Chronicle, are sampled. Furthermore, according to the website "there is a good mix between different sections of the newspaper, such as local news, opinion, sports, financial, etc."(Davies 2010).

[15] Information on the composition of the corpora can be found in the link below:
<https://www.english-corpora.org/coha/>(accessed 29.01.2020).

In addition to the journalistic sub-corpora of the COHA, however, there is also the fictional and non-fictional sub-corpora that can be readily compared to the data from the journalistic sub-corpora. As all three DMs enjoy a significant and seemingly sudden surge in usage starting in the 1990s, which represents the last two decades of the COHA, it was deemed necessary to observe data from the last decade (2010s) as well. To accomplish this end, and to compare the COHA data to similar, yet more informal contemporary journalistic data, the NOW corpus was selected. The NOW corpus is an interesting resource to use for a comparison of the COHA, because many of the traditional news sources (such as TIME magazine, the Wall Street Journal and USA Today) found in it are the same as in the COHA. However, the NOW also includes many news sources that are available online only (such as BuzzFeed, TV.com and Vulture). Therefore, the comparison of these two corpora allows one to not only observe further changes that may be occurring in journalism in the decade following the end of the COHA, but it also allows one to check if the shift from traditional print journalism (which is now also available online) to web-only journalism is leading to any new dramatic changes. Table 2.1 provides the details pertaining to both of these corpora.

Corpus	Size of Corpus	Fixed or Monitor Corpus?	Composition	Text-Types	Period of Time
COHA	• 400 Million Words (total) • 137 Million Words (Journalism)	Fixed	• Fiction • Non-Fiction • Magazines • Newspapers	Published Print Texts	1810-2009
NOW	• 9 Billion Words (total) • I Billion Words (USA)	Monitor	• Magazines • Newspapers	• Print & Web • Web only	2010-today

Table 2.1 Makeup of COHA and NOW corpora

2.7 Methodology

Since this investigation combines quantitative and qualitative methods, it therefore takes a 'Corpus Pragmatics' approach to investigating these DMs in writing. Corpus Pragmatics, according the Rühlemann & Aijmer (2015: 12 emphasis in original) "is more than just

pragmatic research and it is more than just corpus-linguistic analysis in that it *integrates* the horizontal (qualitative) methodology typical of pragmatics with the vertical (quantitative) methodology predominant in corpus linguistics". Furthermore, considering that this is a corpus-based study that looks at diachronic language change to explain the development and spread of particular DM constructions, this study adopts a usage-based outlook on language (cf. Diessel 2014). It is important to combine quantitative and qualitative methods in this investigation of dialogical DMs in journalistic prose because it is not possible to determine in what function(s) these DMs are used based solely on the position in which they appear. Such details can only be determined by manually observing individual concordances.

In particular, this thesis adds to the relatively recent trend of observing DMs through the lens of corpus data (e.g. Müller 2005; Aijmer 2013; Beeching 2016), as opposed to relying on data collected by the researcher him/herself. The main advantage of using corpus data to observe how DMs are used in journalistic writing is that corpus data allows one to analyze how these DMs are used in real life contexts from a variety of sources. Furthermore, corpus data is also advantageous in such an investigation because the primary aim is not to observe how these non-dialogical examples of DMs *could* be used in writing, but instead the primary aim is to observe in what ways these DMs *are* used in writing. In order to investigate these questions, it is therefore best to use corpora that allow one to observe how the DMs in question occur in different text types diachronically. In this particular study, thanks to the availability of the COHA, I am able to first observe how the three DMs in question are used form the 19th to the beginning of the 21st century. I am then able to complement the journalistic COHA data with contemporary data from the NOW corpus, which is a monitor corpus (a corpus that is updated and expanded regularly) that starts in 2010 and continues to the present day.

In the diachronic quantitative portion of each case study, the frequencies per 1 million words are taken per decade in the COHA for the following sub-corpora: fiction, non-fiction and journalism. In order to facilitate more robust comparisons to how the DMs in question are used in each genre, their Kendall's tau τ correlation coefficient scores are calculated for their diachronic developments. When discussing the purpose of calculating a correlation coefficient with numeric variables, Gries (2013:373) states that:

> By computing a correlation coefficient, which usually falls between -1 and +1, one tries to answer the following questions:
> - is there a relationship between a variable x and a variable y such that, on the whole, one can say "the more x, the more y" and/or "the less x, the less y" or, on the other hand, "the more x, the less y" and/or "the less x, the more y"? If the relationship is of the former type, then the correlation coefficient will be >0; if the relationship is of the latter type, then the correlation coefficient will be larger than <0; if there is no relationship between x and y, the correlation coefficient will be ≈0;
> - how strong is this relationship? The more the correlation coefficient differs from 0, the stronger the correlation;
> - is the correlation statistically significant?

Furthermore, Gries argues that Kendall's tau (τ) is more beneficial when looking at diachronic linguistic frequency trends than the others such as Pearson's product-moment correlation. He argues that despite being "a bit less powerful", Kendall's tau (τ) is the best choice, because it is "less sensitive to potentially problematic distributions" (*ibid.*: 375). For the purposes of this study, I will be looking at the diachronic frequency trends of the DMs in question for all of the genres in COHA in order to determine if the changes in how DMs are used are unique to journalism. After the analysis of the COHA data, the data from the 2010s from the NOW corpus is analyzed in the second part of each case study to see if the observed trends continue or differ from what was expected. In addition to looking at the overall frequency changes, when applicable[16], this diachronic quantitative analysis will also study the DM usage in certain syntactic positions; namely, in the sentence-initial, medial and final positions.

[16] This is particularly applicable when looking at *well*, which will be discussed in chapter 3.

In addition to looking for an overall positive diachronic correlation, it is also of interest to analyze the stages of change found in the diachronic development of each DM. To observe these periods of change, the tokens per million words for each decade are run through the Variability-based Neighbor Clustering (VNC)[17] algorithm (Hilpert 2013). This is a form of Hierarchical Clustering Analysis that is applied to temporally ordered data. This means that it only merges two data points if they are temporally adjacent. By applying the VNC algorithm to the DM data, we can see if their developments can be divided into a sequence of distinct stages. Using the VNC is advantageous because of its abilities to detect stages in a data-driven 'bottom-up way', which is more robust and reliable than depending on an analyst's subjective impressions (*ibid.*: 34). Figure 2.2 gives an example of what a VNC dendrogram looks like.

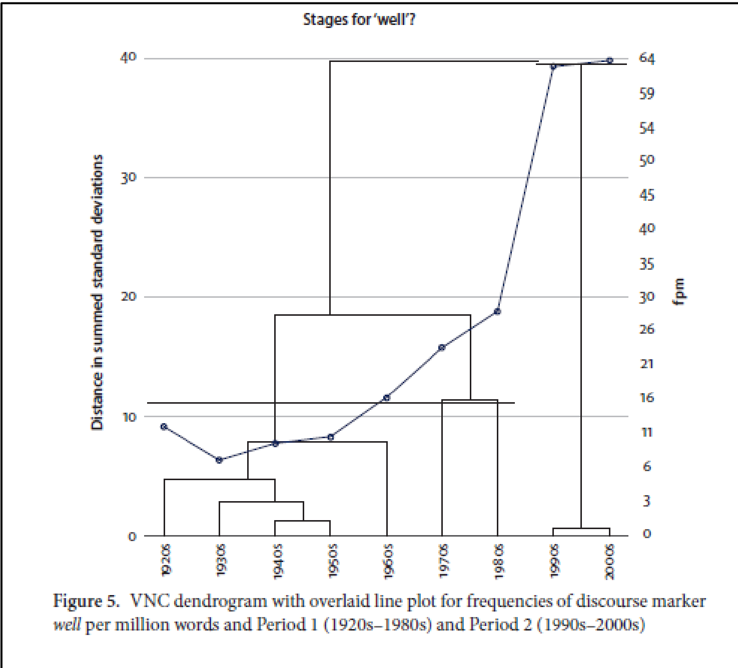


Figure 5. VNC dendrogram with overlaid line plot for frequencies of discourse marker *well* per million words and Period 1 (1920s–1980s) and Period 2 (1990s–2000s)

Figure 2.2: VNC Dendrogram (Taken from Rühlemann & Hilpert (2017:115)

[17] For a detailed discussion of the Variability-based Neighbor Clustering (VNC), see Gries & Hilpert (2008, 2012) and Hilpert (2013).

Acting as a sort of exploratory method to bridge the gap between the quantitative analysis of each DM and the qualitative analysis, the collexeme analysis[18] is used. Collexeme analysis is used to determine which items are most strongly attracted to a given word or construction and generates a ranked list (Hilpert 2012: 132). Furthermore, what it does is determine if the observed frequency of word combinations –in this case the individual DMs and the given collocate – is significantly higher or lower than their expected frequency given how often each word/construction occurs in the corpus overall (*ibid.*). Furthermore, the collexeme analysis is a useful tool because it ranks the attraction of the word in question to its collocates not by the raw frequencies, but by their degree of attraction based on their collostructional strength score that is calculated using the Fisher-Yates Exact test. The Fischer-Yates exact test is used to calculate the collostructional strength because it does not require large frequencies, and it places no particular restrictions on its input data (Stefanowitsch & Gries 2003: 218). This collexeme analysis can potentially be a great tool to help indicate how these respective DMs are being used in writing, especially if they are developing conventionalized functions with fixed patterns (or constructions). To calculate the collocation strength score, I put the following information through the collexeme algorithm as can be seen in Table 2.2 below, which includes the collocates of *well* at 1L in the 1990s-2000s COHA data, the frequency of co-occurrence and the total frequency of the collocate in the corpus. Not included in this table is the size of the corpus and the frequency of the DM *well* in the corpus, which are also put in the collexeme algorithm.

[18] For more information on collocation analysis, see Stefanowitsch & Gries (2003) and Hilpert (2012). The collocation analysis itself was generated with aid from an algorithm for R provided by Flach (2017).

Collocate	Frequency	Frequency in Corpus
A	39	196327
OF	19	194996
THE	17	433454
TO	13	199191
THAT	11	82430
IS	10	74182
I	8	48225
IT	8	65023
AND	7	201468
SAY	6	6829

Table 2.2: Top 10 collocates of *well* at 1L

Once this is run through the collexeme algorithm, it gives the collocates back in the order of their strongest collostructional score rather than their raw frequencies as can be seen in Table 2.3. In the upcoming chapters of this thesis, I will show the results of the collexeme analysis only as in Table 2.3. In the case where words are mutually repelled from one another, they are ranked from the least strongly repelled to the most strongly repelled. The first column, moving from left to right, shows the collocate in question, while the following column states whether the collocates and DM are attracted to or repelled from each other. The next column shows their collostructional strength score that is calculated using the Fischer-Yates exact test and at the very right is their level of statistical significance.

	COLLEX	ASSOC	COLL.STR.FYE	SIGNIF
1	A	attr	12.35414	*****
2	YEAH	attr	11.53126	*****
3	OH	attr	7.81997	*****
4	SAY	attr	5.87122	*****
5	LIKE	attr	4.30434	****
6	BECAUSE	attr	2.91875	**
7	YOU	attr	2.58761	**
8	THAT	attr	2.56451	**
9	DO	attr	2.55701	**
10	I	attr	2.55451	**

Table 2.3 Sample Top 10 of Collexemes of *Well* based on Collostructional strength as generated by the software provided by Flach 2017

However, despite the potential for the collexeme analysis to indicate hints as to how these DMs are being used in the corpus data, the only way to get a complete understanding of how these DMs are used in journalism is to observe the individual concordances. Therefore, the bulk of each case study consists of a qualitative analysis of either all or a sample of the extracted corpus data. In this analysis, the tokens used in non-dialogical functions receive special attention. These non-dialogical examples are annotated for their position in the sentence, their primary function, the type of lexical item or phrase being marked, the type of article it is in and any other distinguishing properties[19]. When it comes to labeling their primary functions, this work is influenced by the labels used in the previous studies of the individual DMs. However, in some cases, the DMs are used in new ways. In these cases, new labels had to be invented. In the qualitative part of each case study, not only is special attention given to the new written functions, but also the categories of articles in which they are used. This special attention is given to the types of articles, because, as explained above, little attention has been given to distinguishing from where in the newspapers and magazines the colloquial elements of language are being used. Furthermore, in the NOW data the source of the news magazine or newspaper, i.e. traditional news source or web-only, is taken into account. This will allow me not only to determine if the usages of certain DM functions are developing and growing in journalistic writing. It will also allow me to determine if such developments are occurring in journalism across the board or only in certain sub-genres of the press.

2.8 Summary and conclusions

In this chapter, I defined the major characteristics and functions of DMs. In terms of how DMs develop, it was found that the textual functions tend to come first before the (inter)subjective (interpersonal) ones develop, as discussed in Traugott & Dasher (2002) and Traugott (2010).

[19] This could be anything from co-occurring with other DMs to marking a word repetition or sequence of word play.

In addition to developing along the (inter)subjective cline, however, DMs have also been found to develop ‘text-type’ specific functions that allow them to expand into a variety of genres of language. Though the focus of DM enquiry has been primarily on oral speech in the past decades, there are indications from a small number of studies that such expansions are possible in writing as well.

This thesis, therefore, straddles the line between studying how dialogical DMs develop specialized functions outside of their prototypical dialogical environments and studying recent change in journalistic writing from the end of the 20th century to today. In order to investigate these changes in how dialogical DMs are used in writing, this study will use the diachronic COHA corpus which will be complemented by the NOW corpus. In the following three chapters, the DMs *well*, *like* and *oh yeah* will be analyzed in separate empirical case studies. After that, some common observations will be discussed together in chapter 6.

Chapter 3:
Non-dialogical *well*: Rhetorical marker of (inter)subjective word choice

3.1 Introduction

The DM *well* is one of the most studied DMs in the English language (Beeching 2016: 51). In the vast body of literature that addresses *well* as a DM, one common characteristic that is mentioned is how *well* is commonly used in informal spontaneous interaction. Furthermore, its most discussed function is summed up as flagging a demurral, often as a dispreferred response (Biber *et. al* 1999: 1086; Müller 2005: 122; Beeching 2016: 52). However, more recent studies have begun to investigate its usage outside of these informal conversational environments and have started looking at monologue contexts (Aijmer 2013; Beeching 2016). This chapter continues this research of *well* in non-prototypical environments by looking at its use in published texts and journalism in particular. Moreover, this chapter will focus on the development of the uses of non-dialogical *well* from the end of the 20th century to today. These specialized functions can indicate intentional repetition or other forms of playful language use in the predicative construction, as in example (1). In other cases, it is also used as a word-choice marker (also referred to as WCM in this chapter) that performs various secondary functions such as highlighting intentional repetition, playful language use and end of list markers. Example (2) shows *well* in the end of list function to mark the final and clearly most important lexical item in a list.

(1) But as the Washington Post's Fact Checker noted, the $21 trillion claim was quite, **well**, fuzzy. (CNN: The Point with Chris Cillizza 01.10.2019[20])

(2) This team is more talented than his previous three. Bigger and stronger and more, **well**, more like that school's across town. This team is his team. (COHA Magazines: Sporting News 2006)

[20] This example comes from a CNN online columnist. It can be read in its entirety below:
< https://www.cnn.com/2019/01/07/politics/alexandria-ocasio-cortez-facts/index.html>

These uses of non-dialogical *well* resemble the use of sentence-medial *well* from spontaneous speech where *well* marks pausing and/or word-searching (e.g. Aijmer 2013: 28). Interestingly, in many studies of *well* based on spontaneous speech data, these sentence/turn-medial uses have been found to be far rarer than the functions of *well* in the sentence/utterance-initial positions. In the sentence-initial position, *well* often guides the addressee on how to react to what is about to come (e.g. Biber *et al.* 1999; De Clerk 2005; Aijmer 2013). However, in the COHA and NOW data, sentence-medial *well* performs four main tasks. In particular, this includes quote *well* (3), clause *well* (4), word-choosing marker *well* (2), and predicative *well* (1).

(3) Are you still going to be our daddy after Mommy gets married again? My response : **Well** , sit right down, Son, and let's chat about that. Never mind about your room. Still, I count certain victories. You want me to cut the crust off your grilled cheese? Here's a butter knife. (COHA Magazines: Parenting 2007)

(4) The creatives say they can help clients sell diet soda and detergent in this radically changed environment because, **well**, that's what they've always done. (COHA Magazines: Fortune 2004)

Quote *well* and clause *well* will be shown to be functions that also exist in oral speech, while word-choosing *well* and predicative *well* were found to be more particular to writing, especially from the end of the 20th century to the present. Quote *well* is used at the beginning of a quote of an 'other' or sometimes even a self or thought quote. Clause *well* occurs at the clause structure boundaries. As (4) above exemplifies, clause well is typically used in positions right before the occurrence of the main clause. As a preference, clause *well* occurs at the boundary between an adverbial and the main clause. On the other hand, the word-choosing marker and the predicative *well* functions operate as highlighters of lexical items and simultaneously serve textual and interpersonal functions. As example (1) shows, the predicative *well* function highlights a lexical item in the subject predicative.

This chapter is motivated by the following questions. In what ways is *well* changing when used in journalistic prose? Are these developments to *well* happening at a specific period? What can be influencing these changes during these specific periods? Are these changes specific to journalism or are they happening within other mediums of writing as well?

The structure of this chapter is as follows. Section 3.2 surveys past studies on *well* and highlights the important theoretical questions that will be discussed in this chapter. Section 3.3 analyzes the diachronic development of *well* in the COHA, and, in particular, it investigates the evolution of *well* in the sentence-medial positions. Section 3.4 adopts a qualitative perspective on the specific functions of sentence-medial *well* in the COHA and pays special attention to the most recently developed functions of non-dialogical *well*. Section 3.5 continues the analysis of the non-dialogical sentence-medial *well* using contemporary journalism data from the NOW corpus in order to determine if the developments observed in the COHA data continue. Section 3.6 analyses the major functions of non-dialogical *well* in the NOW corpus. Finally, 3.7 presents the conclusions of the chapter.

3.2 Previous work on *well*

Besides being one of the most studied DMs in the English language, *well* has also been used as the testing ground to different theoretical approaches to the study of DMs and has been investigated by the means of several methods of study. The reason for this is undoubtedly due to its strong multifunctionality in 'textual' and 'interpersonal' uses as well as its ability to be both retrospective and prospective (e.g. Aijmer 2013; Beeching 2016). In addition to its DM functions, *well* also has interjection uses. Exclamatory '*well*!' or repeated '*well*, *well*!' or even '*well*, *well*, *well*!' are used to express surprise or shock or even to comment on misbehavior in the case of sequenced *well* (Beeching 2016: 51). Further, with falling intonation, *Oh well* expresses resignation (*ibid*.). Due to its informal dialogical nature and its preference to occur

sentence-initially, most early studies on *well* used oral data and concentrated on its functions as a response marker.

For example, Lakoff (1973) found that answers may be prefaced by *well* under certain conditions. Such conditions include if the answer is 'indirect' (1973:458), if the information supplied in the response is only a partial response to the question (*ibid*.: 459), or if the speaker senses some sort of 'insufficiency' in his answer (*ibid*.: 463). Furthermore, Lakoff also found that *well* is used in narrative functions in order to "indicate that details have been omitted" (*ibid*.: 464). Another early study of *well* comes from Svartvik (1980) who agrees with Lakoff's analysis of *well* in answer/response functions, but also analyses other functions. Globally, Svartvik categorized *well* into two major categories: *well* as a quantifier and as a frame (*ibid*.: 173). The quantifier functions include agreement, reinforcement, the non-straight and incomplete answer to the wh-question and non-direct or qualified answers. The frame functions, which he states normally occur non-initially, include shifts of the topic or focus, introducing explanations, the beginning of quoted discourse and functions of editing for self-correction. Similar findings were discussed in Schiffrin's (1987) chapter on *well* that analyzed it from the scope of her model of discourse with five planes of talk. As in the other studies, the primary focus of her chapter is on the interactive response/answer functions where *well* marks indirect, insufficient, qualified, non-straight and incomplete answers, all of which are "not fully consonant with prior coherence options" (1987: 103). In the second part of her chapter, she also briefly discusses functions that do not involve the speakers' responses to their interlocutor's talk, but "to their own talk" (*ibid*.: 123). These functions, according to Schiffrin, "may consist of adjustments to the ideational content of talk (self-repair; …), changes in the deictic center of talk (reported speech) and alternations in the objectivity of talk (reflexive frame breaks) (*ibid*.: 123). Reflexive frame breaks, on the other hand, is explained as displaying one's "shift in

orientation to evaluation of events" within their own just-completed discourse (Schiffrin 1987: 125).

Various studies address the historical development of *well*. In a study of *well* in responses, Finell (1989) argues that early propositional uses of *well* as a predicative adjective were the precursors for the interpersonal response markers uses of *well* that we are familiar with today. On the other hand, Jucker (1997) argues that the origins of modern *well* can be traced back to old English *wella* (well + la), which had an attention-getter function such as "listen-up" that shares some similar functions with Old English *hwæt*. According to Jucker (1997: 106), over its history the DM *well* moves from interpersonal in old English to textual in Middle English. However, it then moves back to interpersonal in Early Modern English and continues into Modern English (*ibid*.). For the purposes of this chapter, however, we are informed by the development of *well* as described by the cline of intersubjectification as proposed by Traugott (1989) and Traugott & Dasher (2002: 175-6) where *well* develops from propositional, over to textual, to having interpersonal meanings which increase in subjectivity and intersubjectivity. In the case of *well*, its diachronic evolution goes from adverbial form only to textual (DM) functions that then develop elaborative senses (subjectification) (*ibid*.: 176). After this it develops into hedges that soften or mitigate what is said by acknowledging the addressee's actual or possible objections (intersubjectification) (*ibid*.). According to this same work, when *well* started to appear with DM meanings, it began to occur only in limited contexts. Over time however, *well* "is then generalized to more contexts" (2002: 176). Traugott & Dasher (2002) also noted that *well* is first anchored in the speech of others (example 5), then it was expanded to the narrator's voice (example 6) and finally it further extended to the narrator/speaker/writer's perspective. Finally, it developed meanings with strong orientation to the addressee/reader's face (example 7) (*ibid*.). In a similar discussion of the development of *well*, Defour (2010: 190) argues that "(a)s a common denominator, the element of acceptance

inherent in adverbial *well* can be seen to be increasingly 'recruited' in contexts where speaker and addressee need a means to express speaker attitude, and to establish a greater interactional understanding between different participants".

(5) "Ye sey welle," sayde the kynge, 'Aske what ye woll and ye shall have hit and hit lye in my power to gyff hit." "**Welll'** seyde thy lady, "than I aske the hede of thys knyght that hath wonne the swerde." "'You say well,' said the king. 'Ask what you will and you shall have it if it lies within my power to give it.' 'Well,' said the lady, 'then I ask the head of the knight who has won the sword."'(before l47l Malory, Morte Arthur, 48 [Jucker 1997: 99])

(6) Moyses was a wonderful felowe, and dyd his dutie being a maried man. We lacke such as Moyses was. **Well**, I woulde al men would loke to their dutie, as God hath called them, and we then shoulde haue a flourisyng Christian weal. (1549 Latimer, 29 [Jucker 1997: 100])

(7) Tom: Yes, you must keep a Maid, but it is not fit she should know of her Masters privacies. I say you must do these things your self.
Ione: **Well** if it must be so it must. (1684 Tom the Taylor, 268 [Jucker 1997: 102]

Though many of the past studies analyze *well* through the lens of different theoretical approaches[21], most of them acknowledge the existence of both textual (or even sometimes 'coherence') functions and interpersonal functions.[22] The interpersonal functions and especially those associated with indirect or dispreferred responses are generally found to be the most frequent uses of *well*, while the textual functions, which include hesitations, self-correction, word-searching and shifting to a quotation, are rarer. This preference to use *well* as a response marker also explains why the most frequent position for *well* is at the beginning of the sentence/utterance. In fact, De Klerk (2005:1190) regards the initial position as the unmarked one for *well* because "this tendency of *well* to precede utterances could generally be regarded as natural, since discourse markers typically act as a guide to addressees as to how to react to what is about to be said, rather than acting retrospectively on what has already been said". When

[21] *Well* has been used as the testing ground for various approaches to DMs including Politeness Theory (Watts 1989), Coherence Theory (Schiffrin 1987), Relevance Theory (Jucker 1993; Blakemore 2002; de Klerk 2005) and Meaning Potentials Theory (Aijmer 2013, 2015).

[22] See discussion in Müller's (2004), Aijmer (2013) and Beeching's (2016) chapters on *well*.

it comes to the textual functions, the position of *well* can vary, though it has been found that pausing/hesitation and word-searching functions are particularly likely to be found sentence-medially.

Until recently, the vast majority of studies on *well* have focused on how it is used in informal speech situations between two or more participants. The reason for this is undoubtedly that these environments are the most prototypical. In an investigation of the use of *well* from a variety of text/conversation types, Aijmer (2013) indeed found that "(w)ell was most frequent in conversation (face-to-face and especially telephone conversation), that is, in more informal types of situations characterized by intimacy and close relationship between participants" (*ibid*.: 27). However, this concentration of the study of *well* in informal conversation gives an incomplete picture of how *well* is used overall. It is for these reasons mentioned above that Aijmer (2013; 2015) stresses the importance of looking at the uses of *well* in various text types and coming up with a theory that is compatible with all uses of *well*.

This idea of *well's* meaning going beyond the discourse context alone also finds support in Norrick (2001:866), where it is suggested that *well*, along with *but*, have specialized functions in the different types of discourse. Norrick's study pinpointed that *well*'s specialized functions in storytelling were highly connected to conventions of highly coded sequentiality *(ibid*.). Other studies that have observed *well* in other discourse types have also found specialized functions of *well* that are not necessarily activated in informal discourse. For example, Innes (2010) investigates how *well* is used in New Zealand courtrooms and suggests that *well* provides not only contextual coordinates but also indicates participant responses and attitudes (2010: 114). Also, in an investigation of DM usage in interviews, Fuller (2003: 44) found that the DMs *well* and *oh* were used more frequently by the interviewer than by the interviewee and also that, "DMs can be viewed as particles which are employed in a variety of speech genres, but show

distinctions in frequency and function that correlate with the speaker's role as either an interviewee or an interlocutor in an informal, symmetrical interaction". Using a more unique source of data, Greasley (1994) looked at the usages of *well* in television commentaries during snooker games. In this specific text type, *well* is used to "address the positional consequences of an error in the play, address a positional problem of apparent concern, remark upon a surprisingly good shot and comment upon an unexpected observation" (*ibid*.: 477). When considering the uses of *well* in these snooker commentaries, Greasley found that earlier studies of *well* that were based on conversation data were incomplete and inconsistent in dealing with the uses of *well* in his data. However, in order to explain the specialized functions in his data, Greasley suggested there may be similarities between the functions of *well* in everyday discourse and those that he found in the snooker commentaries. He argues that these functions of everyday discourse have been expanded upon and have led to the development of functions that are specific to the snooker television commentaries. One such example is that problematic situations in a snooker game that require the commenter to modify his/her speech quickly is similar to a speaker modifying his/her speech in self-correction situations that have been seen in conversational data (Greasley 1994: 493).

Moving on to examples that are more relevant to the discussion about the use of *well* in written texts, Rühlemann & Hilpert (2017) took data from the TIME corpus and demonstrated that there was a major increase of usage of *well* in the 1990s. A large reason for this increase was the sudden spike in usage of *well* in the sentence-medial positions. In this position, *well* took off from the 1990s acting both as analytical marker flagging clausal structure and as a word choice marker indicating playful language, as found in (8) and (9) respectively:

(8) [AdvCl once you've seen Walking with Dinosaurs (Discovery, April 16, 7 p.m. E.T.)], **well**, [mainCl [Subject you] still won't have seen real animals do any of that]. (TIME 2000) (Taken from Rühlemann & Hilpert 2017: 121)

(9) After all, [Subject Soviets] [Copula are], **well**, [Subj.Predicative Soviets]. (TIME1987) (Taken from Rühlemann & Hilpert 2017: 124)

Rühlemann & Hilpert (2017) is an important influence on this work because they offer one of the few studies that observe DM *well* in the realm of journalistic writing. Furthermore, their article discusses three of the four major sentence-medial functions found in this investigation of *well*. This includes quote *well*, clause *well*, and predicative *well*. In their investigation of sentence-medial *well*, they argue that the usage of *well* at the beginning of quotations (quote *well*) highlights the increase of interviews and the insertion of direct speech into journalistic writing, while the increased usage of clause *well* and predicative *well* suggests that the DM *well* is also entering into new roles such as flagging clausal structure and functioning as a word choice marker.

Such uses of non-dialogical *well* point out that *well* is being increasingly used in innovative ways in writing (as is particularly the case where *well* is used as a word-choosing marker). They also display new ways in which *well*, a prototypically dialogical DM (see Aijmer 2013: 26), has increasingly been recruited to display the authors' attitudes toward the information being discussed in their own text (subjectification) as well as to display the author's attention to the reader's potential or imagined reactions to texts (intersubjectification) (e.g. Traugott 2010).

These past studies show overall that *well* is highly multifunctional and that a large number of contextual elements contribute to its specific functions in particular discourse/text types. What relates all of these various functions together, however, is the central role that is played by the speaker/writer and his/her relation to the intended addressee/reader. When it comes to the recent changes found in Rühlemann & Hilpert (2017)'s paper, one factor that has clearly shifted significantly since the 1990s is how the author positions his/her relationship with the reader(s). In the more informal realms of writing, but journalism in particular, from the later part of the

20th century Leech *et al*. (2009:239) have noted "a definite trend away from the cool distancing of traditional written style and towards a kind of spontaneous directness". Looking back at the journalism examples from Rühlemann & Hilpert's (2017) study, these non-dialogical uses of *well* that rhetorically stress the author's subjective attitudes or that directly address the readership in hedge functions or alignment seeking functions appear to represent one area where *well* is quickly evolving in written journalistic texts. Moreover, the fact that these developments appear to be growing out of sentence-medial *well* also indicates that this change is growing out of what is the more salient yet overlooked usage of *well.* In the following sections of this chapter, these developments will be analyzed in greater detail.

3.3 Frequency trends of non-dialogical *well* in the COHA

In order to extract *well* in the DM functions as opposed to those that are adverbial and nominal, I used a search string similar to that adopted by Rühlemann & Hilpert (2017). This search string includes a series of punctuation signs immediately before and after *well*. This is based on the common practice of writers who offset DMs from the neighboring co-text by typographic means (Rühlemann & Hilpert 2017: 109).[23] This search string allowed me to weed out the many non-DM variants of *well* such as the adverbial and nominal variants, examples of which are provided below:

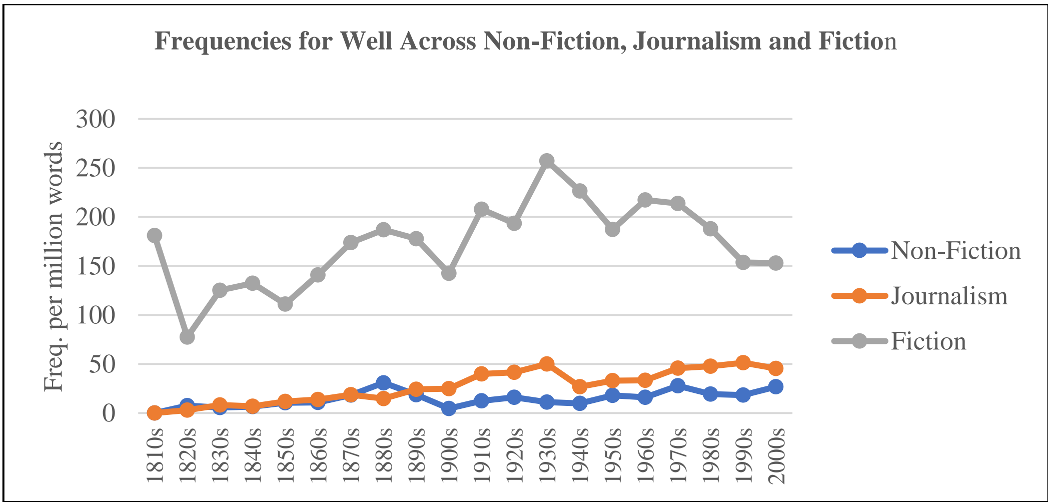
(10) Adverbial Well
DR. KADIDI (Full of himself). As you can see, I am doing quite **well**. My only problem is that I have a lot of work on my hands. (COHA Fiction: Chains of Junkdom 2001)

(11) Nominal Function
The hovercar descended through a **well** in the central courtyard that wits called the Navel of the Earth. (COHA Fiction: Fantasy SciFi 2000)

[23] |(|.|,|:|?|!|... well .|,|:|?|!|...|)|. Similar search strings were used for all DMs in this study for the same reasons

Once extracted, I annotated the concordances for their position in the sentence, their primary function[24], the type of article in which it was found and any other distinguishing properties. Moreover, all cases of nouns and adverbs that the search string did not weed out were removed manually from the data.

When we analyze the usage of *well* in all the COHA sub-genres on Figure 3.1, it is clear that the different sub-genres are not all behaving in the same way from the 19th to the early 21st century. Furthermore, the only sub genres to display a strong positive correlation in terms of its diachronic development (see Kendall's τ score) is journalism. Non-Fiction also shows signs of progression though not nearly as strong.

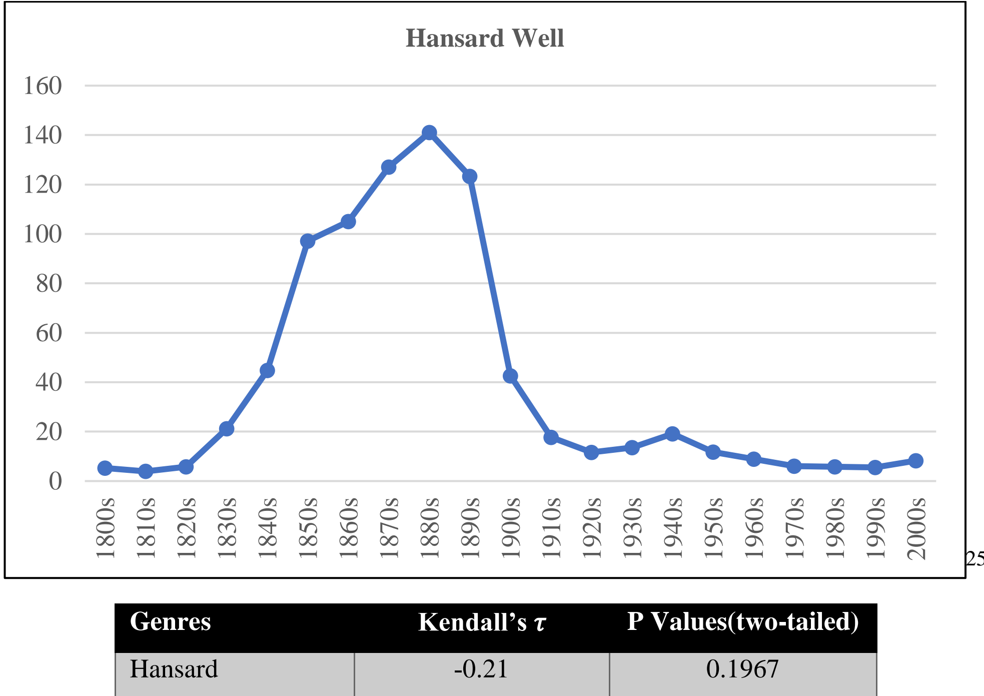


Genres	Kendall's τ	P Value (two-tailed)
Non-Fiction	0.51	0.001427
Journalism	0.82	4.61E-09
Fiction	0.41	0.0111

Figure 3.1: Diachronic Frequencies for COHA Data of *well*

[24] To determine the function of *well* in the individual concordances, the context in which *well* was used in the primary determinant. The fact that the concordances were retrieved based on surrounding punctuation also helped in determining the function as this property is very often a major indicator of how *well* is used.

The signs of something developing in journalism and to a lesser extent non-fiction can be contrasted with oral data from the Hansard Corpus (Alexander & Davies 2015), which shows the usage of *well* acting very differently. In fact, after the steep increase in the middle to late 19th century, there is a sharp decrease in the 20th century.



[25]

Genres	Kendall's τ	P Values(two-tailed)
Hansard	-0.21	0.1967

Figure 3.2: Diachronic Frequencies for Hansard *well*

So far, it has been observed that *well* is used differently in different genres of writing. It is clearly used the most often in fictional writing, though it is also clear that its overall diachronic usage is not showing the same developments as in the other two written genres. In terms of a clear positive diachronic trend, *well* has the highest positive correlation for increased usage over the past decades in journalistic writing than any other. It is used the least in non-fiction, though it is showing some weak signs of an increase in usage over the 20th century.

[25] The Hansard is consulted for comparison with *well* in this chapter due to its high degree of varying behavior in various genres and its differing behavior when used in different syntactic positions, including formal parliament speech data. The Hansard data is not discussed in the chapters on *like* and *oh yeah* due to their virtual absence of use as DMs in this corpus.

This observation of the overall usage in *well* in the COHA sub-corpora, however, is only giving us a vague idea of how *well* is developing. It has already been observed in Rühlemann & Hilpert (2017:118) that many of the changes that pushed the increase in usage of *well* in the Time corpus came from how it was used in sentence-medial positions. This included its usage between syntactic boundaries and as a word-choosing marker, respectively. In order to check if similar behaviors are occurring with *well* in the COHA journalism data, a collexeme analysis[26] (see Table 3.1) was conducted on this data set[27] from the 1990s, as this was the crucial period of change found in Rühlemann & Hilpert (2017). Since *well* can be used in a variety of functions in the sentence medial position, the context window for the analysis was set to two spaces left using the online COHA collocate search application. This context window can potentially catch indications of the use of predicative *well* should this also be used in the COHA data. It is also capable of picking up collocates that could potentially show signs of *well* being used in other medial functions such as verbs indicating quote *well* or even indications that *well* is being used at the boundaries of clause structures.

[26] See the discussion on the collexeme in the introduction chapter section 1.7

[27] As with the collocate searches, the punctuation around the DM is considered part of the construction in the collexme analysis. The same procedure was done for collexeme analysis of the other two DMs in the upcoming chapters.

	COLLEX	ASSOC	COLL.STR.FYE	SIGNIF
1	A	attr	12.35414	*****
2	YEAH	attr	11.53126	*****
3	OH	attr	7.81997	*****
4	SAY	attr	5.87122	*****
5	LIKE	attr	4.30434	****
6	BECAUSE	attr	2.91875	**
7	YOU	attr	2.58761	**
8	THAT	attr	2.56451	**
9	DO	attr	2.55701	**
10	I	attr	2.55451	**
11	YEAR	attr	2.51241	**
12	OF	attr	2.4157	**
13	IS	attr	2.41104	**
14	ALL	attr	2.06249	**
15	THIS	attr	2.01934	**
16	IT	attr	1.81675	*
17	ABOUT	attr	1.71274	*
18	MORE	attr	1.70489	*
19	BUT	attr	1.49497	*
20	BE	attr	1.04119	ns
21	ARE	attr	0.91317	ns
22	TO	attr	0.75785	ns
23	FOR	attr	0.75677	ns
24	‘S	attr	0.57347	ns
25	AND	rep	0.61532	ns

Table 3.1: Collexeme Analysis of Journalism in COHA Well 2L 1990s and 2000s

In this collexeme analysis, there are several indications that *well* is strongly attracted to collocates that indicate it being used in sentence-medial positions. This includes the form *is* of the copular verb *be*. Three other forms of the *be* appear in Table 3.1, but they are not statistically significant. This suggests that the COHA data also involves *well* occurring in the predicative construction (see example 12), and coincidentally this proves to be the case as section 3.4.3 will show. Other examples that indicate action in the sentence-medial positions involve words that indicate that *well* is being used within sentences with complex clausal structures such as

because, *and, say*, and *that*. Specific examples (see example 13) of *well* being used between clausal structure (clause *well*) will be analyzed in section 3.4.2.

(12) How CLA might do that is, well, complicated. There seem to be several mechanisms. (COHA Magazines: Forbes 1990)

(13) His immune system doesn't attack his new kidney because, well, it's not exactly his immune system anymore. (COHA Magazines: Men’s Health 2006)

These indications of *well* being used at high levels within the sentence structure (predicative *well* and clause *well*) is an intriguing result because it gives further evidence to the fact that when used in journalistic writing, *well* is developing functions in areas outside of its typical dialogical uses. This is in contrast to what has been observed in natural spontaneous speech where the sentence-initial position is the most typical and natural environment for it. In spontaneous speech, *well* is most often used as a ‘deliberation signal’ (Biber *et al.* 1999: 1086) or a guiding marker that signals to the addressee how they should react to what is about to be said (De Klerk 2005: 1190). *Well*’s usages in the medial positions of the sentence are much more rarely discussed. In terms of how *well* is used in sentence-medial position, however, Aijmer (2013) states that its functions are primarily related to pausing or word-search (28). Considering these various indications of *well* developing major functions within the predicative construction and in-between clausal boundaries, the focus of attention will now shift to *well* occurring sentence-medially.

Though there was an overall positive correlation of the usage of *well* in writing from the late 20th century on, Rühlemann & Hilpert (2017) and my own collexeme analysis point to developments occurring in the sentence-medial positions. Furthermore, as will be shown below, it is in these positions that the functions of *well* were changing and developing in a much more uniform way than *well* in any other position. This can be observed if one compares the sentence-

initial examples (14) and (15) below to examples (16) - (19), which encompass the four major groupings of the functions of non-dialogical sentence-medial *well*. These four function groupings will be discussed in more detail in the next section.

Sentence-Initial Examples

(14) Remember Duhamel back in 2004? No? **Well**, we're willing to bet that your wife or girlfriend probably does: As the eponymous lust object in the beefcake chick flick *Win a Date with Tad Hamilton!* Duhamel grinned his way into the spotlight, chopping wood sans shirt and outshining his romantic foil, Topher Grace--the putative everyman. (COHA Magazines: Men's Health 2006)

(15) Early last winter I was walking along a creek and jump-shooting ducks when a coyote ran past me, carrying the spine of a fawn deer in its mouth. **Well**, if the coyotes weren't going to play fair, then neither was I. (COHA Magazines: Outdoor Life 2006)

Sentence-Medial Examples

(16) Quote Well
She said, **well**, yes, they did smoke pot. (COHA Magazines: Saturday Evening Post 1984)

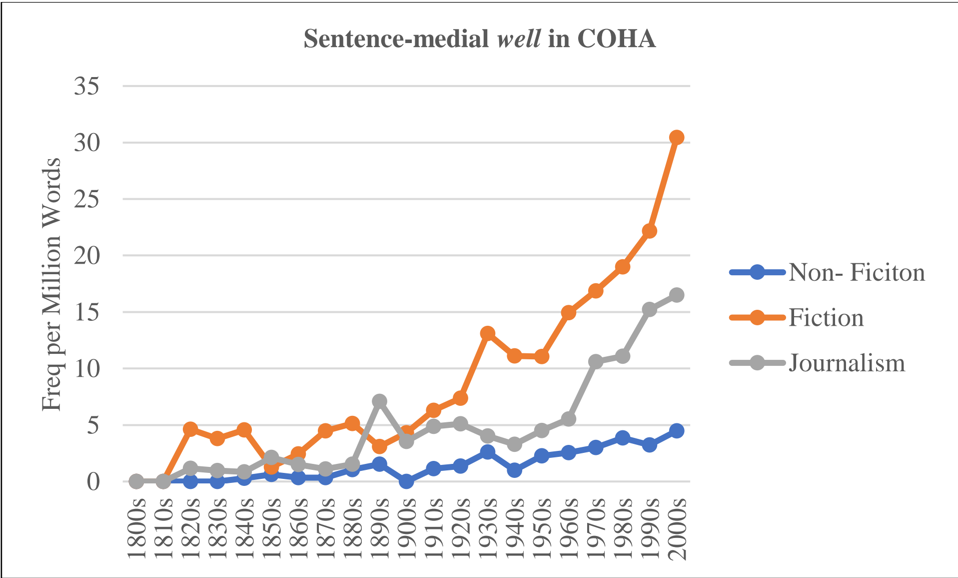
(17) Clause Well
The fear of MS is just terrible. When you lay down in bed at night, **well**, I know those nightmares. (COHA Magazines: Ms. 2006)

(18) Predicative Well
New ads feature punchy lines such as: "Until I find a real man, I'll settle for a real smoke" and "I get enough bull at work. I don't need to smoke it. Anti-tobacco activists say the new product image is, **well**, bull. (COHA Newspapers: Washington Post 1997)

(19) WCM Well
We dove in like, **well**, a wild bunch, hooting and howling. (COHA Magazines: Skiing1995)

Coming back to examples (14) and (15), though they also are used in writing for rhetorical purposes, they also still follow the regular conventions of dialogical *well*. In (14), for example, the author uses the convention of using *well* as a turn taking marker that marks a transition to additional information after the invented exchange between him/herself and the readers who

supposedly did not remember the actor in question from 2004.[28] In a similar fashion, example (15) functions as a subjective contextualization cue (e.g. Aijmer 2013: 35) in which the author explains a situation in the preceding sentence and then begins the next sentence with *well*, which marks the author's attitude or excuse for not hunting the coyote in a "fair" manner. Though both of these uses in the sentence-initial position are also great examples of how authors can use *well* (inter)subjectively in articles in order to directly engage with their readership, they are also used in ways that are carbon copies of how *well* can be used in speech. Furthermore, sentence-initial *well* is not enjoying the same salient increase in usage and conventionalization from the late 20th century onward (see Figure 3 for the diachronic development of sentence-medial *well*). It is for this reason that the remainder of this chapter will be concerned with the discussion of the trends and functions of sentence-medial well.

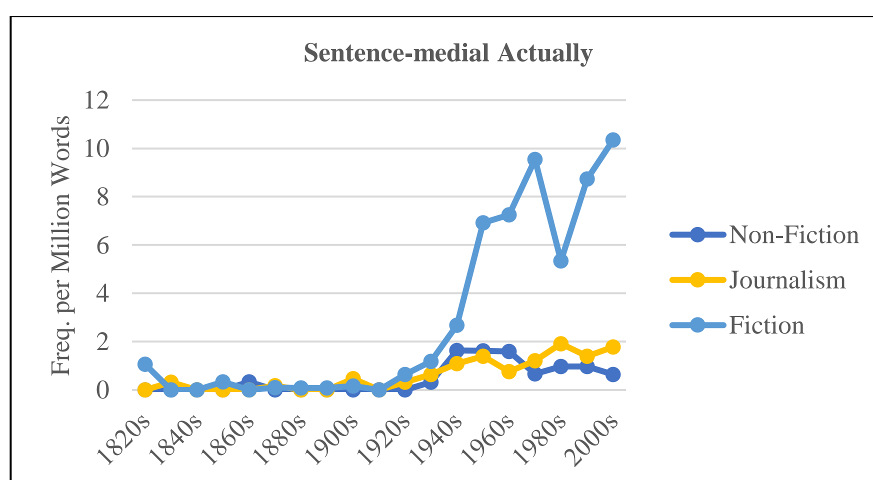


Genres	Kendall's τ	P (two-tailed)
Non-Fiction	0.79	1.408X10-6
Journalism	0.76	1.796X10-7
Fiction	0.77	1.35X10-7

Figure 3.3: Diachronic Frequencies for COHA Data of sentence-medial *well*

[28] See the practical writing guide for the DM *well* from the Cambridge Dictionary's *English Grammar Today* at the following link: <https://dictionary.cambridge.org/grammar/british-grammar/discourse-markers/well>

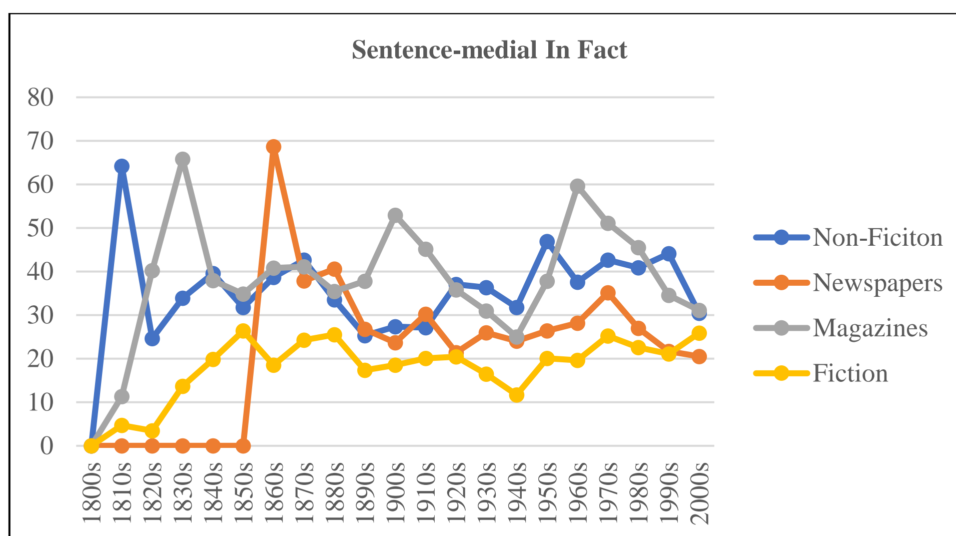
When looking at the diachronic usage and functions of sentence-medial *well* in the COHA data, some trends are easily observable. First, it is clear that its usage is a recent phenomenon that steadily rises in the later 20th century. This is the case of all of the sub-corpora of the COHA data. Furthermore, this development of medial *well* at the end of the 20th century can be distinguished from other DMs that can also serve similar functions in the sentence-medial position, namely *in fact* and *actually*. According to the Cambridge English Dictionary's webpage, *in fact* and *actually* serve many of the same functions as *well* when it comes to showing attitude in oral speech as well as writing.[29] Furthermore, it was observed in the COHA corpus that *in fact* and *actually* were also used in the sentence-medial position in similar functions. However, as can be seen in Figures 3.4 and 3.5 below, these two DMs do not display the similar developments at the tail end of the 20th century as we see with *well*. To begin with, *actually* is used significantly less often than *well* in writing. Moreover, *in fact*, though used far more often than *well* in writing, shows minimal positive development over the 20th century.



Genres	Kendall's τ	P (two-tailed)
Non-Fiction	0.41	0.04
Journalism	0.71	8.00E-05
Fiction	0.67	7.05E-05

Figure 3.4: Diachronic Frequencies for COHA Data of sentence internal *actually*

[29] See Discourse Markers (so, right, okay) at Cambridge Dictionary online via the following link: <https://dictionary.cambridge.org/grammar/british-grammar/discourse-markers-so-right-okay>



Genres	Kendall's τ	P (two-tailed)
Non-Fiction	0.1157895	0.5006
Newspapers	-0.3904762	0.04629
Magazines	-0.031579	8.73E-01
Fiction	0.3263158	4.68E-02

Figure 3.5[30]: Diachronic Frequencies for COHA Data of sentence internal *in fact*

Sentence-medial *well* is further distinguished from the other two DMs when it comes to how it is used in the sentence-medial position in writing. From the 1990s, authors of journalistic texts rapidly adopt functions that indicate the specific selection of lexical items. Therefore, the rest of this chapter will concentrate on the uses of non-dialogical sentence-medial *well* in journalism, though the non-fiction and fiction data will be brought up for comparison purposes.

In Figure 3.6 below, the diachronic frequency trends of non-dialogical sentence-medial *well* for journalism only are further analyzed using the VNC algorithm[31]. With the help of a scree plot (not shown), the decision was made to separate the data into four stages. In the first two stages, there is little indication of a clear-cut positive correlation. However, in the second half of the 20th century, from 1970s in particular, there is a sudden jump in usage (see stage 3 1970s-

[30] Because of the higher frequency of examples and the different behavior of *in fact*, the newspapers and magazines sub-corpora were analyzed separately instead of combined as is usual.
[31] See discussion of the VNC in chapter 2 section 2.7 for details.

1980s). This is followed by another sudden leap in usage in stage 4 (1990-2000s) that also includes the explosion in the usage of *well* in the before mentioned word-choosing functions.

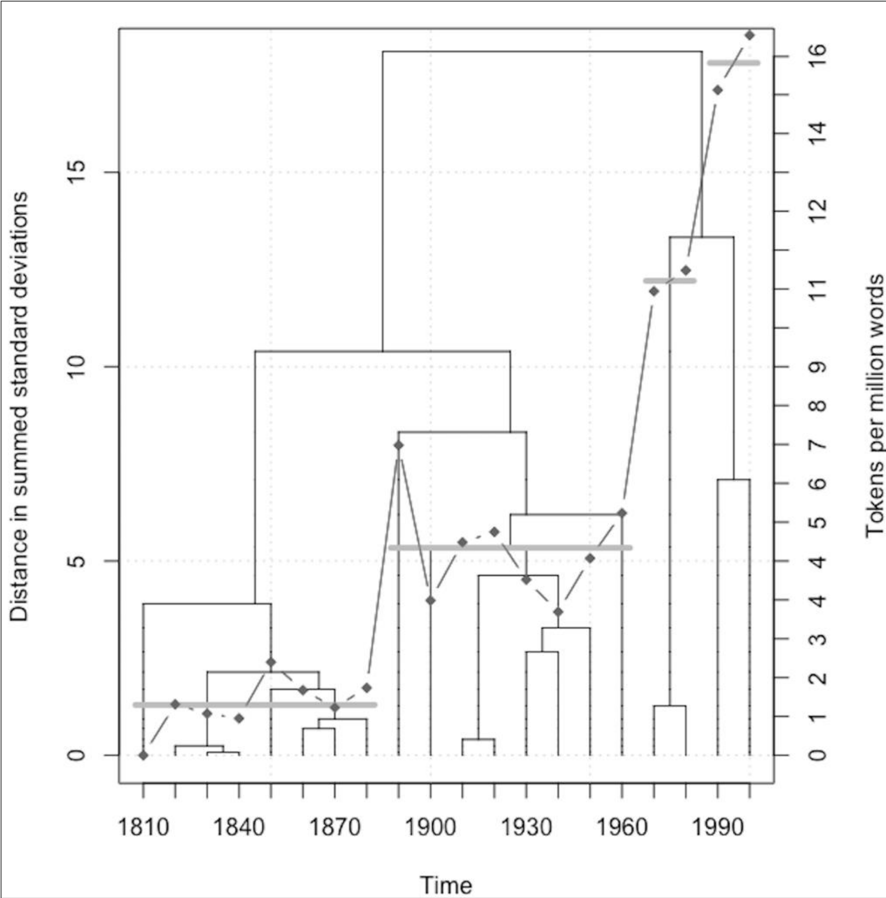


Figure 3.6: VNC Dendrogram with overlaid line plot for frequencies of DM *well* for periods 1(1810s-1880s), 2 (1890s-1960s), 3 (1970s-1980s) and 4 (1990s-2000s) in Journalism COHA data

Before period 3, which is the period where the word-choosing functions show signs of growth, the quote and the clause marking functions were still the primary functions of sentence-medial *well.* Interestingly, in period 4 the proportions of quote *well* and clause *well* drop, while the two word-choosing variants of medial *well* become dominant. This shift in medial *well* becoming more and more a marker of word choice can be seen in Figure 3.7 below, which displays the development of the sentence-medial functions in the 20th and first decade of the 21st century for the journalism data. This sudden rise is pushed especially by *well* occurring in the predicative construction, which will be discussed in detail in section 3.4.3.

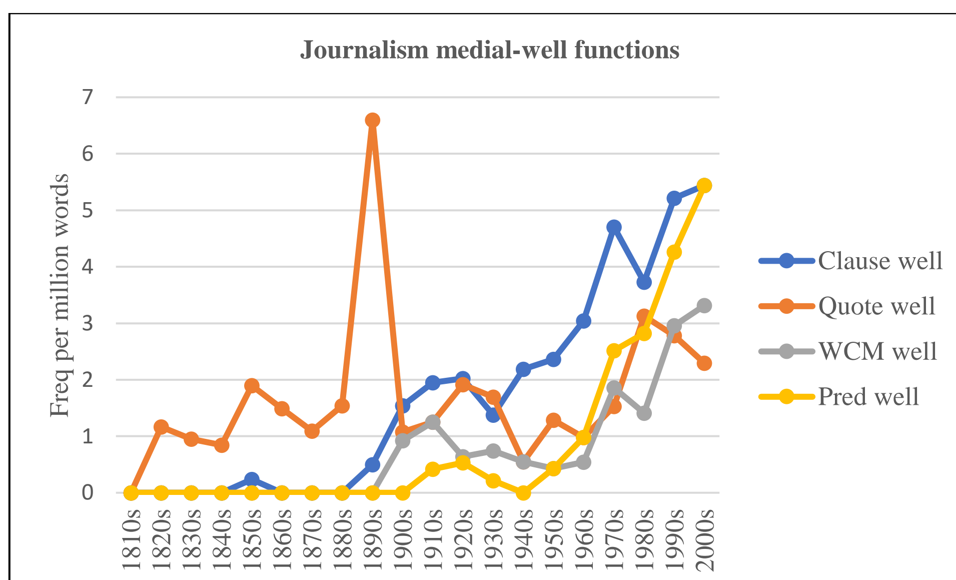


Figure 3.7: Functional Categories of non-dialogical sentence-medial *well* in journalism by decade

This rise of the two-word choice functions, and predicative well (pred *well*) in particular, of sentence-medial *well* is what also what separates the journalism data from the other two sub-corpora in the COHA. Though sentence-medial *well* in non-fiction and fiction also show a significant rise in usage from the second half of the 20th century, the sudden rise in usage of word-choosing functions does not occur at the end of the 20th century. This can be observed in Figures 3.8 and 3.9 that also show the proportions of the major functions of sentence-medial *well* by decade. With the fictional data in 3.8, a 1,000-word sample was taken from the pattern ', well,'. Due to the fact that many of the concordances of this pattern were in fact part of a DM sequence at the beginning of the sentence, only 388 concordances from this sample were actually true sentence-medial examples.

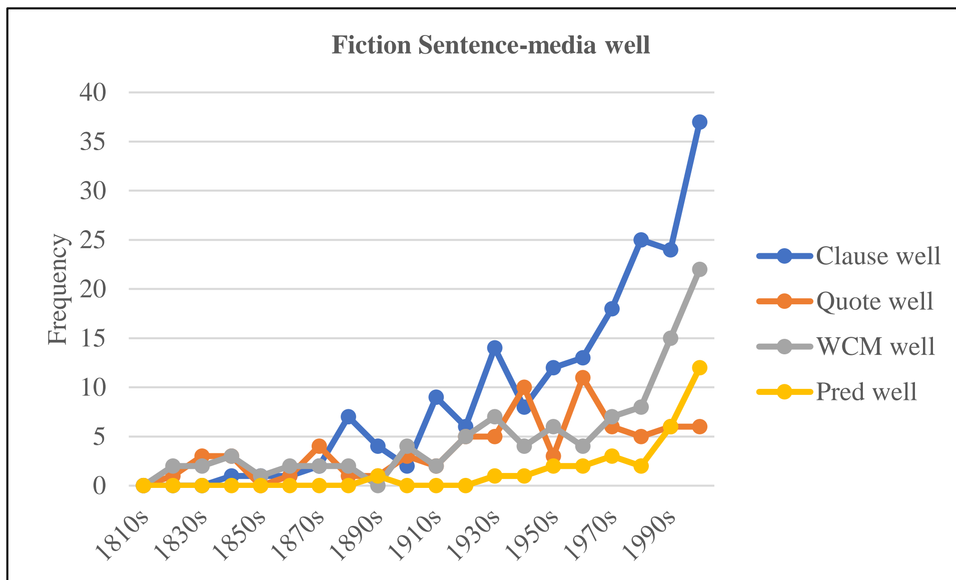


Figure 3.8: Functional Categories of non-dialogical sentence internal *well* in Fiction by Decade based on 1,000 sample of pattern ‘, well ,’

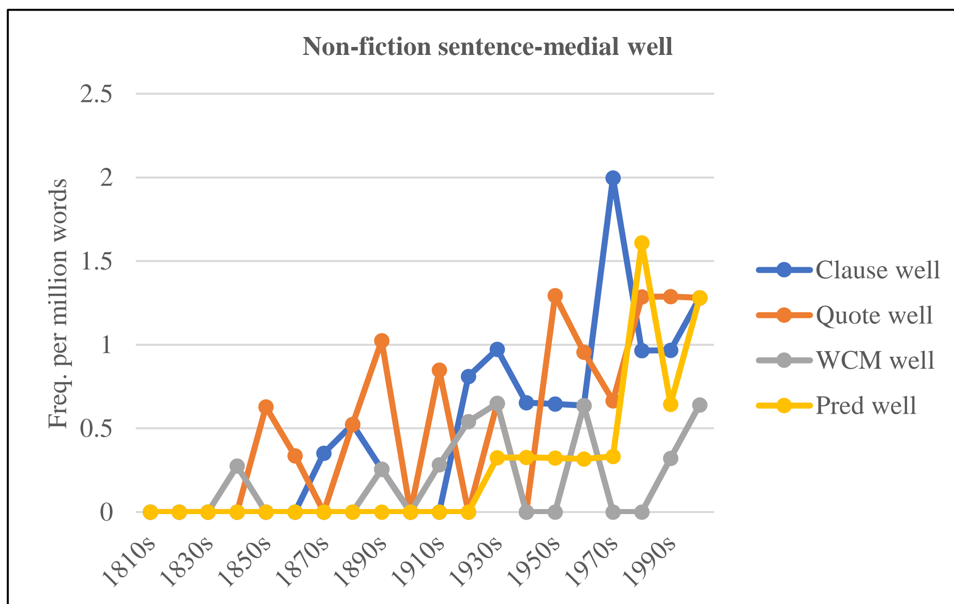


Figure 3.9: Functional Categories of non-dialogical sentence-medial *well* in Non-Fiction by Decade

This section has demonstrated that the usage of *well* in multiple forms of writing is indeed on the rise since the second half of the 20th century and into the 21st century. Furthermore, it has highlighted that this rise is particularly apparent in the sentence-medial positions in writing. This indicates that there was a change in the development of *well* outside of its more prototypical dialogical uses in spontaneous speech where its “tendency to precede utterances could generally be regarded as natural” (de Klerk 2005: 1190). This new development is

particularly strong in journalism. Perhaps because it is so much more salient in the sentence-medial position, authors began using *well* rhetorically in this position from the second half of the 20th century in a variety of functions before it became most popular as a word-choosing marker. Furthermore, these uses of *well* as word-choosing markers undoubtedly can be seen as an outgrowth of the dialogical functions of *well* marking a pause or the author performing a word-search, as has already been attested as one of the rarer uses of *well* in oral discourse (e.g. Aijmer 2013: 28; Rühlemann & Hilpert 2017: 130).

This section has also shown that this substantial change in how *well* is used in new non-dialogical functions takes place at the very end of the 20th century and into the 21st. The major theoretical implication of this in colloquialization studies is that in addition to the changes in the acceptable writing style that began in the second half of the 20th century, there also appears to be a second wave that begins in the 1990s. This essential issue of the 'second wave' will be discussed in detail in the conclusions of the chapter in 3.8.

3.4 Qualitative Analysis of non-dialogical *well* in the COHA data

The last section has gone over the overall frequency trends of sentence-medial *well* and the changes in usage of its four major functions. This section will give a detailed analysis of each of the four major functions of sentence-medial *well* in the journalism sub-corpora of the COHA data. This section begins with the two functions that appear to have similar equivalents in spontaneous oral speech: quote *well* and clause *well.* As discussed above, these functions have a longer history in being used in writing. Furthermore, due to them having direct descendants in speech, their usage in writing is very similar to their equivalents in oral speech. Furthermore, unlike the two categories of word-choosing markers, quote *well* and clause *well* are found in all of the four COHA sub-corpora and even oral Hansard data. Therefore, the bulk of the analysis in this section will be on the usage of *well* in sentence-medial positions that appear not only

much later in writing but are also the functions that are the most unique to writing. This includes pred *well* and the more general word-choosing marker (WCM) *well*. The particularities of these two uses of *well* will be analyzed in detail 3.4.3.

3.4.1 Quote *Well*

Transitioning into a quotation has been discussed in other studies of *well* involving spontaneous oral data. In Aijmer (2013:36), *well* is described as being able to "introduce or 'frame' a quotation". When used in the COHA data, it appears that quote *well* is used very much in the same fashion in writing as it was in studies of informal speech. This is most likely because these examples come from interview quotations in the vast majority of cases. However, even in the rare cases where quote *well* is used in non-dialogical situations, their execution is still identical to their equivalents in quotations. This is evident if one compares the two examples below. Examples (20) and (21) look very similar despite the fact that one is from a transcribed speech (see 20) while the other is from journalistic prose (see 21).

(20) The Financial Secretary wriggled: He said, "**Ah, well**, of course there are all kinds of economic and monetary union:" He nods, but perhaps he is tired. (Hansard: House of Commons 1993)

(21) But I will say that marriage and kids to me, **well**, that's like real life, man. (COHA Newspapers: Chicago Tribune 2000)

Furthermore, in this discussion on the usage of sentence-medial *well* in writing, quote *well* is also the function that demonstrates essentially no innovation in terms of having distinctively written executions. As found in other studies (Aijmer 2013:37; Rühlemann & Hilpert 2017), quote *well* has been found to have two major realizations: transitioning into actual quote (example 22) and presenting a change to a 'quoting voice' in which the speaker constructs a hypothetical 'other', or performs a 'self-quote' (example 23).

(22) He sat down on the ground next to the plugs and said: Look, this is Healey. You told him that you were gonna listen to him. Right?' And they said, **Well**, we can't, Bill, the show's on.' He said: I'll tell you what - we can. (COHA Magazines: Rolling Stone 1991)

(23) People say, **well**, the stadium is always filled. (COHA Newspapers: Washington Post 1992)

As these two examples demonstrate, *well* as a quote marker functions as a flagging device that is designed to mark the fact that the speaker is either transitioning into the dialogue of another speaker, or uses it as a prose device designed to create a conversation-like rapport (e.g. Rühlemann & Hilpert 2017: 128). What is interesting about the use of quote *well* is that it is used both in quotations/transcriptions as well as in monologue texts. In written monologue texts, unlike in speech, the need of the DM coherence function of using *well* is greatly reduced because of the abundance of typological coherence tools such as exclamation marks, colons, and commas (*ibid.*). Therefore, the decision of the author to include this typically oral function lies most likely in the usefulness of either making a fabricated quote look more real and/or making the text more engaging and conversational-like. Such uses are neither surprising nor innovative since the inclusion of interviews or dialogue sequences have also been found to be on the rise in all written text genres since the second half of the 20th century, particularly in journalistic writing and fiction as has already been seen in other studies on colloquialization.

Section 3.2 showed that a variety of co-DMs and interjections were found to be attracted to *well.* As it turns out, it is in this category that such DM and interjection sequencing is most often used.[32] In the COHA data, quote *well* frequently co-occurred with other DMs and interjections such as in: '*oh well', 'ah well',' well well',* and *'yeah well'*. Curiously, however, this DM/interjection sequencing was actually much more common in the Hansard Corpus than

[32] See Lohmann & Koops (2016) for a full discussion on DM sequencing.

it was in the COHA, thus giving further indication that DM sequences involving *well* are more common in oral dialogue than written English.

Due to the similar findings between the written and speech data, coupled with the fact that quote *well* is losing ground to the newer word choice marking functions in writing, quote *well* in writing is best seen as an early use of *well* journalistic writing, but not one that indicates any change to *well*'s meaning or functionality. Unlike the examples that follow, the usage of quote *well* is found in the early 19th century and its frequency of use varies throughout the progression of time. This suggests that this convention of using quote *well* may go back even further than that. Furthermore, this function is heavily connected to the ever-increasing use of interview and dialogue sequences in written texts. This does indeed point to significant stylistic changes in written English (e.g. Mair 2006: 188). However, going as far as to indicate that the elements of these dialogue sequences are signs of oral elements being used in writing itself is debatable, since this is more often an indication of the author attempting to mimic speech within a written text rather than making his/her writing style as such more colloquial-like.

3.4.2 Clause *Well*

Similar to the situation with quote *well*, clause *well* is also well represented in all of the sub-corpora of COHA. In addition, like quote *well*, it loses its dominance in the journalism data in the 1990s and 2000s while the uses of pred *well* and WCM *well* increase. Furthermore, the usage of clause *well* is also found in the Hansard Data, where it looks and functions in exactly the same way (compare examples 24 from the Hansard with 25 from the COHA).

(24) And when they are told that all this other elaborate machinery has to be gone through before anything is decided, **well**, they have their doubts (Hansard: House of Lords 1963)

(25) When the kid picked himself up off the ground after his John Daly -- esque swipe and asked where the ball had gone, **well**, my brother and I were pointing at a rocket zooming high on the horizon toward the 3rd tee, which sat directly behind the 5th green. (COHA Magazines: Sports Illustrated 2004)

As its name implies, clause *well* occurs at the clause structure boundaries. Furthermore, this study finds that clause *well* has a strong preference to be used in positions right before the occurrence of the main clause. By far the single most common clause *well* pattern involves adverbial clauses. In these cases, the main clause is separated from the dependent clause by *well*. More specifically, the prototypical realization of this pattern involves the subordinate *if*-clause followed by *well* which is followed by the main clause. Example (26) below gives a prototypical example of this. The co-occurrences of *if* clauses with clause *well* alone accounted for roughly 40% of all cases of clause *well* in journalism and over 72% (88% for all adverbial clauses combined) in the Hansard data. This preference to use *well* in-between if-clauses and main clauses in both the oral and written data suggest that this is the prototypical exemplar of clause *well*.

(26) This is the worst of all possible times to cut taxes, but Congress is bent on it, mostly for the rich and middle incomes. If the unpatriotic ghetto rebels, **well**, there's always repression. (COHA Magazine: New Republic 1969)

Moreover, besides the obvious example of *well* occurring in-between an if-clause and the main clause, there were also examples involving the when-clauses and because-clauses. A second major category involved *well* being placed in-between two main clauses, with *well* occurring after the appropriate coordinating conjunction. The third category involves *well* marking the border between non-finite clauses, and the main clauses. Finally, the last major category involves well marking topicalization. In these examples, the NP isolated to the left of *well* is followed by the main clause. In order to best illustrate these common patterns of clause *well*, consider examples (27)-(31) below that present the patterns in the order in which they were discussed above. See also Table 3.2 for the respected proportions of these same patterns.

(27) AdverbialCl-*well*-MainCl
In the 1978 voting, 63 women were elected as first-time legislators across the United States. Sixty-one of the 63 women belonged to the same political party. Which party was it? If you answered Democratic…**well**, you were wrong. The correct reply was the Republican Party - you know, the party with the "gender gap." (COHA Newspapers: CS Monitor 1983)

(28) MainCl-Subordinator-*well*-AdverbialCl
His experience with sex, pre- Sophia, was ridiculously limited, and she knew it because, **well**, he'd told her the truth. (COHA Fiction: Dark Night Novel 2009)

(29) MainCl-*well*-MainCl
They staked their own credibility on the success or failure of Reaganism, and, **well**, the negative evidence is mounting. (COHA Magazine: Washington Monthly 1993)

(30) Other Dependent Cl-*well*-MainCl
As for melted middles, **well**, they've been there, done that, this past spring and summer. (COHA Newspapers: Washington Post 1994)

(31) Topicalization
The top editors of the influential Globe live in distant, affluent suburbs, sanctuaries for whitey. And Judge Garrit, **well**, he lives in Wellesley, and his children never schooled with blacks. (COHA Magazines: Saturday Evening Post 1976)

Genres	AdvCl	MainCl-Well-MainCl	Non-Finite (all)	Topicalization	Other
Non-Fiction	40%	30%	6%	6%	18%
Journalism	55%	19%	11%	9%	6%
Fiction	38%	25%	13%	17%	7%

Table 3.2: Patterns of Clause Well

As the five examples above show, clause *well* has a significant preference to be to the left of the main clause, therefore also in close proximity to the subject constituent of the same main clause. Despite being found in both the COHA in non-dialogical and dialogical contexts, as well as in the Hansard corpus, there is little discussion of this specific function of *well* outside of Rühlemann & Hilpert's (2017) paper. This is likely due to the limited amount of discussion regarding *well* in the sentence-medial positions. However, the development of this specific sentence-medial function can be explained through the thetical grammar framework as explained in chapter 2 (cf. Kaltenböck *et al.* 2011) In the case of clause *well*, we see such a combination of the components of the situation of discourse. In particular, if we investigate the

examples involving ‘if clauses’, (see 32 below which is a copy of (26) we see a combination of the two components. The TO (text-organization) component of discourse is foregrounded in the sense that *well* highlights the transition from the dependent if clause to the main clause. In addition to the TO component, the AS (attitudes of the speaker) component is also foregrounded in these examples (*ibid*.: 861). As a DM, *well* often functions as a response marker that encodes the author’s (inter)subjective attitudes. When used between the ‘if clause’ and main clause, *well* carries some of these characteristics. Specifically, the author is responding to the dependent ‘if clause’ before moving on to the main clause, which entails what could be seen as a response – in this case a dispreferred, which the authors attempts to downplay.

(32) This is the worst of all possible times to cut taxes, but Congress is bent on it, mostly for the rich and middle incomes. If the unpatriotic ghetto rebels, **well**, there’s always repression. (COHA Magazine: New Republic 1969)

Clause *well*, like quote *well* has examples in the COHA data going back to the 19th century, though it appears to solidify in writing during the second half of the 20th century. Moreover, like quote *well*, it is used to mark a transition. In quote *well* it marks a transition into a quote, while in clause *well* it marks a transition from the dependent to the main clause (or vice vera). This section has also demonstrated that the usage of clause *well* in oral data (the Hansard Corpus) looks very similar to the examples found in writing. Rühlemann & Hilpert (2017: 18) investigated whether if-clauses that contained *well* were of greater length than a random subset of if-clauses without a usage of clause-well. The result was that there was no significant difference in length between the two groups (*ibid*: 128). However, despite these results they still concluded that the usage of *well* in these inter-clausal positions give “a convenient service for the reader to whom *well* announces the eventual arrival, after a long-winded adverbial prelude, at the sentence subject and its predication” (*ibid*: 128). Based on the findings of this study, it holds that the usage of *well* in these inter-clausal positions involving *if-clauses* is

indeed providing the reader a 'convenient' announcement of the arrival of the main clause to come, though this does not appear to be the main role of the usage of *well* in these sentences. If we consider another example (see 33) of *well* occurring to the right of the adverbial clause and just to the left of the subject of the main clause, it appears that the utilization of *well* is also used to rhetorically flag that the author's subjective attitudes in the main clause to come are being expressed.

(33) And if the other kids made fun of me, **well**, that was their problem. They're just jealous (COHA Non-Fiction: Promises to Keep 2007)

Though *well* does in fact clearly mark that transition from the *if-clause* to the main clause (the syntactical switch), the usage of *well* in this position is dominantly functioning on the subjective plane where it marks the author's subjective reaction to the actions of the other kids.[33]

Though clause *well* is strongly represented in all of the sub-genres of the COHA, it is also well represented in the sentence-medial Hansard data. This means that, like quote *well,* its usage in writing is not a recent innovation that helped push *well's* rise in usage in the 1990s. However, its function of combining the roles of marking the author's subjective attitude and signaling the syntactic border between the dependent and main clause is very likely a precursor to the later development of WCM and pred *well*. As will be shown in the next section, WCM and pred well shift from foregrounding primarily the TO and AS components of discourse to primarily foregrounding the AS and SHI (Speaker-hearer Interaction) components, with only a significantly weakened TO component remaining (cf. Kaltenböck *et al.* 2011: 862-3). Further evidence that clause *well* is the precursor for the two new functions of *well* comes from the fact

[33] It has been observed that *well* and other DMs for that matter can have several functions simultaneously. Schiffrin (1987) discusses how DMs can simultaneously address the participants (participant coordinates) and the text (textual coordinates). Biber et al. (1999: 1086) in their general definition of DMs state they simultaneously have two roles: "(a) to signal a transition in the evolving process of conversation, and (b) to signal an interactive relationship between speaker, hearer and message".

that, unlike quote *well*, which was almost exclusively used in quotations or transcriptions, clause *well* was used more often in non-dialogical contexts than in dialogical ones. In the journalism sentence-medial data, clause *well* is used 57% of the time in non-dialogical contexts.

3.4.3 Word-choosing marker *well* and predicative *well*

Having discussed the two categories of sentence-medial *well* that have carbon-copy oral equivalents and that are more generally found in all of the sub-corpora of COHA, it is now time to move on to the two major categories that are much more particular to journalistic writing. This includes pred *well* and the more general WCM *well*. Though pred *well* was indeed found in Rühlemann & Hilpert (2017), the other WCM *well* functions were not yet mentioned in other linguistic studies. What is interesting about both of these categories is that they appear to have the same origin in word-searching *well*. Word-searching *well* is used to hold the floor while the speakers considers how best to say something or where *well* is used to 'stall for time' while the speaker conducts a memory search (e.g. Allwood *et al.* 1990:11; Aijmer 2013: 32). Another sentence-medial function of dialogical *well* that these written variants sometimes resemble are the self-repair functions in which the speaker uses *well* to mark that they are revising an error that they previously made (Aijmer 2013: 33). Such dialogical examples where also found in the COHA data, though such functions came from interview and other dialogue sequences which are written to represent or imitate spoken conversation. Example (34) illustrates a 'correction marker', and example (35) shows where *well* marks the border between a false start and a revised utterance.

(34) Self-Correction Marker
She never told how her escape was managed but Virginia was certain of one thing, the name Senja used was nowhere near the one she was born to. Not long ago, **well**, some time during the first of the New York period, Virginia told Helene, the True Trotskyite, some of Senja's experiences. (COHA Fiction: Snake Pit 1946)

(35) False Start/Revising Marker
"That's, **well**, I'm not sure what you're saying." (COHA Fiction: Ploughshares 2000)

In contrast to the dialogical functions discussed above, WCM *well* and pred *well* are not used to stall for time or mark that one is making a revision. The dialogical function of revising an error in the linear flow of conversation does not have any practical application to writing since the process of writing allows the author to make revisions and search for the proper terminology before publication ever happens. Therefore, the analysis to follow will demonstrate that the usage of WCM *well* and pred *well* have expanded upon certain elements of the sentence-medial functions of dialogical *well*, but have done so in a way that takes advantage of dialogical flavor of sentence-medial *well* to make rhetorical points about key lexical items in the text. Furthermore, when sentence-medial non-dialogical uses of *well* are used in writing, they give the texts what Schourup (2001) argues is the 'mental state interjection' nature of *well* as it is typically used orally. In his paper on *well*, Schourup (2001: 1058)[34] argues that *well* conveys "that the speaker is actively considering whatever it is relevant to consider in determining what should now follow". When acting specifically as a continuative, which includes what he calls 'clause internal uses', Schourup characterizes that *well* is used so that the speaker can be "heard as considering what to say next" (1043). Examples (36) and (37) from transcribed speech in the COHA data demonstrate this action where the speaker uses *well* when they are considering what to say next. These instances of *well* are also significant because they bear a sticking resemblance to what will eventually become pred *well* and the WCM *well* in written texts at the end of the 20th century.

(36) "What an amazing person Manolo is," she said. "Such pain and hurt and, you know, he's not a very handsome man. But when he smiled up at me, his look was just, **well**, beatific." (COHA: Sports Illustrated 1977)

[34] In his paper, Shourup (2001: 315) rethinks and expands on the approach to *well* by Bolinger (1989) who considers *well* to be epistemic in the sense that *well* is used to "invoke a norm".

(37) They and their parents see themselves in double jeopardy, a minority in their own city, yet too urban and too Italian to be part of the American mainstream, which they characterize as suburban and WASP. “When you really feel this is like when you get into college, “said Lucille Poet, a bright-eyed college sophomore whose father is a foreman in a factory. “You can't get a scholarship because you're not quite poor enough? **well**, really, you're not black. When she finished, a roomful of North Ward kids let the silence hang for a long moment. (COHA: Harpers 1972)

Example (36) resembles written pred *well* (see 38) in the sense that it highlights the adjective located in the subject predicative. However, in this case it comes from transcribed speech and in this case the speaker is using *well* as a way to hold the floor while she searches for the appropriate word to use in the stream of talk. In (37) we see a case where *well* is used in a textual function that marks the speaker reconsidering how to phrase something. This has a strong resemblance to how WCM *well* is used (see 39), though it is markedly less rhetorical. However, at the same time, *well* is also arguably marking the speaker’s hesitation (a hedging marker) to express what she intends to say next, which is undoubtedly due to the controversial nature of her (and perhaps the group’s) opinion. This hedging of a controversial opinion is confirmed by the long silence that hangs in the room after she says it.

(38) His Arabic is, well, shocking, and he is studying to improve it. (COHA Magazines: TIME 1963)

(39) Ridiculed and infantilized in the media for months, Lewinsky was surprisingly sympathetic, confident, unrepentant and, well, telegenic. (COHA Magazines: TIME 1999)

Though such functions are not inherent to many forms of writing, I am arguing here that these expansions of the textual sentence medial functions of *well* have developed since the 1990s, because they allow authors of journalistic texts to make their articles feel more interpersonal. This is similar to what has happened with the other colloquial elements that have been increasingly inserted into journalistic writing since the second half of the 20th century, as mentioned in the introduction (e.g. Leech et al. 2009: 239).

When it comes to the formal properties and functions of WCM *well* and pred *well*, it is important to point out that they have two key similarities. The first of these similarities involves the position of *well* immediately to left of the lexical item(s) that the author wants to isolate for the reader to take notice. The second major similarity has to do with the function of *well* which is to act as a signpost that rhetorically instructs the reader to take special note of the word(s) that immediately follow. Because of these key similarities in their characteristics and functions, it is very likely that pred *well* is one individual sub-category of the more general WCM *well*. Though this could very well be the case, I decided to separate pred *well* from WCM *well* for a variety of reasons. First, in journalism, where WCM *well* and pred *well* are used the most, pred *well* alone is used more often than all of the other WCM *well* variants combined. Furthermore, it is also clear that pred *well* developed later in the 20th century than all other variants of sentence-medial *well* including WCM *well*.

Furthermore, one convention that is particularly noteworthy with pred *well* is the use of *well* to saliently mark the author's use of repetition or other forms of wordplay. Though repetition and other forms of wordplay are also noted with WCM *well*, this phenomenon is much more common in the pred *well* construction. This can be observed in Tables 3.3 and 3.4 below. Therefore, even if we make the argument that pred *well* is a sub-category of WCM *well*, it is also clear that pred *well* has become much more conventionalized in journalistic writing in the later part of the 20th century and into the 2000s. It is not only used more than all the other variants of WCM *well* combined, but it is also much more selective in what kind of lexical items it marks (nouns and adjectives) due to its function as marking the predicate construction as well as marking rhetorical word repetition and other forms of wordplay.

Type of Creativity	Freq.	% of pred *well*
Repetition	64	32%
Wordplay	6	3%
Synonyms	3	2%

Table 3.3: Journalism Predicative Well Marking Repetition and other Wordplay

Type of Creativity	Freq.	% of WCM *well*
Repetition	22	17%
Synonyms	4	3%
Wordplay	1	1%

Table 3.4: Journalism WCM Well Marking Repetition and other Wordplay

3.4.3.1 Word-Choosing Marker well (WCM well)

Having explained the similarities and differences between WCM *well* and pred *well*, it is now time to discuss the properties and functions of these two categories separately. WCM *well* occurs in various patterns, including split infinitives, stressing a noun or the head of a noun phrase and at the end of a list of lexical items, as examples (40) through (42) demonstrate below:

(40) Split Infinitive
And when light filtered through that color and those forms, it was enough to**, well,** make an uninnocent eye like Andre Malraux's reach back eight hundred years to the Romanesque for a point of comparison. (COHA Magazines: Atlantic 1963)

(41) Inter NP Marker
Fashioning successful careers in both professions no doubt requires a lot of, **well**, patience. (COHA Newspapers: Washington Post 1992)

(42) End of List
And the idea of paying homage to the old serials was fresh when Spielberg and Lucas cooked it up, but since then we've seen The Mummy, Tomb Raider, and, **well**, even Hidalgo. (COHA Magazines: Esquire 2004)

As already discussed in the quantitative section of this chapter (section 3.3), WCM *well* is a 20th century phenomenon for journalism that rhetorically stresses important word choices for a variety of reasons. Similarly to pred *well*, which will be discussed in more detail below, WCM

well is most often used to by authors to call attention to nouns. In Table 3.5 below, it can be noted that WCM *well* is most often used to mark the usage of nouns and the last item in a list, which also almost exclusively involve nouns. Together these two categories make up 66% of all cases of WCM *well* in journalism.

Type of Lexical Item	Freq.	%
Noun/Noun Phrase	71	53%
End or Beginning of List Marker	17	13%
Adjective Marker	16	12%
Verb/Verb Phrase Marker	16	12%
Adverbial Marker	15	11%
Rhetorical Word-Searching	6	4%
Other	5	3%
Conjunction	1	1%

Table 3.5: Type of Lexical Items being Rhetorically Stressed by WCM Well in COHA Journalism

As the examples above show, WCM *well* is used in ways in which the author can both highlight an important lexical item to his/her readership, as well as mark his/her subjective attitudes towards the information being discussed. Though the usefulness of such strategies is self-evident when reading the examples, it is also important to investigate in what kind of articles such uses of *well* are used. Another related question that must be addressed is whether this increase in the use of WCM *well* is occurring across the board in all categories of journalism or if its usage is more frequent in some categories of journalism over others. It was already pointed out in chapter 2 that the COHA journalistic data includes a good mix of different media sources and different news sections. Therefore, if these functions of *well* were used in all categories of journalism equally, we would expect the representation of article types to be evenly spread across all areas of journalism. However, Table 3.6 shows that non-dialogical *well* has a preference to be utilized in the less formal categories of journalism. At the top of the list is the category of entertainment news. Entertainment news covers everything from movie reviews, celebrity gossip, popular music, and other themes involving popular culture. This is followed

closely by sports articles in the second position. In the third position are editorial/opinion articles. Below the top three categories, the numbers quickly drop. However, it is still worth noting that a diverse variety of article types are covered.

Categories of Journalism	FREQUENCY	%
Entertainment	24	16%
Sports	22	15%
Editorials/Op-Eds	18	12%
Short Story	9	6%
International News	8	5%
Business	8	5%
Politics	6	4%
Literary Review	6	4%
Culinary	5	3%
Travel	4	3%

Table 3.6: Top Ten Types of articles where WCM Well is used in COHA Journalism

3.4.3.2 Predicative well (Pred Well)

The case of pred *well* is more uniform when it comes to how it is used. Furthermore, its usage is by far the most recent of any of the other functions of sentence-medial *well*. For example, 62% of the totality of pred *well* comes from the 1990s and 2000s alone. Moreover, 88% of its total use comes from periods 2 and 3 of Figure 3.2 (1970s-2000s). Despite being such a major part of the contemporary usage of *well* in written journalism, however, it has seemingly only been discussed in one other study to this date (e.g. Rühlemann & Hilpert 2017). As its name indicates, pred *well* is used to highlight the subject predicative constituent when it is a noun or adjective. In other cases, it highlights the noun constituent of a prepositional phrase after a copular verb or the object predicative. For examples, see (43)-(45) below.

(43) Subject predicative
He seemed**, well,** troubled. (COHA Magazines: Omni 1994)

(44) Object predicative
Army grunts resent the fly-boys' sexy scarf-and-leather-jacket image; the jet jockeys consider the cadets**, well,** grimy. (COHA Magazines: Sports Illustrated 1988)

(45) Subject predicative with Prepositional Phrase
But according to sources in the Shady secret service, Encore will sound a lot like, **well**, Eminem: crude wit, intricate rhymes, and once again a disc produced by Marshall himself and Dr. Dre. Oh yeah, one more thing: It's gonna be huge! (COHA Magazines: Entertainment 2004)

Predicative Type	Freq.	%
Subject	152	81%
Object	19	10%
Subject with PP	16	9%
Total	187	100%

Table 3.7: Types of Predicative *Well* in COHA Journalism

As Table 3.7 indicates, the prototypical construction in which pred *well* appears is the following: *[Subject] [Copula], Well, [Subject Predicative]* as can been seen in example (38). *Well* functions as an indicator that the reader should take special note of elements found to the right. As already discussed above, pred *well* often occurs in cases of rhetorical word repetition or other types of wordplay. One such example can be seen in Image 3.1 below, which is the long version of example (39) above:

Why we can't wait to hear Eminem's latest

Encore (Nov. 16)

By EW Staff | September 18, 2004 at 04:00 AM EDT

Why we can't wait to hear Eminem's latest

He's given white rap street cred, powered a Hollywood flick, and turned a so-so crew of Detroit MCs (D12) into chart-topping stars. Now America's favorite speed-rapping, profanity-spewing, button-pushing Caucasian continues his incredible streak with a fourth solo album — and the product is under tighter security than the Bush twins. But according to sources in the Shady secret service, *Encore* will sound a lot like, well, Eminem: crude wit, intricate rhymes, and once again a disc produced by Marshall himself and Dr. Dre. Oh yeah, one more thing: It's gonna be huge.

Image 3.1: Extract from Entertainment Weekly magazine article of Eminem album 2004

In this example, the repetition is not of a word used immediately in the same vicinity or in the immediately preceding sentence. In this case, it is the repetition of the artist's name to describe how the new album will sound. Coincidentally, another interesting phenomenon that occurs in this example is the usage of another construction using another prototypically dialogical DM, *oh yeah* in the near vicinity of pred *well*. This construction with *oh yeah* is used to highlight the end of a list. It works much like the end of list function of *well* that was discussed in the previous sub-section on WCM *well*. The usage of DM constructions involving *well* and *oh yeah* near each other is a phenomenon that will come up again in chapter 5.

However, other examples of repetition are used when the word that is repeated is proximal to a word marked by *well* (sometimes it is even in the subject constituent of the same sentence). Example (46) below shows how pred *well* can mark the repetition of a person's name that is both the subject constituent and predicative constituent:

(46) This is basically a 10th-tier rehash of Indiana Jones, laced with moments that are actually clever and exciting. Dawson is alluring, Walken is terrific, and The Rock is, **well**, The Rock. (COHA Newspapers: CS Monitor 2003)

The previous two examples repeated the artist/actor's name to explain their respective performances. Other examples can be more playful such as the use of repetition of homonyms as exemplified by example (47). In this example, the author intentionally uses the verb 'bond' in the same sentence where the film hero James Bond was already mentioned.

(47) The 21 Bond movies have collectively earned more than $1.3 billion at the North American box office, and international sales and home video amount to far more. Bond is a hero who spans generations, including fathers and sons, many of whom, **well**, *bond* over the spy movies. (COHA Newspapers: USA Today 2006)

Such repetition is generally seen as bad writing style, which one would normally be advised to avoid whenever possible (e.g. Rühlemann & Hilpert 2017: 130). Therefore, the question that

must be addressed is why authors should be using this convention in the first place? What is evident in all the examples given so far, is that pred *well* is used primarily to convey to the readership that the word to appear immediately to right of *well* is something which he/she wants the readers to understand to be both intentional and important. Furthermore, it is also evident that pred *well* is often used to mark repetition and other wordplay for comic or sarcastic effect. Rühlemann & Hilpert (2017: 130) point out that the examples that involve wordplay are reminiscent of *well*'s dialogical function of prefacing and marking dispreferreds, which are used to hedge the fact that the new speaker responds in some way contrary to the expectations raised by the preceding speaker (see also Levinson 1983: 334). In these cases, pred *well* could plausibly be functioning both subjectively, in the sense that the author is using blatant repetitions to mark his/her subjective stance towards the information being discussed in the article (for example the artists references in 46 and 47 have such strong unique characteristics that their names are sufficient to explain their performances), and intersubjectively, in the sense that that author is using *well* to make explicit to the reader that the repetition/wordplay to follow is indeed intentional and that the reader should not view them as a sign of bad writing or a mistake. Moreover, the casualness of using *well* to highlight playful language is also reminiscent of how *well* is used amongst peers in informal casual conversation. Therefore the usage of *well* in these cases where language is used playfully is arguably taking advantage of its casual and friendly nature to index that they are seeking alignment with their readers.

In other words, the use of these DMs in these rhetorical uses cannot truly be operating at the first order of indexicality, which would involve the usage of these dialogical DMs in their typical spontaneous speech environment. Instead *well* in these two functions is operating either at the second order of indexicality or the third, if not both. When indexing the second level of indexicality, they are utilized because they enable the author to 'pretend' that he/she cannot find the right word in order to indicate their subjective reservation/opinion about the word or

the topic being highlighted. The example concerning the actor, The Rock (46), is a good example of this. However, example (47), and many of the other examples discussed in this section, are also operating at the third level of indexicality where *well* operates much more at the intersubjective level in which the author attempts to also get the reader on his/her side and therefore are used to seek alignment with the readers as well. The double use of *Bond* the person and *bond* the verb indicate not only the author's intentional use of word play, but it also serves an alignment seeking stance with the readership in the sense that it includes them in the evaluation of Bond films as a multigenerational phenomenon. Beyond *well* being used intentionally to take up subjective stance (cf. Du Bois 2007), the recognition of sentence medial *well* as a marker of someone who is searching for the proper terminology has led to the development of these specific sentence medial positions into proper written functions of *well* in their own right. The development of these uses into functions in their own right, therefore, operate at the third level of indexicality in which it marks a rhetorical informal and colloquial style of intersubjective stance taking in writing. (cf. Silverstein 2003; Du Bois 2007).

The next question to address is in what categories of journalism pred *well* is used the most. Table 3.8 shows the top ten categories of articles in which pred *well* is used. Unsurprisingly, the top two categories of journalism are also entertainment and sports. Business articles are a close third. After that, the numbers quickly drop, as was the case with WCM *well*, and a wide variety of article types are covered.

Categories of Journalism	Frequency	%
Entertainment	33	18%
Sports	23	12%
Business	22	12%
Editorials/Op-Eds	15	8%
Travel	12	6%
Politics	11	6%
International News	8	4%
Literary Review	8	4%
Arts	7	4%
Financial	7	4%

Table 3.8: Top Ten Types of articles where Pred Well is used in COHA Journalism

3.5 Discussion of written functions

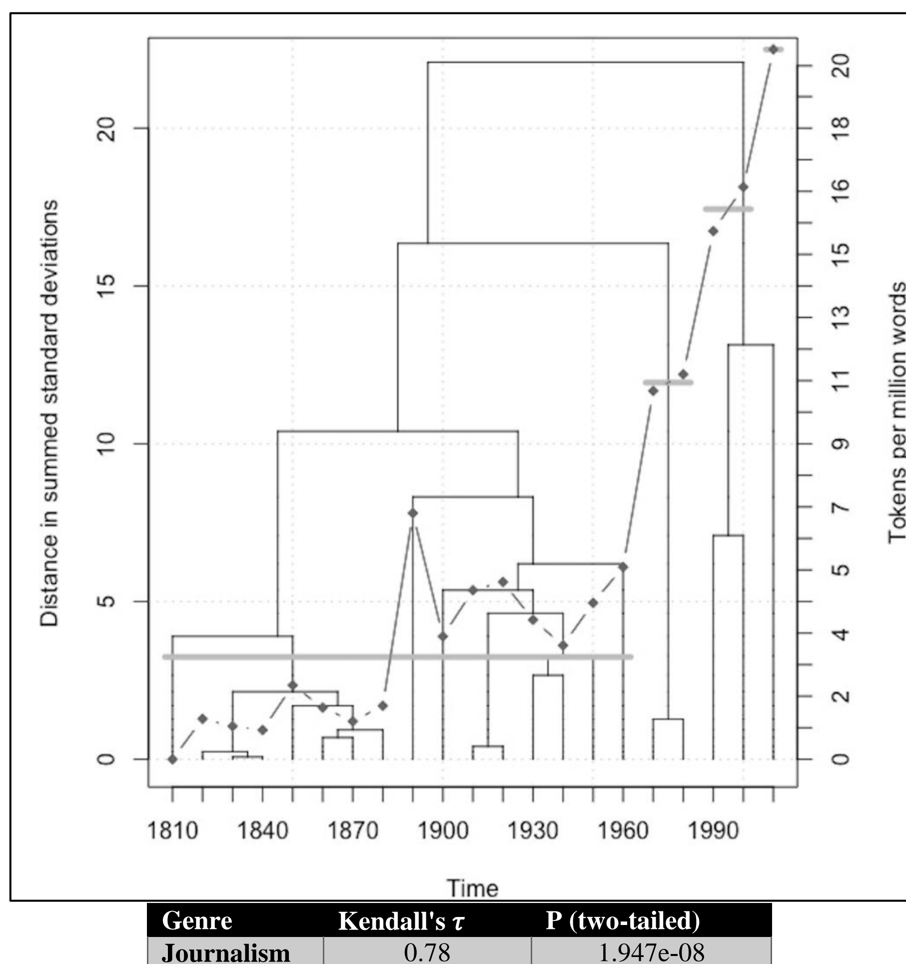
This section has demonstrated how WCM *well* and pred *well* are much more unique to writing than quote *well* or clause *well*. It has also demonstrated that these two categories share two fundamental characteristics: the fact that they involve *well* being used as a salient word-choosing marker and the fact that their position is always to the immediate left periphery of lexical items. This section argued that the origins of these word-choosing categories are evolutions of the oral functions of marking word-searching sequences or self-repair sequences. In these examples, such as in the cases where *well* is marking a lexical choice that is also an intentional repetition of a nearby word or a play on words, it is also reminiscent of the oral function of prefacing and marking dispreferreds. The major distinguishing factor between the oral characteristics and these two written ones is that, unlike a speaker in spontaneous speech, a journalistic author would have no need to use *well* in order to hold the floor or mark a self-correction. This is because the author has the opportunity to both pre-plan what he/she wishes to say as well as to go back later and revise the article. Rather, WCM *well* and pred *well* appear to be used in cases where the author would like to mark the usage of a particular lexical item for rhetorical (inter)subjective purposes. Therefore, in these cases the original text organizing (of coherence) function is still present, though its importance has been weakened as the

subjective stance taking function (along with the attempt to seek readership alignment) is now much stronger. Such revelations stress one of the central ideas of this study, namely that colloquialization can involve more than the simple usage of so-called oral elements of language in written texts. In studies using dialogical data, the (inter)subjective functions that have been identified have typically been those that occur sentence-initially and precede responses (often indirect or dispreferred responses) and face mitigating hedges. This section has shown that the resemblance that the WCM and pred *well* have with the oral word-searching uses of *well*, which encode a speaker 'actively considering what to say', are utilized in journalistic texts to saliently express their (inter)subjective attitudes. Utilizing sentence-medial *well* in this way is especially useful for authors because *well* is highly salient in this position.

WCM *well* and pred *well* are excellent examples of how these oral elements of language can be adapted to suit the needs of writing. However, the use of *well* in these adapted functions is not completely alien to the use of *well* in speech. Instead, these written uses may be seen as expansions or innovations of oral *well* that are better suited for their new environment of written texts. These innovations in the 1990s and 2000s in particular can be seen as a new trend or second wave of change in cultivating a writing style that is more open to using colloquial elements in writing. Furthermore, these two functions can be seen as the next step in development after clause *well* which, as already discussed, constitutes an earlier function of sentence-medial *well* that developed by way of 'cooptation' where the grammatical resources available in one domain are recruited for designing information units in another domain (Kaltenböck *et al.* 2011: 878). The developments of WCM *well* and pred *well*, however, are different from clause *well* particularly because they are niched into writing specifically and not used so freely in all forms of communication.

3.6 Non-dialogical *well* in contemporary journalism data from the NOW Corpus

The previous sections have demonstrated that *well* has developed specific uses in writing and particularly in journalistic writing. These specialized written functions (WCM and pred *well*) were only sparingly used until the 1990s. A major question that must now be addressed is whether these functions are continuing to be used and if any new written functions have been developing. Furthermore, it will also be essential to see whether *well* is continuing to be used increasingly in journalistic texts in the 2010s. As mentioned in the introduction, an additional advantage to using the NOW corpus is that it includes both traditional news sources as well as web-only based news sources. As the findings in the analysis of the COHA data indicated that non-dialogical *well* was used the most in the informal sub-genres of journalism, it will be important to see if *well* starts to expand into other areas of journalism. Moreover, it will be especially key to see if there is a difference in how *well* is used in the web-only based news sources. To answer these questions, contemporary data was extracted from the U.S. sub-corpus of the NOW corpus using the same specially designed search string that was used in the COHA. Figure 3.10, which adds the 2010s of the NOW corpus data to Figure 3.6, shows that in the 2010s the usage of sentence-medial *well* indeed continues to grow. As Figure 3.10 shows, the usage of sentence-medial *well* surges up again like it did between the 1980s and 1990s, going from 16 tokens per million words to just above 20 tokens per million words.



Genre	Kendall's τ	P (two-tailed)
Journalism	0.78	1.947e-08

Figure 3.10: VNC Dendrogram with overlaid line plot for frequencies of DM *well* for periods 1(1810s-1880s), 2 (1890s-1960s), 3 (1970s-1980s) 4 (1990s-2000s) and 5 (2010s) in Journalism COHA corpus data for 1810-2009 and NOW corpus data 2010-2019 with Kendall's τ score

Considering this jump in usage in the 2010s NOW data, the next question to address is how sentence-medial *well* is used in the NOW. A sample of 1,000 concordance lines was taken from the total 23,382 concordance lines of sentence-medial *well*, and each line was annotated in the same way as the COHA data. In particular, it was annotated to mark whether the example was from an interview (or an oral quote) or if it was indeed an example of *well* used in journalistic prose. The rest of the annotation looked at its formal characteristics such as its overall function, whether its occurrence marked an instance of word repetition or other forms of word play and the type of article in which it was used. Of the sampled concordances, 714 (roughly 71%) of tokens were used non-dialogically. Of these 714 examples from journalistic prose, 13 were used as part of DM sequences that were used at the beginning of the sentence. These uses such as

example (48) below are used as a direct copy of the oral convention giving a dispreferred response to a collocutor with DM or interjections such as '*well*, *well*' or '*oh well*'. Though such uses certainly interject a colloquial/interpersonal flavor to the texts in question, these uses will not be addressed in this discussion of the innovations of *well* in the sentence-medial position and have therefore been removed from the data being analyzed in this section.

(48) **Well, well, well**, everybody showed his or her true hand in "Anatomy of the Joke," didn't they? (NOW: Paste Magazine 2012)

3.7 Functions of Non-dialogical *well* in the NOW Corpus

When we analyze the non-dialogical uses of *well* in their entirely (see Table 3.8), 49% of them are used either in the pred *well* or general WCM functions. This is slightly lower a percentage than was seen in the 2000s COHA data. Furthermore, it is also noteworthy that the pred *well* function is no longer the more frequent of the two word-choosing functions, as it was in the 1990s and 2000s COHA data. However, the NOW data also has an additional category that was not observed before. This will be called the meta-textual comment category. In these examples, the author utilizes a parenthetical comment that begins with *well* in order to qualify or put into context the information that was just stated in the text. Similar to the conventions of pred-*well* and WCM *well*, Meta-textual *well* has similarities to the oral convention of medial *well* where it is used by speakers to rephrase or correct their speech. These meta-textual comment functions of *well* will be discussed in 3.7.3. Taken together, pred *well*, WCM and the meta-textual comment functions represent approximately 61% of all the uses of non-dialogical medial *well*. This indicates a continued rise in the importance of these new functions in journalistic texts.

Function	Frequency	%
Clause *well*	257	37%
Quote *well*	17	2%
Pred *well*	113	16%
WCM *well*	228	33%
Meta-textual	86	12%

Table 3.8: Functional Categories of non-dialogical sentence internal *well* in NOW USA data 2010-2019

In the following sub-sections, the five functions of sentence-medial non-dialogical *well* will be looked at in detail. Clause *well* and quote *well* will be discussed briefly in 3.7.1. Section 3.7.2 studies WCM and pred *well* and compare how they are used in the contemporary NOW data with the COHA data from the 1990s and 2000s. Section 3.7.3 will analyze meta-textual *well*, which appears to be the most recent function of medial *well* in journalistic writing.

3.7.1 Quote *well* and Clause *well*

The usage of quote *well* and clause *well* in the NOW data function essentially the same as was observed in the COHA data. When it comes to quote *well*, it was primarily used in sequences involving interview articles or verbatim oral quotes. In seventeen cases (2% of the total sentence-medial uses), however, quote *well* was used in non-dialogical contexts. As was seen in the COHA data, these uses look and function in the exact same way as they do in oral speech and are therefore best regarded as direct transfers of an oral convention into writing. Example (49) presents one such use from an article on the classic Hitchcock film “Psycho”. The author uses *well* to introduce the hypothetical thought of a person:

(49) You liked the heroine? “Psycho” killed her 40 minutes in. You had hopes for the detective? Good, “Psycho” killed him, too. You thought, **well**, maybe now Marion’s boyfriend and her sister would fall for each other? No, they won’t. Whatever your expectations were, “Psycho” ignored them at every turn. A happy ending? There’s only one person smiling at the end of the movie, and it’s dear dead Mom. (NOW: The Star-Ledger - NJ.com 2010)

When it comes to clause *well*, things also look much as they did before in the earlier COHA data with a few changes. Table 3.9 indicates that of all the clause categories, *well* is by far most used in-between an adverbial clause and the main clause. This category alone counts for 70% of the clause *well* examples in the NOW data. Like in the COHA data, the most common type of adverbial clause that *well* is used next to is the if-clause, where *well* is located to the right of the if-clause and the left of the main clause. This pattern alone consists of 40% of the clause *well* examples. Example (50) below exemplifies a typical example of *well* occurring after an if-clause.

Function	Frequency	%
AdvCl.	180	70%
Conj. Main Cl.	38	15%
Topicalization	28	11%
Other Dep Cl.	11	4%

Table 3.9: Categories of Clause *well* in NOW Data

(50) Hemsworth's dreams of a different, wackier Thor movie are likely to come true, as Ragnarok is being directed by New Zealand filmmaker Taika Waititi, the man behind Boy and the amazing vampire comedy What We Do in the Shadows (seriously, you need to watch that sh*t). So yeah, if the next Thor is mostly Rhys Darby from Flight of the Conchords doing improv with Kat Dennings, **well**, you'll know why. (NOW: UPROXX 2015)

Besides the concentration of usage of clause *well* with adverbial clauses, one new observation that has been made with the NOW clause *well* data is the occurrence, though only to a limited extent, of word repetitions similar to what was seen with pred *well* and WCM *well* in the COHA data. In these rare situations, *well* serves the role of marking the author's intentional usage of a word repetition. One such example of this can be seen in example (51) below:

(51) Kids sometimes act impulsively because, **well**, they're kids. (NOW: Concord Monitor 2017)

Overall, however, the number of cases where *well* marks the occurrence of an intentional use of word repetition or other cases of word play is relatively small (only 12 cases of such

repetitions occur in the totality of clause *well*). Furthermore, the vast majority of these repetitions continue to occur with pred *well* and the general WCM functions. For this reason, word repetitions and word play with the NOW data will be discussed in the next section.

3.7.2 WCM and pred *well* in the NOW data

In the analysis of the COHA data, it was demonstrated that WCM and pred *well* were the functions of *well* that were the most specific to journalistic writing. Furthermore, it was also speculated that pred *well*, which appeared later yet still outnumbered WCM *well* due to its surge in the 1990s, was in fact an individual sub-category of the WCM functions. The usage of *well* in these functions rose in frequency during the 1990s and 2000s undoubtedly because of their salience and usefulness in highlighting nouns and adjectives.

These two functions continue to make up roughly 50% of all the non-dialogical uses in the NOW data, though there is a noticeable drop in the proportion of pred *well*. The fact that these two functions have not jumped up in the proportion to the overall non-dialogical uses of *well*, however, may be in part due to the emergence of the newest function to journalistic writing – the parenthetical meta-textual comment. Coincidentally, the meta-textual comments also indicate word selection. Moreover, they mimic/resemble the convention where one corrects or amends something that was previously said. When it comes to how pred *well* and WCM *well* function in the NOW data, things are similar to what was seen in the COHA data (see example 52 for an illustration of pred *well* and 53 for WCM *well*). However, there was a notable rise in the WCM category marking nouns/NPs where they represent 68% of the lexical items being marked as opposed to 47% in the COHA data (see table 3.10).

Type of Lexical Item	Frequency	%
Nouns/NPs	155	68%
Verbs/VPs	36	16%
Adjectives	22	10%
Adverbs	12	5%
Interjection	3	1%

Table 3.10: Type of Lexical Items being Rhetorically Stressed by WCM Well in NOW data

(52) Radiohead was starting to shift toward a more experimental sound, losing some mainstream melody in the process. Creed was, **well**, Creed. (NOW: MTV.com 2014)

(53) Those warm months in the mid-point of the decade were a golden era for pop culture — and particularly music. Sure, the summer of '69 may have had Woodstock, Abbey Road and "Honky Tonk Women," but the summer of '85 had Live Aid, Brothers in Arms and, **well**, "Summer of '69." Strap on your Walkman and read on. (NOW: Rolling Stone 2019)

As these two examples demonstrate, the usage of pred *well* and WCM *well* to mark the author's use of intentional repetitions or word play continues in the NOW data. Example (52) with pred *well* uses the band name to explain their sound, a convention that was seen with pred *well* in the COHA data already (see examples 45 and 46). Example (53), which marks the end of a list and repetition of 'summer of 69' is used playfully as a label for both a period of time (a famous summer in American history for the counter-cultural movement) and a song title (by the artist Bryan Adams) within the same sentence.

Along with the end of the dominance of pred *well*, the use of *well* before these playful repetitions of words is no longer preferentially used with pred *well*. The details of how frequently *well* was used to mark repetition and other forms of wordplay can be found in Table 3.11. What is evident from Table 3.11 is that the use of *well* to indicate repetitions is far more common than using *well* for other forms of word play. Furthermore, though there were some instances of repetitions used with clause *well,* as mentioned above, the use of *well* to mark repetitions is still used most often with the word-choosing functions.

Functions of *Well*	Repetitions	Other Wordplay	Total	% of Function
Pred well	26	7	33	29%
Other WCM	80	5	85	37%
Meta-Textual	6	1	7	8%
Clause Well	11	1	12	7%

Table 3.11: Repetition and other Wordplay in NOW data by function of *well.*

3.7.3 Meta-textual comment *well*

The only new function of *well* that was found in the NOW corpus data is the meta-textual function. This use of *well* can be differentiated from all the other functions because of its formal properties and primary function. As example (54) demonstrates, when *well* is used in this function, it appears at the front of a parenthetical comment. Furthermore, it resembles and in the case of (54) actually pretends to be used as a marker of self-correction. However, as we have seen with the usage of pred *well* and WCM *well*, the use meta-textual *well* in writing does a different task than the oral function that it resembles. Instead of marking the 'mental state' task or surveying and correcting one's speech, it is used to act as a signal to the reader that the author is intentionally amending a statement that he/she has just made in the article. The usage of the meta-textual comment function is highly intersubjective in that it not only expresses the author's subjective stance or playful use of language, but it also engages with the imagined reaction of the readership to the text as well and as such seeks alignment with the readers.

(54) With the character set for his solo cinematic debut in a few years, it's about time that DC Entertainment's Aquaman got some recognition for his constant, under the radar (**well**, under the sonar, technically speaking), onscreen appeal over the last few decades. (NOW: Hollywood Review 2014)

The emergence of these meta-textual comments in the 2010s represents the newest shift in how sentence-medial *well* is expanding its functions into journalistic texts. Leech *et* al. (2009: 246) have shown that 'brackets' round and square have increased immensely in standard English writing (British and American) in the later 20th century. These uses where *well* is used in journalistic prose in such parenthetical comments is thus a later development as far as

journalism conventions are concerned. As was the case with pred *well* and WCM *well* in the 1990s and 2000s, the meta-linguistic function of *well* continues the trend where interpersonal elements of language are increasingly adapted into journalistic texts.[35]

3.7.4 Discussion of changes from COHA to NOW

Overall, in the 2010s NOW data the frequency of sentence-medial *well* continues its rise from the prior decade. Furthermore, a new written function emerges that was not seen in the COHA data at all: the meta-textual comment function. When the three written functions are added together in the NOW data, they account for 61%, which is up 8% from the previous decade. Therefore, by comparing the COHA and NOW data, the case is made that the shifts in the writing style of these specialized areas of journalism have not only been maintained, but they are also expanding further. This continuation of change demonstrates that *well* is still evolving in journalistic texts since the period of major change that first started in the 1990s. Furthermore, as was the case in the COHA data, Table 3.12 demonstrates that the non-dialogical uses of sentence-medial *well* continue to be used especially in the informal types of articles, (entertainment, sports and columns/editorials). Therefore, there has not been a clear expansion of *well* into the more formal areas of journalism.

Categories of Journalism	Frequency	%
Entertainment	217	30%
Sports	101	14%
Column/Editorials	81	11%
Tech	44	6%
Comment Section	34	5%
Science	25	4%
Business	23	3%
Cuisine	22	3%
Politics	19	3%
Lifestyle	16	3%

Table 3.12: Top ten Types of articles in NOW data where non-dialogical *well* is used

[35] One exception to this is the occasional availability of comment sections at the end of the online articles. However, even when these comment sections are available, the author of the article does not respond directly to the commenters.

However, when we consider the origin of the news sources in the NOW corpus, it is also noteworthy that there is no overwhelming preference for these non-dialogical uses of *well* in the web-based sources. As Table 3.13 shows, though these non-dialogical uses of *well* are indeed used more often in articles coming from web-only news sources, this advantage over traditional print and web-based news sources is very slim. The fact that written functions of *well* started to gain traction in the early 1990s, before such web-only news sources became mainstream, likely plays a role in this near equality between traditional news sources and the web-only based news sources.

Source	Number	%
Web only	326	46%
Print & web	315	44%
Television & web	59	8%
Radio & web	14	2%

Table 3.13: Origin of news sources for well NOW data

Returning to the functions of non-dialogical *well*, the analysis of the NOW data also demonstrates that the preferred function of non-dialogical *well* for one decade can fall out of favor after a certain period of time. This is certainly the case with pred *well*, which was by far the most used new written functions during the 1990s and 2000s. After this two-decade period, however, its dominance dropped dramatically in the 2010s. In the NOW data, it represents only 26% of the word choice functions. In the 2000s, for example, pred *well* constituted roughly 2/3 of them. This drop in the usage of the pred *well* function coincides with the emergence of the meta-textual functions that appear at the beginning of parentheticals that serve to inject a sort of *faux* spontaneous qualifier or amendment to the text. These meta-textual comments, like the pred *well* function, occur in a very set pattern. Namely, it occurs immediately to the right of the first parenthesis in a parenthetical comment and immediately to the left lexical item (and in some cases clause) being marked. In this pattern, the usage of *well* to mark word choice is the most salient due to the fact that it is literally separated from the main body of the text by

parenthesis. This sudden emergence of a new function and the limited period of dominance of pred *well* in the 1990s and 2000s followed by its decline in the 2010s, is reminiscent of what occurs with many slang terms, which in many cases have characteristically short-lives after experiencing a sudden burst in popularity due to their notoriety (Coleman 2014: 5). Coleman 2014: 7) defines slang as the following:

> (S)lang is informal, non-technical language that often seems novel to the user and/or listener, and that challenges a social or linguistic norm. It can also imply complicity in value judgments and thus play a performative role in defining personal or group identity.

Obviously, *well* must be distinguished from slang, which usually involves lexical words and constructions. *Well* and DMs in general do not carry semantic content. However, it is noteworthy that what was once the most novel and creative function of non-dialogical *well* drops in frequency at the same time that a similar but newer and more salient function of *well* appears. This indicates that the evolution of the usage of DMs in written functions is continuing to evolve with newer and more salient usages coming into use. However, unlike some slang terms, which often disappear altogether, the older functions of non-dialogical *well* continue to co-exist with the newer variants, at least for the time being.[36] In fact, the broader WCM category has doubled in importance from the 2000s to the 2010s.

3.8 Conclusions

This chapter has demonstrated that in the four COHA sub-corpora there has been an increase in the usage of the DM *well* in sentence-medial positions. Due to the strong tendency for *well* to be used at the beginning of a turn, *well* in the sentence-medial position has received limited attention in previous studies for its use in spontaneous speech. In the COHA data, sentence-medial *well* is used in one of four major categories: quote *well*, clause *well*, WCM *well* and pred *well*. Quote *well* and clause *well* are variants of pre-existing functions of *well* that are also used in spontaneous oral speech. The usage of quote *well* and clause *well* can therefore be seen as

[36] See Adams (2009) for a more detailed discussion on slang.

evidence of an increase of more obviously dialogical features of language in writing, which would fall into the more traditional definition of colloquialization (e.g. Mair 2005; Leech *et al.* 2009). WCM *well* and pred *well* however, though resembling the word-searching and self-repair functions of spontaneous oral speech, clearly do not serve the same functions in writing as their closest dialogical relatives, since the use of such functions are not necessary in the realm of writing. Instead of marking the author's cognitive process of searching for the right way to say something or repairing an error made in the speech process, *well* serves word-choosing functions that mark the author's desire to stress important lexical choices rhetorically. The usage of *well* to indicate one's subjective stance toward the item that follows in a way that 'looks like' the author is considering what to say carries a second-order index in the sense that it takes advantage of the spontaneous speech practice of *well* to be used in situations in which a speaker searches for the right word. Though WCM *well* shows signs of being used early in the 20th century, the much more conventionalized pred *well* is not used significantly until the later part of the second half the of 20th century. When these functions start to encode alignment seeking strategies, the written functions of *well* therefore start taking on third-order level indexicality in which the strategy of using *well* is not only to indicate the author's subjective reservation about using such a word, but it is also attempting to get the reader on their side through the use of a linguistic feature of 'closeness' or 'friendliness'. Furthermore, in the 1990s there appears to be a surge in pred *well'*s usage that continues into the 2000s when pred *well* becomes the single largest function of sentence-medial *well.* In the 2000s, together with WCM *well,* these word-choosing functions constitute over 50% of sentence-medial *well* for the first time. This rise in the overall use of *well* in the second half of the 20th century followed by the second wave in growth starting in the 1990s marks a time of change in how *well* is used in journalistic writing. In particular, this second wave of change is pushed by the sudden growth of pred *well* and WCM *well*. When data from the 2010s using the NOW corpus were added to the analysis, pred *well* lost its dominance to the broader WCM category, though the most

common type of lexical marking still occurs with nouns/noun phrases. Furthermore, in the NOW data, *well* is used in a new and even more salient interpersonal function. This new function is the meta-textual comment function. It includes the use of *well* in a parenthetical comment, and it resembles the author making a self-correction or spontaneous amendment to their text. As was seen with pred *well* and the WCM functions, the usage of the meta-textual comments with *well* are often used to mark the subjective stance of the authors as well as attempt to seek alignment with the readers by giving attention to the readers' imagined reactions to the information in the text.

Therefore, in addition to *well* being used more frequently in writing since the end of the 20th century, it is also being used in innovative ways that retain some of its dialogical characteristics, but function in ways that are more suitable to the specific purposes of writing. In a literal sense, these newly evolved functions of *well* cannot be considered colloquialization, because the functions that they serve do not come from the conventions of oral speech. In fact, the oral speech functions that they resemble serve different functions. Instead of marking that the speaker is searching for or correcting something that they are saying, in writing *well* is marking the importance of the authors' lexical choices as well as their stance (cf. Du Bois 2007). Therefore, the argument made here is that this wave of change in the usage of sentence-medial *well* represents a new shift in the style of contemporary written language where journalistic writing is taking on a more oral-like appearance. However, it is doing so in a way that works to the advantage of the genre in question.

There are several factors that can be contributing to this type of change at this time. First of all, the shifts in various popular written genres towards writing styles that are more 'involved', 'situation dependent' and non-abstract'(cf. Biber 1988) could help foster a writing style in the popular sub-genres of journalism that is in Mair's words more "tolerant of informality and even

allows for anti-formality as a rhetorical strategy" (2006: 187). Therefore, the usage of *well* in these word-choosing functions could be the result of the writing style of journalism becoming increasingly interpersonal-like in its composition. Obviously, the texts that have been observed in this chapter are still monolog texts that cannot directly engage with the readership. However, it is also clear that from the 1990s there is a major shift in how these texts attempt to create the illusion that they can do so. The usage of the DM *well* in these word-choosing functions is one such way that the authors are able to use this 'anti-formalism' as a rhetorical strategy in a way that could also be interpreted as the authors accommodating their writing style to their readership in a way that appears more casual and even more intimate. Evidence for this anti-formalism strategy is found in the types of articles in which *well* is used. In the majority of cases, the articles where these functions of *well* are used pertain to entertainment news, sports news and personal columns or op-eds. Therefore, in a manner similar to Aijmer's (2013: 27) observation that *well* is used most frequently in "more informal types of situations characterized by intimacy and close relationship between participants", *well* is used in journalistic prose in articles that are also inherently informal and written in a style that is more interpersonal than other forms of journalism.

Moreover, the second surge of the word-choosing functions of *well*, particularly pred *well*, starting in the 1990s coincides with the mass adoption of e-mail and other forms of online based language that has only increased into the 2000s and 2010s. This ever-increasing influence of digital discourse has exposed the population to a writing style that is explained by Crystal (2001:47) as "written language which has been pulled some way in the direction of speech". This shift makes the English of the internet neither exactly like spontaneous oral speech nor like formal writing. In fact, it is an expansion of vernacular written English on a massive scale (cf. Barton & Lee 2013). With such a writing style evolving online, it is plausible to imagine that digital discourse could influence various writing styles and journalism in particular. For

example, journalism has already been observed by Mair (2006: 188) to respond more strongly to the social and cultural pressures of society to develop a more colloquial writing style in the decades prior to the advent of the internet age. With this in mind, I therefore argue that the increasing availability and influence of this much more informal and colloquial style of e-mail and social media writing is encouraging journalism authors to continue to innovate and use certain oral-like features of language in ways that mimic orality in order to make their texts more engaging. However, I am not arguing that internet writing is the cause of these changes. This connection between rapid change from the 1990s and internet writing will be discussed in detail in chapter 6.

Chapter 4:
Like: from stigmatization to rhetorical marker of word choice

4.1 Introduction

Of all the DMs discussed in this thesis, *like* is the one that suffered the most stigmatization in recent history due primarily to a variety of language myths that have surrounded it, particularly in the 1980s and 1990s (e.g. Dailey-O'Cain 2000; D'Arcy 2007; Beeching 2016.) Moreover, as most studies of this DM have concentrated on its use in spontaneous speech data, it has, to my knowledge, yet to be studied in published journalistic texts. This chapter will investigate the non-dialogical uses of *like* in journalistic texts. In contrast to what was observed with the investigation of *well* in the previous chapter, *like* is only used sparingly in these same genres until it suddenly appears in the 1990s. Some uses of non-dialogical *like* in writing involve focuser hedging functions in which the author displays his/her subjective stance, or makes alignment-seeking comments while also highlighting lexical items. This is exemplified in example (1) below, which comes from a movie review of the Kevin Smith film *Dogma*:

(1) In a film carefully designed to upset everyone except those with purple and green hair, who will think it's, **like** , cool, "Dogma" is surprisingly restrained in its images. (COHA Magazines: America 1999)

In addition to the focuser hedge functions, *well* is also used in journalism in the exemplifying function. In this function, an example (often something unexpected) is presented in the text. This can be seen in example (2) where the author gives a list of what one should (or should not) do every time one gets new medication:

(2) Your Safest Rx: Before beginning a new prescription, ask your pharmacist for a patient information leaflet, which is available for most medications. At the very least, take the extra 30 seconds to read that little sticker on the bottle. If you have a seemingly silly question (**like**, "Does an iced latte count as a dairy product?"), don't be afraid to ask your pharmacist. (COHA Magazines: Redbook 2004)

A final example involves exaggerated uses of the approximative function of *like*. In non-dialogical contexts, this function is used less for the purposes of actually making numerical estimates and more for the purposes of marking their subjective stance about information, often in a sarcastic manner. Example (3) gives an example where the author adds a positive element to a list of generally negative things that happened to a character in the television program *Game of Thrones*, only to add the comment on the brevity of this 'happy event' (the character in question kills his half-brother immediately after his birth).

(3) In case we forgot how sick and demented Ramsay Bolton is,' Home' gave us a graphic reminder. He hit a low when Theon and Sansa escaped, with another punch to the ego coming when his hunters were found dead. On the bright side, he has a new baby brother. For, **like**, an hour. (NOW: Tampa Bay Times 2016)

This chapter will investigate if *like* is experiencing similar trends as *well* has in journalistic writing. For example, it will determine if these uses of *like* in writing are special to their usage in journalistic prose. Furthermore, considering that *like* was stigmatized during the same period that it first started to appear in journalistic texts (1980s and 1990s), it will also analyze what such uses of *like* say about the nature of colloquialization into the 21st century.

This chapter will begin with a review of past studies of the DM functions of *like* in section 4.2. Section 4.3 analyzes *like* diachronically across the three major sub-genres of the COHA. It will show that the trends of non-dialogical *like* are especially strong in journalistic writing. Section 4.4 investigates the COHA data qualitatively and identifies the major functions of non-dialogical *like* in journalistic texts. Section 4.5 continues the analysis of the non-dialogical uses of *like* in contemporary journalistic data from the U.S. sub-corpus of the NOW to see if the surge of usage continues. It will also check how the non-dialogical functions have evolved. In section 4.6 the individual functions of non-dialogical *like* are discussed. Section 4.7 will offer some concluding remarks.

4.2 Previous work on *like*

Like is one of the most ubiquitous words[37] in the English language according to D'Arcy (2006). Besides functioning as a verb, noun, preposition, conjunction and adverb, it also serves a number of DM functions, as has already been shown above (cf. D'Arcy 2006: 339). Though there is no consensus on the precise number of functions of *like*, most studies on it as a DM state that its core function indicates approximation or looseness of meaning (e.g. Schourup 1985; Andersen 1997, 1998;). Overtime, *like* has also developed functions that act as a focuser or a means of hedging discourse (Fuller 2003: 369; Beeching 2016: 127).

When compared to other common DMs, *like* is perhaps the one that is perceived as the most recently arrived and is the most likely "to be the subject of social comment" (Beeching 2016: 126). Meehan (1991: 39) proposes that *like* began in Old English as the adjective *gelic* that had the meaning 'having the form of'. *Ge* (or *ga*) was a prefix with the meaning 'with', while *lic*' is from the reconstructed proto-european from **liko* meaning 'body' or 'form' (*ibid.*). The adverbial comparative or the 'similar to' meaning that is found in writing as early as the 14th century is thus argued to derive from the adjectival use which means 'in the same manner or to the same extent as' (*ibid.*). In Middle English, Romaine & Lange (1991: 262) find that *lic* becomes syntactically mobile and can appear after the entity being compared: One such example being, "Hire sune wass hym lic (Her son was like him)". This usage is maintained in contemporary English with the suffix in adjectives such as 'god-like' and so forth. Another major function of *like* that dates back to the 14th century is the approximative function, which marks an approximation (Meehan 1991: 40). In common usage today the approximative function is used with quantity phrases such as the example below from Meehan (1991: 40).

(4) I wrote it in like ten minutes.

[37] See also Romaine & Lange 1991: 244

In the 19th century, Romaine & Lange (1991: 244) find examples that use *like* as a conjunction, what Meehan (1991: 41) refer to as the 'as if' function (see 5).

(5) None of them act **like** they belonged to the hotel (taken from Meehan 1991: 41)

Romaine & Lange (1991: 246) suggest that it is this usage of *like* that paves the way for *like* to develop as an interpersonal DM as well as the quotative function. The exemplifying function of *like* is attested at its earliest in the 19th century by Meehan (1991: 42), which comes from Robert Louis Stephenson's *Letters* (1899) as reproduced in (6).

(6) A critic **like** you is one who fights the good fighting, contending with stupidity.

The true DM functions of *like* involve the focuser and hedge functions, which are used to indicate the speaker's attitude toward the material in the clause. This development is in line with Traugott and Dasher's (2002) cline of (inter)subjectivity from non/less subjective to subjective and then intersubjective. For example, Beeching (2016: 154) points out that "*like* arises historically from its core sense of 'similar to' and 'such as'" and then expands to "express exemplification and approximation and thence to develop both a quotative and hedging usage". In her discussion of what she calls the 'focuser' function, Meehan (1991:50) states that "*like* not only functions to highlight the new information, but also allows the speaker to pull back from the assertion in a rather non-committal fashion". Beeching (2016: 154) comments on this most recent stage of development in the following passage:

> The reduction of commitment mentioned by both Meehan and Romaine and Lange is what I have referred to as the 'hedging' use of *like* (Section 6.2.4) and is, to my mind, this hedging quality which promotes its spread, both as an approximative and as a quotative. Because *like* has a rich polysemy which includes prepositional *like* ('s/he looks like, sounds like, is like…') the 'similar to' cognitive core of *like* remains salient even in its more discourse connective and textual uses as a focuser or quotative.

This double function where *like* both highlights new information and serves hedging functions is something that continues to be controversial in the study of *like* as a DM. Moreover, this phenomenon will be analyzed in depth in the analysis of how *like* is used in journalistic writing.

Moving on to studies that look at more recent uses of *like*, primarily from a sociolinguistic perspective, it is immediately clear that the vast majority of approaches look at how *like* is used in spontaneous speech as opposed to how it is used in writing. This lack of research into the use of *like* in modern writing is obviously due to the association of it (not to mention most DMs) with oral speech and especially informal speech. For example, it has been shown in Jucker & Smith (1998: 197) that speakers used 'presentation markers', and particularly *like*, much more among friends than with strangers. Müller (2005: 232) reports similar findings. These findings complement sociolinguistic studies on the perceptions of *like* which consistently show that it is associated with the younger generations (particularly females) and more informal speech in studies throughout the English-speaking world (e.g. Blyth Recktenwald & Wang 1990, Romaine & Lange 1991; Dailey-O'Cain 2000; Buchstaller 2006; D'Arcy 2007). Before it was associated with young Valley Girls however, it was associated with the counterculture groups (e.g. the Beat Generation) in New York City during the 1950s and 1960s. Furthermore, Chapman (1986: 259), in his dictionary on American Slang describes *like* as a characteristic of "1960s counterculture and bop talk" using the example below:

(7) Like I was like groovin' like, you know?

Moreover, it is popularly assumed that it makes people sound uneducated, lazy, or even like 'an airhead' (e.g. Dailey-O'Cain 2000: 69-70; D'Arcy 2007; Beeching 2016: 150-1). In addition to these negative evaluations of *like,* there is also the popularly held misconception that *like* is an exasperating tic and a monolithic and meaningless entity (e.g. Dailey-O'Cain

2000: 69-70; D’Arcy 2007; Beeching 2016: 150-1). These views of *like* are mostly based on the usage of quotative (*be*) *like*, perhaps because this construction is one of fastest-spreading constructions in English today (cf. Mair 2009: 22)

Evidence for these negative feelings toward *like* in the 1980s and 1990s can be found both in popular media[38] as well as in linguistic studies. However, D’Arcy’s (2007: 402) apparent time analysis of the use of *like* in Toronto English finds that people in “their seventies, eighties, and nineties use *like* in the same ways speakers more than sixty years their juniors do.” The main difference between the age groups, however, is that adolescents “use them at higher frequencies than older age cohorts within the population.” (*ibid.*). This therefore provides evidence that the various forms of what she calls ‘vernacular *like*’ were not all invented by the Californian Valley Girls in the 1980s as popular belief holds because all of the variants of *like*, with the exception of quotative *like* (*ibid.*: 406), are used by the age groups that came before them. D’Arcy therefore concludes that, “(a)ll together, the results suggest that while the Valley Girls were not solely responsible for launching these forms in discourse, they did seem to get a boost during the height of the Valley Girl period, which may in turn have helped to propel them into the vernacular faster than they would have otherwise” (2007: 405). One final myth about *like* that must be dispelled is that it is used by females only. D’Arcy (2007: 396) also disproves this myth by pointing out that males and females prefer using certain forms of vernacular *like* differently. In her analysis, women favor the quotative and discourse marking functions, while men favor what she calls the particle functions. Finally, the approximative exhibits no gender pattern at all.

[38] See Beeching (2016: 149) for a discussion of a comedian’s criticism of ‘young people’s language’ on a BBC 4 radio program.

This association of *like* with young females is one characteristic that differentiates *like* from *well.* During the same period of study, for example, the attitudes toward *well* are not nearly as negative. In fact, Watts (1989) views *well* as not being particularly salient and easily overlooked. Furthermore, in terms of gender Müller (2005:139) finds that *well* "seems to be fairly neutral in terms of the sex of its user" when it comes to American speakers. Based on an additional survey conducted by Beeching (2016: 69), *well* "is neither marked for social class/level of education nor is it perceived to be associated with a particular social class/level of education". Ironically, however, this attitude survey went against evidence from her corpus-based data that showed that the usage of *well* in British English actually is sociolinguistically stratified with respect to class, gender and age (*ibid*.: 68). For instance, it is used more by older females who were categorized professionally as "skilled manual" (*ibid*.). This difference in who uses which DM more is undoubtedly a major influence in why *well* and *like* are perceived differently, and this also appears to play a role in why *well* switches into writing earlier and at a higher frequency than *like.*

These observations about the recent development of the *like* and stigmatization of the groups who are associated with it are consistent with the idea of 'change from below'. According to Labov 1994: 78) changes from below:

> (A)re systematic changes that appear first in the vernacular, and represent the operation of internal, linguistic factors. At the outset, and through most of their development, they are completely below the level of social awareness. No one notices them or talks about them, and even phonetically trained observers may be quite unconscious of them for many years. It is only when the changes are nearing completion that members of the community become aware of them. Changes from below may be introduced by any social class, although no cases have been recorded in which the highest-status social group acts as the innovating group.

Ultimately, changes from below lead to shifts in vernacular norms. Once this process becomes stable and salient to language users, they are often stigmatized. (*ibid.*) On the other hand, changes from above are introduced by the dominant social class, often with full public

awareness. According to Labov (1994: 78), they represent borrowings from other speech communities that have higher prestige in the view of the dominant class. It was in the 1980s and early 1990s that *like* started to become mainstream, which coincidentally is also the time that many of the negative evaluations about *like* started to receive attention in mainstream society (and media). The fact that this increasing use of *like* was only commented upon once it entered the mainstream is a quintessential characteristic of 'change from below' (Labov 1994:78). As already pointed out by D'Arcy (2007), though all age groups appear to use various forms of *like*, the groups that use it the most, the younger generations, are the ones who are blamed for its usage and development.

For the purposes of this study, one key element that must be taken into account is that *like*, in all of its DM and other 'vernacular' functions, did not become mainstream in American English until the 1980s/early 1990s. This therefore makes it less likely that *like* will be as present as far back in the COHA journalistic data as was seen with *well*. Taking a synthesis of the most generally agreed upon sub-functions of *like* as well as on what was observed in the COHA and NOW data itself, we find that there are five main sub-functions[39] of DM *like* that warrant discussion before looking at how *like* is used in journalistic writing specifically. The five sub-functions[40] are listed in (8) – (12). *Like* can be used to introduce an example (exemplifying *like*). It can function as an approximative, which can be associated semantically with 'similar to' or 'such as' (approximative *like*). It has the quotative construction or the (*be*) *like* construction (quotative *like*). It has been found to function as a focuser, which highlights the

[39] Underhill (1988), Romaine and Lange (1991) and Dailey-O'Cain (2000) only distinguish two functions of the *like,* the 'focuser' and 'quotative' functions. Fuller (2003: 366-9) in her study based on interview data, highlights some of the issues with the broad label of 'focuser' and insists that this 'focuser' function should be further broken down to include the 'focuser on salient new information' and 'inexactness of meaning' or even both. Müller (2005: 204), in her investigation of conversational data, cites 5 functions (4 discourse marking functions, with the quotative function as a separate category). This disagreement on the number of functions highlights the fact that DMs such as *like* can be used in varying ways depending on the different text-types where it is used (e.g. Aijmer 2013: 18).

[40] This list matches that given in Beeching (2016: 128).

upcoming item (focuser *like*). Finally, there is the hedge (hedge *like*) function. As a hedge, *like* plays a face-saving role that often downplays 'dogmatic delivery' and shares some functional overlaps with other DMs such as *you know*, *I mean* and even *well* (Beeching 2016: 132).

(8)	Exemplifying like:	In the past first dates I've had, we talked about our experiences with OkCupid and Match. (**Like**, how many people we've met, bad/good experiences, etc.) (NOW: Philadelphia Inquirer 2015)
(9)	Approximative like:	Gambon's Dumbledore? Much more shouty. **Like**, at least 5000% more shouty. (NOW: MTV.com 2015)
(10)	Quotative like:	Jenny, of course, is, **like**, "whatever, I don't care." (NOW: The Frisky 2010)
(11)	Focuser like:	there're loads of stories in the newspapers recently about um a couple who went abroad and **like,** SHE'S BEEN FOUND DEAD IN THE DOCK OF A CAR (taken from Beeching 2016: 128)
(12)	Hedge like:	They're just dead. And they'd like you to be dead too. That's, **like**, it (NOW: NJ Pen2019)

However, there is still much debate as to which of these functions are truly DMs. For example, D'Arcy (2006: 342) argues that when *like* is used in the approximative function involving numerical quantities, it is, in fact, functioning as an adverb and not as a DM as suggested by others (e.g. Smith & Jucker 1998; Andersen 2001). In many cases, particularly in cases where the numerical value is a reasonable approximation, the analysis of D'Arcy appears to be correct. However, there are also approximative uses with numerical quantities that clearly also encode the speaker/writer's attitudes. This can be done by blatantly exaggerating the numerical approximation (see example 9). Another fashion in which this is done is by using the approximation function to present something unexpected and/or to humorous effect (see example 3 above). With regard to quotative *(be) like*, most analysts agree that it is related to the DM functions of *like*, especially in cases where (*be*) *like* suggests that the quote is not

necessarily verbatim. At the same time, however, it cannot be considered to be a DM based on the fact that it cannot be omitted from a sentence without rendering it ungrammatical (e.g. Dailey-O'Cain 2000: 61; Müller 2005: 226; Beeching 2016: 131.)

Exemplifying *like*, though usually easily associated semantically with *similar to* or *for example* (see 13), also has examples that are considerably semantically bleached (see 14). Beeching (2016:129) argues that it is the "weakening of this exemplifying usage which leads to the possibility of *like* being used in other ways to hedge inadequacy of expression". Furthermore, it is precisely such semantic bleaching that pushes the process of the (inter)subjectification of DMs. In the COHA and NOW data, some of the most problematic cases of distinguishing functions occurred when the exemplifying function was weakened and the author's (inter)subjective attitude was strongly encoded into the use of *like* (see 15).

(13) why do you want to work in a big company **like** Boots? (Beeching 2016: 128)

(14) we **like** could go to Nepal **like** (*ibid.*)

(15) Does all the basic things you'd' want from a text editor (**like**, you know, the ability to edit text), but with a whole host of scripting and automation tools that take care of a lot of the busy work and save you time in the process. (NOW: Gizmo 2013)

In the case of (15), *like* retains its semantic association to *for example* or *similar to*, yet it is also pragmatically intersubjectified in the sense that it highlights the near repetition of items *text editor/edit text* which resembles a hedging function similar to what was observed with pred *well* in the previous chapter. Furthermore, this example also occurs in a sequence with another DM, *you know*, which is another common characteristic of hedging. In such cases, it is not always obvious which of the simultaneous functions is stronger. In this particular case, it is labelled as an exemplifier though this example is much closer to the hedging function than the example found in (8), for instance.

When it comes to the focuser and hedge functions, however, there is a consensus in favor of labeling them as true DMs. The focuser as well as the hedge functions, which are sometimes combined into a broader category as will be the case in this study, have been analyzed as true DMs in that they share similarities with other common English DMs such as *you know*, *I mean* and *well* (e.g. Underhill 1988; Dailey-O'Cain 2000; Fuller 2003; D'Arcy 2006; Fox Tree 2006; Beeching 2016 etc.). Furthermore, hedge *like* is the sub-category that has the most potential to display intersubjective stance in journalistic texts. Beeching (2016: 132) argues in favor of this, when describing hedg*e like* by stating that "(t)hough it may have a discourse marking or focusing (textual) function, the use of *like* modalises the text, downplaying potentially dogmatic delivery and thus playing a face saving role". Due to this dual focuser and hedge functionality in the COHA and NOW data, the two functions have been fused together and will be called focuser hedges for the remainder of this chapter.

At the same time, however, Fox Tree did an experiment in which speakers retell narratives successively and found that despite these developments toward DM functions, *like* still retains more of its lexical content residue than other similar 'collateral signal' DMs such as *um*, *ah*, *you know*, *oh, well*, *I don't know* and *I mean* (2006: 732). In analyzing the retellings, Fox Tree (2006: 732) found that *like* was recycled at a much higher rate than the other DMs. This suggests that *like* retains more of its lexical content residue than the others, especially in the cases where it functions to indicate imprecision (*ibid*.: 741). These findings corroborate other observations that indicate that the discourse marking functions of *like* are more recent than other common DMs such as *well*. This is also a potential reason for the strong negative stereotypes that persist with *like* when compared to the other DMs observed in this thesis. The newer (inter)subjective functions of *like* will be the focus of the present study. As was seen in the previous chapter, the most innovative and more commonly used functions of non-dialogical *well* from the 1990s

onward were highly (inter)subjective. If a similar trend should occur with *like,* then the most innovation in this case study should occur with *like* used in the focuser hedge functions.

Finally, there is a particular sub-category of focuser *like* that is used to mark the "searching for the appropriate expression" (Müller 2005: 208). In these cases, *like* is used in the same manner as the word searching use of *well* that was discussed in the previous chapter. Similarly to the situation with sentence medial *well,* this use of *like* in spontaneous speech, which is usually surrounded by pauses, is only discussed passingly in most studies (e.g. Underhill 1988; Meehan 1991; Romain & Lange 1991;Dailey-O'Cain 2000; Beeching 2016). However, it is noteworthy that Schourup (1985) and Müller (2005) discuss them in more detail and Müller even classifies it as a function in its own right.[41] In an early recognition of this function, Schourup (1985) proposes the following reason for using *like* in this way: "A speaker saying *like* during a pause to formulate a continuation subtly suggests a reason for the pause: the material about to follow is difficult to formulate appropriately or precisely" (1985: 56). Similarly, Andersen (2001: 249) also found the use of *like* in a way that indicates the speaker's cognitive process, which indicates that the speaker is expressing "'I have something on my mind, but I don't know how to put it'". Müller perhaps explains it best when she says that *like* can be used much like the DMs *well* and *you know* in the function where the author is "thinking about what to say next" or is "searching for the appropriate expression" (2005: 208). This similarity between this function of *like* and the word searching functions of *like* in the previous chapter can be observed in the example provided by Müller (2005: 208) below:

[41] D'Arcy (2007: 394) refers to this function of *like* as a discourse particle that is distinguished from *like* used as a DM. According to D'Arcy the DM functions of *like* fill the syntactic adjunct slot, adjoining in English to the left periphery of CP, the functional projection that dominates the clause"

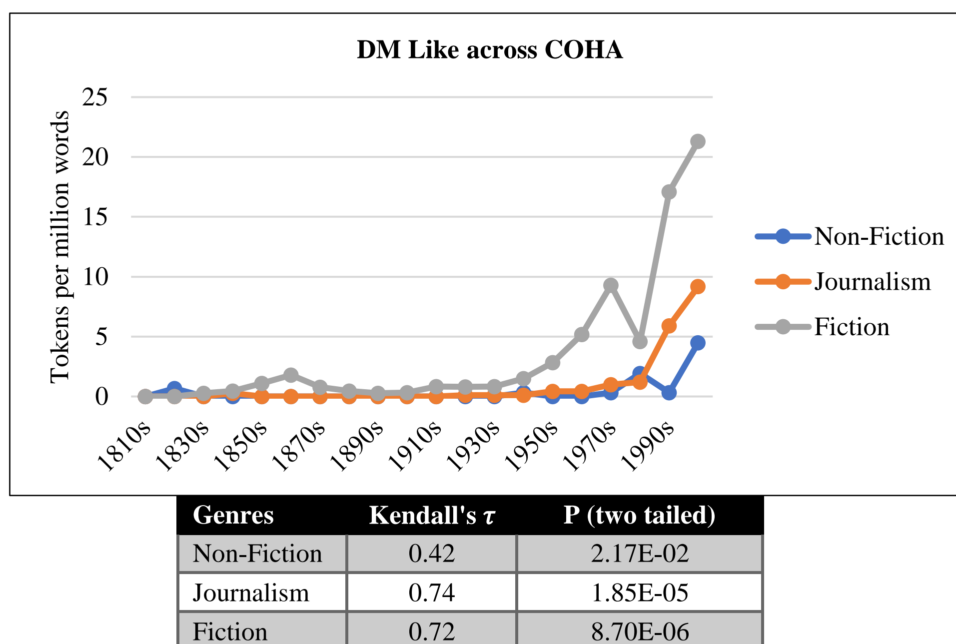
(16) 52 And then .. all of a sudden this … *like* a= &
     53 & .. guy who just found th .. the money also

The use of *like* in this word searching function is also similar to oral sentence medial *well* in that it is not the prototypical DM usage of *like*. For example, Müller shows that *like* as a word searching marker is used significantly less often by American native English speakers than the function which "highlights the following word or expression" (2005: 225). This overlap between the word searching functions of *well* and *like* could potentially result in the use of *like* in similar word-choosing function in journalistic prose. Whether or not this proves to be the case will be discussed further in the qualitative analysis of non-dialogical *like* in sections 4.4 and 4.6.

4.3 Trends of *like* in the COHA Data

Despite these particularly overt negative stereotypes for *like* in speech, the diachronic development for the COHA data in Figure 4.1 shows a steep rise in the usage of *like* in all the written genres under study in the 1990s.[42] This surge in usage in the 1990s is similar to what was observed with *well* in the previous chapter. Coincidentally, this frequency increase also occurs around the same time that the quotative and DM functions of *like* started to receive much of its negative attention. This is most likely the reason why *like* is sometimes used in the 1990s and 2000s in the double sense of mimicking (and often ridiculing) the speech of youth culture and as a word-choosing marker that highlights important lexical items.

[42] The following search string was used in the COHA and NOW searches in order to weed out the lexical uses of *like* :|(|.|,|:|?|!|... like .|,|:|?|!|...|)|



Genres	Kendall's τ	P (two tailed)
Non-Fiction	0.42	2.17E-02
Journalism	0.74	1.85E-05
Fiction	0.72	8.70E-06

Figure 4.1: Diachronic Frequencies for DM *like* in COHA

As the diachronic data in Figure 4.1 shows, *like* was only very sparingly used in all genres of the COHA data until the end of the 20th century. The jump in usage appears to take off from the 1990s for fiction and journalism and in the 2000s for non-fiction. As we have seen with the *well* data in the previous chapter, *like* is used by far the least in the non-fiction sub corpus when compared to the other two. When one also observes how often they are used in non-dialogical contexts, however, it also appears that, once again, the fictional data has the lowest percentage of non-dialogical uses (only 15% of uses were found outside of quotations and other types of fabricated oral speech situations) when compared with the journalism and even non-fiction. This difference in how DMs are used in fiction is undoubtedly due to the significant differences in writing styles and purposes that exist between fiction, journalism and even non-fiction. While most forms of writing are prepared ahead of time and have their primary focus on communicating information rather than developing personal relationships, the narration and dialogue sequences in modern fiction can take on a variety of styles and also has, on the whole, gone in the direction of being written in a "simpler and more colloquial style". (e.g. Biber & Conrad 1999: 156-7). This includes an increase in 1st person narration as well as a significant increase in dialogue between characters. Furthermore, even the examples that are used outside

of quotations and other types of fabricated oral dialogue often come from passages were 1st person narration is being used, a style of narration that is often written to resemble a person's speech stream.

Genres	Non-dialogical	total	%
Non-Fiction	9	24	38%
Journalism	53	209	25%
Fiction	128	897	15%

Table 4.1: Non-Dialogical *like* Percentage for sub corpora of COHA

Considering that journalism is the genre that appears to be going through the most change when it comes to the development of its non-dialogical uses of *like*, this diachronic data was also run through the VNC algorithm. The results of this analysis can be seen in Figure 4.2. Diachronically *like* goes through three distinct stages: stage one ranges from 1900 to 1980 where *like* is hardly used at all, stage 2 comprises of the 1990s alone where its usage surges and stage 3 consists of the 2000s where this surge continues. The data from the 2000s alone makes up 52% of all of the journalism data. If we add the 1990s and 2000s data together, it accounts for 84% of the total.

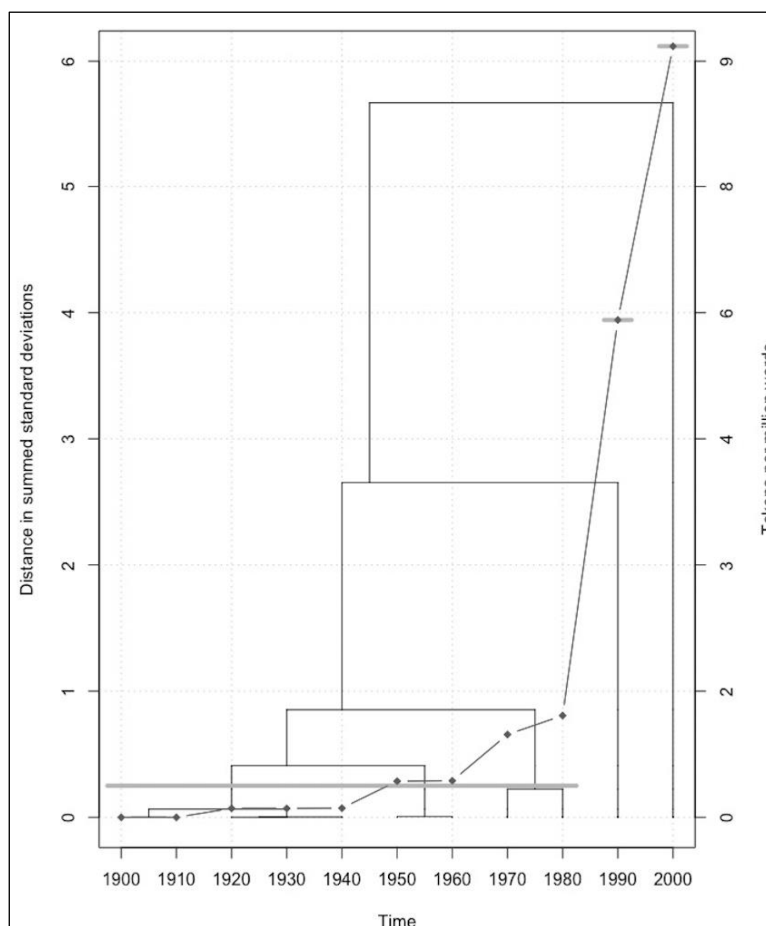


Figure 4.2: VNC dendrogram with overlaid line plot for frequencies of DM *like* for periods1 (1900-1980), 2 (1990) and 3 (2000) in COHA Journalism data

The obvious characteristic of Figure 4.2 is that the frequency per million words of *like* is minimal until the 1990s. In the literature review of the past studies of *like,* it has already been shown that the DM and quotative uses of *like* are more recent overall than, for example, the usage of *well* or other DMs such as *actually* or *in fact*.

However, the fact that the timing for the surge for *well* and *like* is the same indicates that there is something special happening during the 1990s as far as shifts in writing are concerned. Therefore, this rise of *like*'s usage in speech during this same period alone cannot explain why *like* is suddenly used more in writing at this time. It is important not to forget that *well* was an established and common DM long before its surge of usage in writing in the 1990s. Therefore, this chapter will continue search for the answer as to why the 1990s is such a major turning point for how dialogical DMs are used in journalistic prose. The following sub-section will

investigate how these non-dialogical DM uses of *like* are used and, in what sort of articles they are used.

4.4 Qualitative Analysis of non-dialogical *like* in the COHA (1990s-2000s)

The analysis of the COHA data has demonstrated that the usage of the DM functions of *like* have gone from barely used at all in writing prior to the 1990s, to suddenly enjoying a sharp rise in usage afterward. This surge is pushed by its increased use in both quotations and by the usage of *like* in journalistic prose. These non-dialogical examples are used most particularly in journalistic texts in functions similar to what was seen with *well*. The timing of this surge in usage coincides with the spread of *like* into the American mainstream in oral usage as well as the increased usage of the DM *well* in journalistic prose. Furthermore, the apparent specialization of the non-dialogical functions also show similarities with what is happening with *well*.

In this section, the functions of non-dialogical *like* will be analyzed qualitatively in detail. This will include examples from journalism and some from non-fiction. However, the focus will remain on how *like* is used in journalism. Furthermore, this section will investigate the types of texts in which *like* is used in journalistic prose. In analyzing the article types, it will be determined if the non-dialogical uses of *like* are used across all of the sub-genres of journalism similarly or if these uses of *like* have niche functions in specific sub-genres journalism. In the chapter on *well*, it was found that *well* had a preference to be used in articles concerning entertainment news, sports news, and columns/op-eds. The question to be answered here, therefore, is does *like* behave similarly?

The first aspect the non-dialogical usage of *like* that will be addressed in this section is one that differentiates it from the non-dialogical uses of *well*. Specifically, this is the practice of using

like to mimic and poke fun at the group most associated with it, namely young people. In the non-dialogical COHA journalism data, there were 10 such examples (19% of the total non-dialogical examples). The earliest of these examples comes from 1987. The rest come from the 1990s and 2000s. This explicit reference to youth culture, usually with pejorative connotations, further supports the sociolinguistic findings on the attitudes that exist in regard to *like*, especially in terms of its association with young people. Of the 9 examples of non-dialogical *like* from the non-fiction data, two are also used to mimic the language of young people and/or 1960s counterculture. As seen in examples (17) – (22), *like* is used sarcastically in a variety of its typical functions. This includes *like* functioning a word-choosing marker, similar to what was seen with *well*. Furthermore, these uses resemble the focuser and hedging functions of *like* found in speech that both highlight the lexical item to come as well as downplay one's authority or dogmatic stance. In these four examples, *like* is shown to mark the subjective uses of adjectives (16 and 17), interjections (18), noun phrases (19 and 20) and verb phrases (21).

(17) Where a mole in 1966 was mainly an animal, now it is also, thanks to John le Carre, a spy who burrows into the enemy's bureaucracy. A window is not only something to gaze out but also an interval during which rockets can be launched or any opportunity seized. And in addition to all its other 1966 meanings, like has become an interjection, breaking out like acne all over adolescent speech, as in, "It's, **like** , ubiquitous." (COHA Magazines: Time, 1987)

(18) Grind (PG-13): Director: Casey La Scala. With Mike Vogel, Vince Vieluf, Adam Brody, Joey Kern. Four Chicago skateboarders follow a professional' boarder's tour across the country, hoping he'll sponsor their team. They run out of money and a rival team steals their van, stranding them, but they refuse to give up. The cast is young and exuberant, the road-trip gags are plentiful, the skateboarding is, **like**, awesome, but it's often slow going in between. By M.K. Terrell (COHA Magazines: CS Monitor 2003).

(19) One of the many fun games to play with the Starr Report is to pull quotes at random and guess whether it's Bill or Monica talking: "If I had known what kind of person you really were, I wouldn't have got involved with you." The airhead Valley Girl? No, the father of our country. **Like**, duh. (COHA Magazines: American Spectator 1998)

(20) Giroux's understanding of this is entirely partisan, however, as though the only significant variables here are increasing commercialization, the flight of capital and jobs overseas, and the reduction of government services. Reading Giroux, one would think the Left had no complicity in making life harder for children. Wasn't it the counter-culture that glamorized personal indulgence and made raising the young seem, **like**, a drag, man? (COHA Non-Fiction: Academic Qs 2001)

(21) The main reason the film drew a rating as low as "R" is the vocabulary. Those of us who live on campuses and regard language as something sacred surely bridle at the realization that the F-word is second only to "like" in frequency counts for teen-talk, but that's, **like**, a fact, not a moral or political statement. (COHA Magazines: America 1999)

(22) Teens aboard Disney's Magic cruise ship can bond in a Friends-style coffeehouse called Common Grounds, shown at left, and then try their hand at deejaying karaoke or watch first-run screenings of Disney flicks. With so many options, teens might, **like**, even have a good time on a trip with the' rents. (COHA Magazines: TIME 2003)

The examples above in which the language of youth culture is mimicked tell us a lot about the stereotypes that these authors have about the language of the youth in the late 1980s – early 2000s. However, such examples do not make up the majority of the data from the COHA. Most of the non-dialogical uses of *like* are used to represent the author's own desire to mark special lexical items and/or their own (inter)subjective stance. This is done by isolating the words or phrases being highlighted with the use of *like* at the left periphery. The most used function in the COHA data is the focuser and hedge functions as discussed above (see example 23).

(23) Cowboys don't cry. It's, **like**, this rule. Baseball players didn't cry, either, until they became multimillionaires and began to remove themselves from the lineup every time they felt a little stiffness in a shoulder. (COHA Magazines: Sports Illustrated 1993)

Such examples account for 44% of all non-dialogical uses. These examples are labeled as focuser hedges because in the non-dialogical data they appear to combine both the focuser function of highlighting the upcoming item (e.g. Underhill 1988; Romaine & Lange 1991; Dailey-O'Cain 2000) while also performing hedge-like characteristics. Beeching (2016: 128)

gives the following examples in order to distinguish focuser *like* (example 24) from hedge *like* (example 25):

(24) There're loads of stories in the newspaper recently about um a couple who who went abroad and **like** SHE'S BEEN FOUND DEAD IN THE BACK OF A CAR

(25) I wondered **like**, I wonder if your sister could like [pause] find out for us.

However, in her own analysis, Beeching also points out that most of the hedging examples found in her corpus data had both a focusing (textual) function in addition to serving interactive hedging functions (2016: 132). Furthermore, she also states that in her data, there were very few concordances that served the focusing function without an element of hedging also present (*ibid.*).

In the COHA, similar observations are made. Much of the reason for this combining of the focusing and hedging functions in the journalism data may be due to the fact that *like* specializes in roles that are similar to the WCM and pred *well* functions that were studied in the previous chapter. The fact the examples from the COHA data tend to combine the textual focuser element with interpersonal hedging is in line with the observations already made about how the dialogical DMs under investigation in this thesis take advantage of the textual functionalities of DMs from spontaneous speech and use them in writing with added (inter)subjective meaning that facilitates stance taking in the written medium. Therefore, in these cases the authors use *like* to both highlight the upcoming lexical item as well as mark their (inter)subjective stance toward the use of said lexical item, often in ways that resembles how speakers use a hedge in speech.

Besides the focuser hedge functions, non-dialogical *like* is also used in the exemplifying function (see example 26), the approximation function (27) and as a hypothetical quotative (28) The complete breakdown of functions can be seen in Table 4.2 below.

(26) I know what you people want. You want someone to tell you what a tough year it was and what brave little soldiers you are for getting through it, wacky election and all. Or maybe you're expecting cute little phrases to sum things up, **like**, "The Dawn of the New Millennium" or "History in Our Hands" or "The Year Anne Heche Went Nuts in Some Poor Woman's House." Well, I'm not going to help you. You came to the wrong place. (COHA Magazines: TIME 2000)

(27) Oddly enough, that perfectly describes my life in a job I once had. I reacted like the caged rat: by showing my belly and losing sleep. **Like**, 2 years' worth. I feared that the circles under my eyes might knit into a noose around my neck. A shrink prescribed an antidepressant/sleep inducer called Trazodone, and it got me through the dark nights. (COHA Magazines: Men's Health 2006)

(28) The Jazz are down by five. Stockton laces a bounce pass through two sets of legs. He throws a pass football-style, the length of the court. Time out. The Sonics are on their heels, confused. They know they have to keep Stockton from getting the ball, so they put Desmond Mason on him, a forward with a four-inch height advantage. Stockton immediately puts his hand on Mason's solar plexus. Mason pushes it away, **like**, "Get that shit away from me." Stockton gets the ball and hits a three-pointer. (COHA Magazines: Esquire 2002)

The exemplifying and approximative functions are similar to the focuser hedge function in that they are used to highlight particular lexical items in the articles in which they appear. Table 4.2 shows that the majority of non-dialogical *like* is used either as a focuser hedge or as an exemplifier, both of which can also be found in speech. What differentiates these uses of *like* in writing from its use in speech is that *like* in writing is used primarily as a word-choosing marker in ways that are similar to how *well* is. With *like,* the lexical item being marked is found directly to the right of it. The focuser hedge examples also show signs of (inter)subjectivity in the sense that *like* is used in order to make the use of the lexical item(s) being highlighted both salient to the readership while also downplaying their directness or indicating some form of attention to the reader(s) potential reaction(s) to the usage of the lexical item. *Like* is useful in

this function due to its strong association with orality, which therefore makes it salient when used in writing.

Functions	Concordances	%
Focuser Hedge	23	44%
Exemplifying	17	32%
Approximation	7	13%
Hypothetical Quotative	6	11%

Table 4.2: Function distribution of non-dialogical *like* in journalism COHA data

The analysis of non-dialogical *like* shows how it is used to highlight important lexical items while it is often also simultaneously used to show the author's (inter)subjective stance, especially with the focuser hedge function. This preference to use *like* in situations that encode the attitudes of the speaker (AS) as well as the speaker-hearer (SHI) interaction more than the textual organizing (TO) components of discourse is reminiscent of what we saw with *well*.

Furthermore, the appearance of *like* in these written functions from the 1990s shows how, as seen with *well*, textual (or coherence) uses of oral *like* that specialize in highlighting text (particularly lexical items) are the items which the new written usages resemble. For example, (29) shows how a focuser hedge can be used in speech. This example comes from an interview article. In this example, the author takes a critical stance towards one's evaluation of a particular beef stew. Like other examples that are found in journalistic prose, this use of hedging *like* is part of a DM sequence. In this particular case, *like* is preceded by *I mean*. The item being highlighted in this case is the verb phrase "gag me with a spoon" which exemplifies the speaker's rejection of the previous evaluation of the beef stew. Based upon such uses in spontaneous speech, it becomes clearer why using *like* in a similar fashion to mark one's (inter)subjective stance in writing would be so useful.

(29) His second, “Lost Honor,” is a sort of What I Did While I Was Writing “Blind Ambition,” and, unfortunately, the segment about Marlon Brando is typical. He notes, for example, that Mo' s beef stew is “the best in Beverly Hills.” The best beef stew in Beverly Hills! **I mean**, **like**, gag me with a spoon! (COHA: Boston Globe 1983)

Moreover (30) shows the exemplifying use of *like* from an interview sequence from the COHA. If you compare (30) to the written usage in (26), it appears that overall they serve similar functions i.e. they could be replaced by *similar to* or *for example*. The expansion of exemplifying *like* into journalistic prose is, similarly to quotative *like*, one of the areas where the oral uses and the written uses differ very little. Instead, this appears to be more of a case where an increasingly frequent oral convention is being adopted into writing as well.

(30) In the second collection, there is more blandness than bite, although Golden does return to the subject of segregation: “Free of charge, I offered the $64,000 people an idea to help get an additional ten million viewers in the South: Ask the questions they ask Negroes in Mississippi to qualify them as voters. They're interesting questions, **like**, How many bubbles in a pound of soap.” (COHA: TIME 1959)

Finally, (31) displays how approximative *like* is used in transcribed speech. Like (27), it is used to approximate an amount of time. However, unlike (26), this example is used to give a realistic estimation of time as opposed to an exaggerated one, as is more often the case with the usage of approximative *like* in journalistic prose

(31) HOWARI.: Chris, I want you to re -- ally hit hard on the importance of the game when we come on. You know, the old championship, two-undefeated-teams thing. Then, after you've had, **like**, a minute for this, with the trains running on, and the hands and the cheerleaders and so forth, I want you to go down to Bill, who will talk about last week's game.(COHA: New Yorker 1972)

However, *like* is also different from what we saw with *well* because it is also used to represent the author’s interpretation of the opinions of ‘others’ in addition to their own. The question that must now be addressed is in what type of articles non-dialogical *like* is used the most. In the

previous chapter, it was noted that non-dialogical *well* was used more often in the informal sub-genres of journalism such as entertainment and sports news. Table 4.3 shows that *like* is used most often in informal sub-genres of journalism. For example, entertainment, columns/editorials and sports articles make up the top three categories, which account for 61% of all uses of non-dialogical *like*.

Types of Articles	Frequency	%
Entertainment	16	30%
Column/Editorials	11	20%
Sports	6	11%
Politics	4	7%
Health	3	6%
Parenting	3	6%
Faits Divers	2	4%
Travel	2	4%
Biography	1	2%
Dating	1	2%
Journalism	1	2%
Literature	1	2%
Outdoors	1	2%
War	1	2%

Table 4.3: Types of articles where *like* is used in journalism COHA data

What differentiates *like* from *well* in this regard is therefore not the preferred types of articles in which it is used, but the concentration of their uses in these text types. While the top three categories of article types in the *well* COHA journalism data (entertainment, sports and business) represented 44% of non-dialogical uses, it is significantly higher with *like*. This concentration of use in these informal article types could be due to a few factors. First, *like* is used significantly less overall than *well* during the same period (17 per million words in the 2000s for *well* vs. 9 per million words in the 2000s for *like*). This lower frequency aligned with the fact that *like* has been used in writing for a significantly shorter period than *well* may mean that *like* has not yet had the time to conventionalize as much. This stronger concentration of

usage may also be influenced by the negative stereotypes that were linked to the usage of this DM and those who used it. Whether or not the high percentage of usage in the informal sub-genres of journalism continues in the 2010s will be investigated in the following sections on the NOW data.

As mentioned in the overview of colloquialization studies, shifts have already been found to be occurring towards the development of a writing style that is not only tolerant of informality, but even allows the usage of anti-formality as a rhetorical strategy (Mair 2006: 187). The rhetorical usage of *like* in these stigmatized forms in the informal areas of journalism from the 1990s is yet another example of how writing norms of contemporary journalism are becoming more oral-like. However, only the functions of *like* that are useful in writing are developing in journalistic texts, particularly those that can be used to highlight lexical items.

4.4.1 Discussion of *like* in the COHA in the 1990s and 2000s

What primarily differentiates the use of *like* in writing from what has been observed in the previous chapter is that in some notable cases the authors use it in order to imitate and, in some cases, even ridicule the speech of the younger generations/counterculture groups. The reason for these explicit connections between *like* and youth culture in the 1990s and 2000s is most likely due to the strong (and often negative) association between the use of *like* and youth/countercultures (often teenage female culture in particular) in the 1980s and 1990s

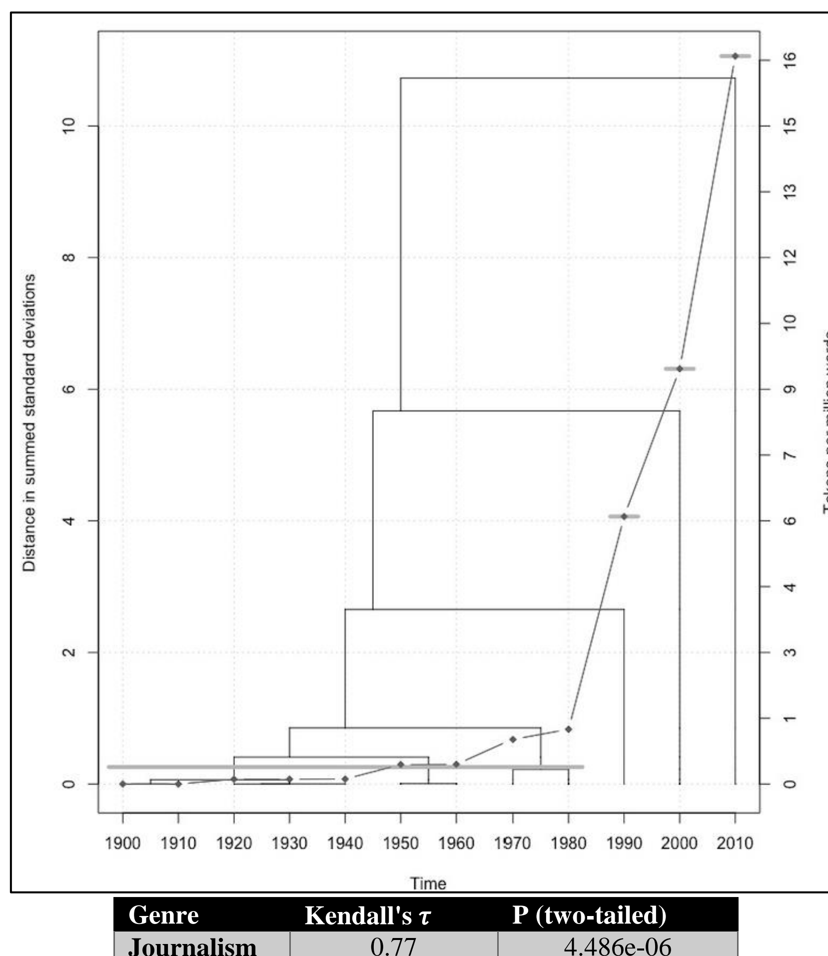
In the majority of cases, however, especially after the 1990s, the authors in the journalistic COHA data use *like* to express their own (inter)subjective attitudes. These evolved word-choosing markers are primarily offshoots of the focuser hedge functions of *like*. However, as was observed in the examples above, there are also examples where the exemplifier and to a lesser extent approximative functions are also taking on similar (inter)subjective word-choosing

marker roles. In these word-choosing uses, we continue to see this combining of second-order and third-order indexicality where the author uses the DM to pretend that he/she is searching for the proper term and furthermore using this as an intersubjective strategy to get the reader on their side (cf. Silverstein 2003). To achieve this, they rhetorically use spontaneous speech styled elements of language that are associated with closeness.

Considering that *like* has been used much more sparingly over the past century than *well*, the next question that must be addressed is whether its usage has continued to rise in the 2010s and if these (inter)subjective word-choosing marker functions are continuing to develop during this time. Another noticeable difference between *well* and *like* is that *like* does not seem to have developed a singularly dominant construction like pred *well*, which dominated how *well* was used non-dialogically journalism from the 1990s to 2000s. The next section will also check to see if the mimicking of youth culture continues in the 2010s. Finally, the next section will also investigate if any change is occurring with the types of articles in which the DM functions of *like* are used.

4.5 Trends of *like* in the News on the Web Corpus (NOW)

In this section, data from the U.S. sub-corpus of the NOW is analyzed and compared to the COHA data. At the time of extraction, the size of this sub-corpus was a little above one billion words. When the frequency per million words is calculated for the overall usage of the DM functions of *like* and compared with the data from the COHA, it is clear that the usage of *like* is still on the rise. Figure 4.3 combines the 2010s NOW data (2010 - June 2019) with the COHA data (1900-2009) in the VNC dendrogram. This rise in overall usage is quite significant as its frequency per million words surges from just above 9 in the 2000s COHA data to 16 in the 2010s NOW data.



Genre	Kendall's τ	P (two-tailed)
Journalism	0.77	4.486e-06

Figure 4.3: VNC dendrogram for the journalism data with NOW 2010s data added with overlaid line for frequencies of DM *like* for periods 1 (1960-1980s), 2 (1990-2000s) and 3 (2010s) with Kendall's τ score

However, when the focus shifts from overall usage to the non-dialogical examples only, things look similar to what was observed in the COHA data in the 2000s. A random sample of 1000 concordances was taken from the total NOW data that consisted of 19007 tokens. After annotating the 1000 token sample, 224 of them (22%) are used in journalistic prose.[43] When the percentage of non-dialogical uses of *like* is compared with what was found in the COHA data of the two previous decades, it is evident that despite this surge in overall usage, the proportion of non-dialogical *like* remains proportionally almost the same as what was seen in the 2000s. In fact, it is the 1990s that has the highest percentage of non-dialogical usage.

[43] As the concordances in the NOW data come from journalism texts, the examples used in quotations primarily come from interview sequences.

Decade	%
1990s	32%
2000s	21%
2010s	22%

Table 4.4: Non-Dialogical *like* percentage for journalism data 1960-2000s (COHA) and 2010s (NOW)

There is, however, one major concession that must be made before moving on from the quantitative analysis of the NOW data. One major difference between online journalism and print journalism is the presence, in many cases at least, of a comment section that immediately follows the main text of an article. Whoever has the ability to comment on an article can differ between different websites. Generally, however, anyone who has an account with a given website is able to do so. The comment board shares some similarities with the traditional 'letter to the editor' section of a newspaper or magazine except for the fact that it is far easier to make a comment online than it is to take the time to physically write a letter or send a formal e-mail to the media source. In the NOW data, there are some examples of non-dialogical *like* that come from these comment sections of the news articles.[44] For example, in the non-dialogical data alone 17 examples come from comment sections. This represents roughly 8% of the non-dialogical concordances. Considering that the examples coming from comment sections do not constitute a significant chunk of the of the total usage of *like* and considering that these online comment sections are similar to the letter to the editor sections in newspapers and magazines that have a long tradition in journalism (not to mention that they are also included in the COHA data as well), the decision was made to include these comments in the analysis.

4.6 Qualitative Analysis of non-dialogical *like* in the NOW data

The COHA data indicated that *like* is most often used non-dialogically in word highlighting functions that included focuser hedges, exemplifiers and to a lesser extent approximatives.

[44] There were also cases of *well* being used in the comment sections of articles in the NOW corpora, though to a lesser extent.

Other uses included the quotative functions for hypothetical or faux quotation sequences. In Table 4.5, there are some indications of *like* being used in the quotative function. Table 4.5 shows the top 25 of collocates of *like* according to the collexeme analysis of all lexical collocates found to the left of all occurrences *like* as a DM. In this list, there are six variants of the copular verb *be* which has already been found to be part of the quotative *be like* construction and a common verb to co-occur in the word-highlighting functions of *like* such as the focuser hedge, exemplifying and even approximative functions. Furthermore, the two forms of the verb *say,* is part of a common construction that is used in the exemplifying function where the common pattern includes the following construction 'say, like + empty slot'. Another similar construction involves *mean* which can also be used in the exemplifier function and sometimes focuser hedge function in the following construction, 'I mean, like + empty slot' Further investigation into the verb *know* in the concordance data was found to often be part of the two word DM *you* know. Coincidentally, *you know* is a common component of DM sequences[45] involving *like*. The use of DM sequencing is more common in oral speech (in this data set they are most commonly found in the interview quotes) but there are also examples of the DM sequence 'like, you know' found in the non-dialogical uses as well. Overall, the collexeme analysis, confirms much of what the qualitative analysis of the COHA data already showed. *Like* often occurs immediately after a copular verb, particularly the verb *be*. It also frequently occurs after other particular verbs such as *say* or *mean*. Furthermore, it can also co-occur with other DMs in sequences, one particularly common co-occurring DM is *you know*. How *like* is used more particularly in its individual functions will be discussed in detail in 4.6.1.

[45] For a discussion on DM sequencing, see Lohmann & Koops (2016) and Haselow (2019).

	COLLEX	ASSOC	COLL.STR.FYE	SIGNIF
1	WAS	attr	Inf	*****
2	JUST	attr	Inf	*****
3	KNOW	attr	258.65205	*****
4	'M	attr	216.15221	*****
5	ME	attr	195.24401	*****
6	MEAN	attr	173.3117	*****
7	'S	attr	166.03986	*****
8	'RE	attr	144.88386	*****
9	WERE	attr	98.03943	*****
10	BE	attr	97.61066	*****
11	QUESTIONS	attr	77.03346	*****
12	SO	attr	58.7128	*****
13	SAY	attr	54.59864	*****
14	SAYING	attr	48.90822	*****
15	THINGS	attr	47.2714	*****
16	MYSELF	attr	43.30544	*****
17	STUFF	attr	38.7914	*****
18	IS	attr	37.12965	*****
19	THING	attr	36.82009	*****
20	BEING	attr	35.52149	*****
21	HIM	attr	34.79954	*****
22	THINKING	attr	31.28417	*****
23	HAD	attr	29.4327	*****
24	RIGHT	attr	28.28293	*****
25	BUT	attr	27.77899	*****

Table 4.5: Collexeme analysis of NOW *like* 1L

Shifting the focus now to the breakdown of the non-dialogical functions of *like*, Table 4.6 presents the four major non-dialogical functions. Overall, things in the NOW data are very similar to what was observed in the COHA data. As the collexeme analysis suggests, the main functions of non-dialogical *like* involve it acting as a word-choosing marker in the focuser hedge, exemplifying and to a lesser extent approximative functions. By far the most common lexical items marked by *like* in the NOW data are nominal elements. Other common lexical items include adverbs, adjectives and interjections. Similar to what was seen in the analysis of *well*, it is the word-choosing functions that become more prominent in the 1990s and 2000s and

then continue to be the most important in the 2010s NOW data. This explains why the focuser hedges represent the biggest chunk of the non-dialogical uses of *like* in NOW.

Function	Frequency	%
Focuser Hedge	104	47%
Exemplifying	79	35%
Approximative	21	9%
Hypothetical Quotative	20	9%

Table 4.6: Functional categories of non-dialogical *like*

The other function of *like* that is found in the NOW data is the quotative function. As in the COHA data, this function is used to introduce faux quotes or thoughts when used in non-dialogical contexts. One big difference that does exist between the NOW non-dialogical examples and the COHA data, however, is that the practice of using *like* to mimic (and make fun of) the language of adolescents or counterculture groups nearly disappears. In continuation with what was seen in the COHA data, however, *like* continues to lack a singular dominant construction, as was the case with pred *well* for a time. When we look at the elements that are marked by the first three categories of non-dialogical *like*, however, we see that nominal elements are the most likely to be highlighted, followed by adverbials and adjectives. Table 4.7 displays how *like* is used in the word-choosing marker functions. Each of these major functions listed in Table 4.6 will be discussed in the following sub-sections.

WCMs	Frequency	% of total non-dialogical data
NPs	90	40%
Adverb(ials)	21	9%
Adjectives	20	9%
Verb Phrases	21	9%
Interjections	2	1%
Word-Searching	1	0.50%
Total	155	69%

Table 4.7: Word-choosing *like* (focusers and hedges, exemplifying and approximatives) in non-dialogical NOW

4.6.1 Focuser and hedge functions of *like* in the NOW data

As discussed in section 4.2 and the qualitative analysis of the COHA data in section 4.4, the focuser hedge DM functions of *like* are the ones that are most strongly correlated with (inter)subjectivity. In journalistic writing, *like* is often used in this function in the combined fashion that resembles a textual function of searching for the upcoming lexical item and the (inter)subjective functions of marking the author's stance and/or acknowledging the readership's potential reaction to using said lexical item. In the NOW data, these functions can be used in a variety of syntactic contexts, but, they appear to be used most often (73% of the time) as word-choosing markers for lexical items such as NPs (see example 32) and adjectives (33). These functions show some similarities to how *well* is used in WCM and pred *well* functions. In fact, example (32) actually co-occurs with *well* in a DM sequence where they both serve as hedges to introduce that the main character, Sabrina, is in a coven that worships the Devil. Furthermore, *like* can also be used to mark the subject predicative (this occurs both in quotations and in the non-dialogical data) though in this case it is not as frequent a function as it is with *well* in the 1990s-2000s. Another similarity is the use of *like* to mark the repetition of a lexical item or a play on words. In the focuser hedges, 15% of them involve repetition or word play (see example 33).

(32) Forget the smirking 1960s comic-book character and the blissful family-of-women '90s sitcom starring Melissa Joan Hart. In 2018, teens have nothing to smile about, so Kiernan Shipka's sorceress lives in two traumatizing worlds. At regular high school, one friend gets bullied; another protests a book-banning. Now, it's not all bad — her boyfriend loves comic books; the best baes always will — but Sabrina double-lifes in the Church of Night, a coven that worships, **um, like, well**, the Devil. The hoof-horned Dark Lord demands that young witches sign their name in his book on their 16th birthday. It's a gender-abasing Faustian bargain. He gives her power if she gives him everything. (NOW: Entertainment Weekly 2018)

(33) This vibe feels familiar. **Like**, Lindsay Lohan Rumors familiar, or Ashlee Simpson's Autobiography familiar. I mean, the Ed Hardy-approved font was the same on the cover of Simpson's 2005 record I Am Me. And while this isn't the first time Swift's played the villain rather than the victim (she had pictures to burn and blank spaces to fill, remember?) the move reminds me of a recent history of pop good girls-next-door who went bad. (NOW: Jezebel 2017)

Both of these examples come from articles that discuss entertainment news. These articles on popular entertainment (especially television, movie and pop music reviews) have already been shown in the previous discussions to be the area of journalism where many of these new functions are being used the most. Coincidentally, it is also the informal sub-genres of journalism that share the most characteristics with internet writing, which has been found to include more interpersonal elements of language (e.g. Zappavigna 2012). For instance, in example (32) the author uses the hedging strategy *um, like, well* in a way that resembles downplaying the fact that the author is about to use a proper noun, the Devil. However, the author is not using this DM sequence to hedge in the same sense that one would in oral conversation. Namely, the author is not actually afraid to actually say 'the Devil' or afraid to offend someone by saying it. Instead, the author uses this hedging strategy as a rhetorical device to make the article feel more interpersonal and inject humor into the article.

A similar strategy is also used for marking word repetitions or wordplay. Example (33) shows how *like* is used to mark the intentional repetition of a lexical item in order to stress the author's subjective stance about the nature of the then newest album released by Taylor Swift. In this case, the author describes this new album as being very 'familiar' to albums released by other artists in the past who wanted to show off a 'bad girl' image. This usage of *like* to repeat and add additional parts to the adjective 'familiar' once again resembles how someone would use *like* in oral discourse to add further details to an utterance they just made. In formal writing, however, such repetitions would traditionally be criticized as being overly wordy and

colloquial. In fact, the adjective *familiar* is repeated not once in the immediate vicinity, but twice. The usage of this example of *like*[46] serves two purposes simultaneously. First, it encodes the author's subjective attitude towards the album being discussed. Secondly, it functions intersubjectively in that it is using *like* to make explicit to the reader that the repetition/wordplay is indeed intentional. Furthermore, this strategy also adds a sort of rhetorical spontaneity to the article that also serves to seek alignment with the readership who the author assumes has the same background knowledge that he/she does.

Other examples include *like* marking the use of adverbials such as example (34) below. In other cases, verb phrases are marked with *like* as in example (35). These examples also resemble the dialogical functions of *like*. For instance, both (34) and (35) use *like* in way that resembles the spoken language hedge function that lessens the dogmatic delivery of giving an order and therefore plays a face-saving role (Beeching 2016: 132). However, since these examples come from writing, the purpose in these contexts have obviously developed into newer roles that are specific to the needs of writers of journalistic texts. In these cases, though *like* is used in a way that looks and still functions like an oral hedge, it is additionally serving the function of seeking alignment with the readership.

(34) The lesson to be learned here? Two of them. First, I'd rather watch Back to the Future for the 51st time than spend one second (or penny) on the latest Saw installment. And second, go out and buy the Back to the Future trilogy on Blu-Ray. **Like**, now. (NOW: Chicago Tribune 2010).

(35) I don't bring that up to, **like**, adshame Common. Nothing wrong with commercials: I never want anything to do with Wal-Mart boxes, but that Dee Rees Wal-Mart box "short film" sure looked cool! Though, as a revolution gets televised, its aesthetics become part of the sales pitch. "What was once considered groundbreaking soon becomes the norm, right?" said Jane Fonda—hopefully or cynically, I couldn't quite tell.

[46] This usage of *like* also represents the ambiguity that sometimes exists between the hedge functions and the exemplifying functions. *Like* in example 33 could also be analyzed as an exemplifier. However, the fact that it was used to mark repetition in an alignment seeking function with the readership and that it also encodes the author's (inter)subjective attitudes towards the album makes its interpretation as a hedge more appropriate.

However, not all of the focuser or hedge functions of *like* are word-choosing markers. In fact, 27% of the focuser or hedge examples are also used in the sentence or clause-initial position where they function similarly to response markers in a manner that encodes the author's subjective stance toward their 'response' or where their evaluation is made implicit (see example (36). This is also one of the rare examples in the NOW that uses *like* in a fashion that is meant to make fun of the speech of young adults/counterculture college students. In fact, there were only two in the non-dialogical examples, which could be interpreted as a sign that *like* is no longer as stigmatized as was back in the 1990s and early 2000s.

(36) Seriously, go read the whole thing. [But, **like**, that was a really long paragraph of unindented text, man!!–ed. Then buy the book — it looks much better on the printed page.] (NOW: Foreign Policy Magazine 2010)

This section has analyzed the focuser hedge functions of *like*. Through the analysis of the NOW data, it has been shown that there are a variety of examples where *like* is used as a focuser hedge DM that highlights important lexical choices. Such uses are used more in articles pertaining to entertainment, sports and columns/opinion pieces than in the more formal sub-genres of journalism. A breakdown of the types of articles in which non-dialogical *like* is most used will be looked at in more detail in the discussion section of the NOW data in section 4.6.4.

4.6.2 Exemplifying and Approximative functions in the NOW data.

As discussed above, the exemplifying function of *like* is the most similar to the canonical meaning as a preposition and conjunction. Therefore, it could be easily associated semantically with *similar to* or *for example.* This would also explain the common collocate, *say* (as in like, say + open slot) when *like* works in this function. However, in many cases this association with exemplifying is weakened and the distinction between this function of *like* and certain hedges, especially those that hedge the inadequacy of expression, is not always clear (e.g. Beeching

2016: 129). In the majority of cases with the non-dialogical data, the exemplifying function of *like* was also used as a word choice marker (see example 37 below).

(37) "A man does not choose his companions," Jaqen says. Now if you're cool like Jaqen, you can get away with referring to yourself in the third person, but don't go trying it at, **like**, the Olive Garden, you'll just sound bizarre ("A man wants a refill on his salad and breadsticks…"). (NOW: Entertainment Weekly 2012)

In this example, the author uses *like* to introduce one such example of a place where speaking in the third person like the *Game of Thrones* character Jaqen would be strange. This example is clearly serving an exemplifying function as it could easily be replaced by "for example". However, it is also flagging the sarcastic use of a location, the American restaurant chain Olive Garden, to the readership, since it is later explained to them that this is not a good place to try speaking in the third person. On the intersubjective level, *like* is functioning similarly to the focuser hedge functions discussed above. In particular, it is highlighting the importance of the lexical item to come while also highlighting the fact that the ridiculousness of the proposition of selecting the Olive Garden as an example is intentional and meant to be humorous.

The approximative function is used significantly less than the focuser hedge and exemplifying functions. However, as in the other two categories, *like* is primarily used to highlight lexical items and noun phrases in particular. As discussed in section 4.2 of this chapter, the approximative function of *like* is prototypically used to give approximate number values. Though D'Arcy (2006) states that the approximative function is not a DM function, but instead an adverbial, this work has shown that in certain cases the approximative use of *like* is more about encoding the author's subjective stance than about giving a reasonable approximation of the numerical value. This includes cases where the author clearly exaggerates the numerical approximation to the point where it cannot be taken as an actual estimation of the numerical

value of something. This was the case with example (9) above that is reproduced below for convenience in (38) as well as example (39).

(38) J.K. Rowling's Dumbledore was ancient, wise, and always in quiet control (albeit infuriatingly vague at times.) Gambon's Dumbledore? Much more shouty. **Like**, at least 5000% more shouty. (NOW: MTV.com 2015)

(39) You have to realize if we've been talking on the phone, emailing, and flirting, it means I am interested in you and your chances of being rejected are much, much lower. **Like**, .003 percent, should Clive Owen show up on my front stoop with a package of Magnums. Are you really so afraid to ask me out for a $4 cup of coffee at Starbucks and risk a .003 percent chance of rejection that you just won't do it? (NOW: The Frisky 2011)

These cases where the numerical value is exaggerated, which is how it is primarily used in these non-dialogical cases, demonstrates that when *like* is used in this way it is not functioning as an adverbial but as a DM. This also explains why these approximative *like* examples are separated from the main clause and surrounded by commas. In these cases, *like* is used subjectively by using one of the common dialogical strategies of *like* for the purposes of marking important lexical items in the text(s). Furthermore, this function of *like* also includes non-numerical examples uses such as (40) below:

(40) The real gem in the Swedish Invasion line, though, is Kolbein Karlsson's debut book, "The Troll King," which is slated to ship in April. Trading in a different kind of exotic other-ness, Karlsson's story is a surreal exploration of a fantastical forest kingdom where hairy (**like**, Cousin It-hairy) homosexual mountain men pray to the gods of the forest for butt babies, old gnome-like men fall asleep and cavort with spirits, carrots that walk like men sprout into trees, and green trolls grow themselves from buried skulls. (NOW: Comics Alliance 2010).

These examples show that approximative *like* in non-dialogical contexts has developed a subjective function that resembles the oral convention of *like* to signal that the following value/comparison should not be taken at face value. However, in these non-dialogical uses *like* is not encoding an approximate value only but also the subjective stance of the author. This

explains why the numerical values in particular are often extreme (e.g. 5000%, 0.003, etc.). In these uses, the function giving a realistic approximation is weakened and the subjective stance of the author is strengthened.

4.6.3 Hypothetical quotative *like* in the NOW data

In the non-dialogical data, quotative *like* continues to be the least used function of *like*.[47] Furthermore, it does not to appear to change in how it is used at all between the COHA and NOW data. Example (41) below shows the author representing the potential reaction of the readership. Example (42), on the other hand, uses quotative *like* to introduce the author's own inner thoughts. Of all the fashions in which it is used, quotative *like* in non-dialogical contexts is a direct transfer of the conventions of its dialogical function.

(41) I should have also been in that dog-walking scene you just read: I was sitting on a bench fifty feet away, yo! But I guess that at that point I hadn't been sufficiently established as a character in my own right, so I would've seemed like some random scruffy Italian-American sweetie-pie lurking in the background, and readers would've been, **like**, "What's up with Mark Ruffalo on the park bench?" (NOW: New Yorker 2012)

(42) Manning's has never been a silky or speedy or even that strong an athleticism. But it was clean, and perfect, and therefore beautiful. And it somehow inspired a string of unfootbally thoughts, **like**: *It must be so very quiet in Peyton Manning's head... I shall henceforth refer to him as Peyton Motherfuckin' Manning. (*NOW: GQ 2013)

As these hypothetical quotative examples appear to be direct transfers of a colloquial option of representing the quote of a supposed 'other speaker' or a self or thought quote into writing, this function is very similar to the quote *well* and clause *well* functions in the previous chapter. Furthermore, as already stated in the previous chapters, studies on the colloquialization on the written English on the micro-structural level (e.g. Mair 2006; Leech et al. 2009) show a rise in the preference for informal options over formal ones where both are available. The quotative

[47] When all uses of *like* in the NOW sample are accounted for (including the examples from interview quotations) quotative like accounts for roughly 25% of all uses of *like*.

function is a clear example of an informal option being used over various other more formal ways of representing hypothetical quotes or thoughts. For example, if one considers the strong (and often negative) associations that the general population holds (those arguably less so than in the 1990s and early 2000s) for the usage of this function of *like*, the fact that it is still used in writing shows that it is becoming an increasingly popular way to represent speech (e.g. Dailey-O'Cain 2000; D'Arcy 2007; Haddican & Johnson 2015 and the discussion in section 4.2).

4.6.4 Discussion of non-dialogical NOW data

Comparing the COHA and NOW data show both similarities and differences. In terms of the dialogical development of *like,* its rise in usage continues in the 2010s data where its frequency per million words is around 16, up from 9 in the 2000s. When it comes to the use of *like* in journalistic prose, similarities between the COHA and NOW data include the same major functions of *like*. In both corpora, non-dialogical *like* is used primarily as a word-choosing device that displays the author's subjective stance while also often seeking alignment with the readership. Such uses come mostly from how *like* is used in a focuser hedge function, but also the exemplifying and approximative functions. *Like* is used the least in the quotative function. Furthermore, how it is used in the quotative function mirrors how it is used in orality.

Another similarity involves the types of articles where non-dialogical *like* is used. With the exception of the comment section category, which is discussed above, the types of articles that include *like* the most are entertainment, sports and columns/op-eds articles. These three categories of articles include over 50% of all the occurrences of non-dialogical *like*. Interestingly, with the exception of the presence of examples from comment sections, the article types and their percentages look similar to what was seen with *well* in NOW data in the previous

chapter. The only major difference is how much more concentrated the uses of *like* are in entertainment articles.[48]

Categories of Journalism	Frequency	%
Entertainment	93	42%
Sports	18	8%
Comment Section	17	8%
Column/Opinion	14	6%
Politics	9	4%
Science	8	4%
Business/Financial	8	4%
Tech(Computers)	8	4%
Technology	6	3%
Literature Reviews	4	2%

Table 4.8: Top ten Types of articles in NOW data where non-dialogical *like* is used

Considering that non-dialogical *like* is used in similar article types as we saw in the COHA data, it is also essential to see if *like* is used significantly more in web-only based sources in the NOW. Table 4.9 demonstrates that there is indeed a preference for using *like* in articles from web-only sources of journalism. However, it is also noteworthy that many non-dialogical uses of *like* also come from traditional sources of written journalism as well as the websites of television and radio stations that also include written articles. However, this clear preference for *like* to be used in web-only news sources differentiates it from *well*. In the previous chapter, it was shown that *well* was spread nearly evenly amongst traditional news sources and web-only ones.

Source	Frequency	%
Web only	132	59%
Print & Web	74	33%
Television & Web	14	6%
Radio & Web	4	2%

Table 4.9: Journalism source distribution for non-dialogical *like* in NOW

[48] These are also the same categories of journalism from which the most examples of non-dialogical *oh yeah* come.

A qualitative difference in the NOW data and the COHA data was the near disappearance of *like* used to represent and denigrate the imagined speech of adolescents or counterculture groups. While such uses of *like* were noteworthy in the COHA data, they are not present to the same extent in the NOW. This could be taken as an indication of a lessening of the stigmatization of *like* that was observed in the previous decades. Furthermore, the youth of the later 1980s and early 1990s are now adults and very likely authors of many of these articles in which *like* is used. However, while the usage of such forms of *like* are used far less to ridicule the speech of youth, these vernacular forms of *like* are still primarily used in the informal article types of journalism. This therefore demonstrates that their usage in articles that pertain to entertainment news, sports news and personal columns may be deemed an appropriate place to use informal DMs such as *like*. Furthermore, *like* is used to a higher extent in web-only press sources than was the case with *well*. In other words, these forms of *like* are intentionally used precisely because they are colloquial and anti-formal in order to create and maintain a form of intimate rapport with their readership. Furthermore, this high preference to use these forms of *like* in articles covering entertainment news also indicates that these forms of *like* are becoming part of the shared repertoire of entertainment news journalists and internet-news bloggers. These participants in this specific sub-field of journalism could be regarded as a separate 'discourse community' that does things differently than authors in other sub-genres of journalism (cf. Watts 1999).

Another difference involves the emergence of non-dialogical *like* in parenthetical comments that perform alignment-seeking functions with the readership. These uses are similar to the meta-textual comments that were discussed in the previous chapter. With *like*, the author also uses the parenthetical comment to address the reader(s) directly outside of the context of the main text. However, unlike in the last chapter, in which meta-textual *well* was discussed as a new function of *well*, with *like* it appears to act according to the major functions already

identified. In the majority of cases, *like* is executed as a focuser hedge (43) or as exemplifier (44). The breakdown of the meta-textual *like* function can be seen in Table 4.10 below.

Function	Frequency	%
Focuser Hedge	10	34%
Exemplifying	15	52%
Approximative	4	14%
Total	29	100%

Table 4.10: Functions of meta-textual *like*

Though the parenthetical comments make up only a small proportion of the non-dialogical examples (13% of all non-dialogical uses), their highly (inter)subjective nature and the fact that they are found in the NOW data point to it being a newly emerging property of *like* in journalistic writing.[49] Furthermore, this practice of using interpersonal aspects of language in writing shares similarities with the online language that, as already discussed, enters into the mainstream in the 1990s and becomes increasingly incorporated into everyday life thereafter.

(43) This time last year, we were still nervously anticipating the release of J.J. Abrams' Star Wars: The Force Awakens (**like**, really anticipating) and just praying to Yoda (he hears prayers, right?) that this new phase of the franchise wouldn't let us down. It didn't. (Thanks, J.J.!) (NOW: Wired 2019)

(44) Does all the basic things you'd' want from a text editor (**like**, you know, the ability to edit text), but with a whole host of scripting and automation tools that take care of a lot of the busy work and save you time in the process. (NOW: Gizmodo 2013)

The fact that *like* is hardly used in non-dialogical contexts until they suddenly appear in the 1990s along with the finding that they are used most often in the informal areas of journalism, such as entertainment, sports and columns/op-eds gives, confirms a great deal of what was found in the previous chapter. Furthermore, it is particularly the word-choosing functions of

[49] Similar comments were also found in the previous chapter and the next chapter discusses parenthetical meta-textual comments involving *oh yeah* in more detail.

like that have developed specifically written functions. Moreover, the fact that 59% of all the non-dialogical concordances come from web-only based news sources, as well as the fact that the majority of all non-dialogical uses come from the informal sub-genres of journalism, show that *like* is a DM that is most at home in informal journalism even more so than *well* is. It was suggested in the previous chapter that the increasing influence of online-based writing could be one factor that is pushing these changes. The trends found in this chapter give further evidence that this could be the case.

4.7 Conclusions

This chapter first looked at how the DM functions of *like* are used in writing in a variety of text-types using the COHA data and in journalism in particular. The diachronic analysis showed that the usage of *like* in journalism has surged since the 1990s. As was the case with *well*, *like* is used in journalistic prose most often in functions that highlight lexical items. However, unlike *well*, *like* does not have a pred *well* equivalent. The majority of these uses of *like* function as either focuser hedges or, to a lesser extent, exemplifiers. Coincidentally, *like* as a focuser hedge most strongly encodes the authors taking a subjective stance in the sense that it highlights the lexical item/phrase following it in a way that indicates that author's attitude toward the item. It also displays the author's (intersubjective) attention to the readership while also trying to get them to align with him/her. Furthermore, the focuser hedge function has the double role of mimicking the speaker-based practice of conducting a word-search or pause while also minimizing dogmatic delivery or over-directness. This observation is further supported by the fact that the less (inter)subjective functions of *like* were less represented in the non-dialogical data of the COHA and NOW corpus searches.

This signifies that *like* has also developed specific journalistic writing functions that maintain a resemblance to the text monitoring functions. In journalistic prose, they take on interpersonal

characteristics and specific written functions are developed which are utilized in a way that uses the association that *like* has with casual friendly conversation to evoke the positive politeness strategies of stance taking in which the author seeks alignment with the readership. The usage of *like* in these texts and its continuous rise in usage signifies that the changes occurring in published journalistic texts can also include stigmatized linguistic changes currently occurring in speech as well. Despite being stigmatized in the 1980s and 1990s, *like* still crossed over into journalistic prose starting in the 1990s. This even included quotative (be) *like* (cf. Haddican & Johnson 2015), which Mair (2009: 22) identifies as one of the fastest spreading constructions of English, despite the fact that it was perhaps the most stigmatized form of *like* into the early 2000s (cf. Dailey-O'Cain 2000). This is different from the case of *well* in the previous chapter. In contrast to *like, well* has been found to have no association with low prestige groups and it is not salient when used in speech in the sentence initial position. This difference in attitudes is most likely the reason why *well* has transferred into writing earlier. Furthermore, this is also likely why it is used at a higher frequency despite the fact that these two DMs have overlapping word-choosing/hedge functions.

However, the fact that *like* is also being increasingly used in writing despite being a stigmatized form in informal genres shows that writing has the potential to become not only more colloquial, but it also shows that these shifts can involve the use of low-prestige elements. Authors can use elements such as *like* as an anti-formality strategy in order to create the feeling of a personal rapport with their readers in which they attempt to convince them of accepting their stance. For example, Jucker & Smith's (1998: 197) finding that people use 'presentation markers', which includes *like*, *you know* and *I mean,* much more among friends than with strangers indicates that the usage of a DM such as *like* has the potential to create a friendly-like atmosphere within the article. A similar point could be made in terms of *well* in the previous chapter. Even though the written functions of *well* resembled mostly the word-searching/self-

correction uses of *well*, its usage as a word-choosing marker was often used playfully to hedge the intentional repetition of a word or to indicate word play (cf. Rühlemann & Hilpert 2017). The fact that these sorts of phenomena are being increasingly used in these article types of writing also indicates that these texts are continuing to shift in ways that make them appear more 'socially interactive'(cf. Crystal 2006: 20) or more 'involved' (cf. Biber 1988) while still being monologue texts. This brings us back to the whole argument that these journalistic texts are not actually becoming more like actual speech. Instead, they are progressively adopting elements of dialogical language, but coopting them into the written medium. In the case of dialogical DMs such as *like*, they are being used in salient ways that look like speech although they act in different ways.

The fact that these changes are occurring in the entertainment news genre so significantly more than the others, however, shows that these colloquial shifts are not occurring in journalism across the board. In the COHA and NOW data, common article themes that included these specialized uses of *like* involved articles about popular music, television programs, movie reviews and celebrity gossip news.

Chapter 5:
Oh Yeah: Rhetorical marker of (inter)subjective information management

5.1 Introduction

This chapter moves on to the non-dialogical uses of the DM *oh yeah* in journalism. Overall, *oh yeah* as a DM has been largely overlooked in its own right. However, in journalistic writing it demonstrates the development of genre-specific functions similar to what was seen with *well* and *like*. There are various functions of *oh yeah* that are found in journalistic writing. The dominant example of a highly subjective use of *oh* yeah in contemporary journalistic writing is shown in example (1) from COHA. In this example, the author uses *oh yeah* in a fashion that resembles how one would use it in speech in a fashion that indicates the author suddenly remembering something that he/she would like to add. This would fit into how Schiffrin (1987) views *oh* as an information management marker. However, in this case it is used to highlight what is, in fact, the element about the Raiders football team that the author considers the most important. Therefore, despite resembling the oral information management function of *oh yeah*, which marks shifts in the speaker's orientation as the flow of information occurs during discourse, it is used in (1) as a pre-planned rhetorical tool to highlight what the author considers to be the essential element in the list of features that describe this American Football team.

(1) The NFL ultimately might remember the 2002 Raiders as champions, but for now, this is a crack team built on experience, poise and confidence. **Oh yeah**, and a fire-stoking measure of outrage (COHA Newspapers: San Francisco Chronicle 2003).

Another function involves a sort of 'staged' rhetorical response marker in which the author comments on and intensifies the importance of the information to come (see example 2 below). This also resembles a very common oral form of *oh yeah*. In this case, *oh yeah* literally mimics an utterance launcher (Biber *et al*. 1999: 1076), which also expresses some degree of suddenly remembering something. This use is similar to (1) in that the function that it resembles and the function that it actually serves is the exact opposite. In this particular case, the author is not

suddenly remembering that he/she started writing the column in the 1990s, rather the author is rhetorically using this DM to highlight the fact that it was at this time that she/she started to write the column.

(2) What happened in the 90's? Something must have but it's not coming to me at the moment. **Oh yeah**, I started writing this column (and the world would never be the same, hahaha) (NOW: Orange County Register 2017).

A third major function involves the author displaying an intersubjective use of *oh yeah* where they respond to a supposed question that could be posed by the reader(s) in response to information presented in the article. Example (3) below from a recap of an episode of the popular television series Game of Thrones gives an example of this function. In this example, the author aligns with the reader(s) by making a parenthetical comment that touches on the common problem of keeping track of who is who.[50]

(3) After the death of Oberyn Martell last season, Cersei is concerned about her daughter Myrcella (**oh yeah**, remember her? More on her below) who has been living in Dorne since season two (NOW: USA Today 2017).

In the examples above, it was observed that *oh yeah* is used non-dialogically in a way that resembles the author suddenly remembering something, while the function that it serves is precisely the opposite. The author is not using *oh yeah* in a textual manner, instead *oh yeah* is used as an interpersonal function that highlights for the reader what it is that they find important. Though similar functions were found with *well* and *like*, this practice of using a DM in a function that looks dialogical yet is different is especially salient with *oh yeah*.

[50] This television program involved many intertwining sub-plots and the introduction and killing of many characters throughout the series.

This chapter therefore investigates the development of these non-dialogical functions of *oh yeah* that share an appearance with the oral functions of monitoring what is being said, but in fact act at the interpersonal level that also index the author's attitude or viewpoint (subjectivity) or the author's attention to addressee self-image (intersubjectivity) if not both (Traugott 2010: 32). This chapter will investigate if *oh yeah* is also diachronically developing these journalism-specific (inter)subjective written functions as was the case with *well* and *like*. As with the typical cline of (inter)subjectivity, will there be a diachronic shift where *oh yeah* is increasingly shifting away from functions in writing work primarily at the textual level toward functions that primarily encode the author's subjective stance? Furthermore, this chapter will continue to investigate in what sub-genres of journalism these uses are most common. Another question that will be investigated is whether the written functions of *oh yeah* share any similarities with the development of slang terms that are inherently informal and also rapidly appear and disappear within single generations, as was found with pred *well* in chapter 3.

This chapter will begin by reviewing past studies of *(oh) yeah* in section 5.2. It will be shown that although *oh yeah* itself has only been mentioned in a few papers so far, various studies that have analyzed *oh* and its variants shed some light on the usage and functions of dialogical (*oh) yeah*. This section will also discuss how some of these oral functions are able to transfer from dialogical environments to non-dialogical ones. Section 5.3 analyzes *oh yeah* diachronically using data from all of the sub-genres of the COHA. This section will compare how *oh yeah* is used in journalism texts with non-fiction and fiction. Taking this into consideration, section 5.4 presents a qualitative analysis of *oh yeah* in the COHA journalism data and identifies how it is used in these texts. Section 5.5 analyzes data from the U.S. sub-corpus of the NOW corpus and compares it with the diachronic COHA data. Section 5.6 analyses the NOW data qualitatively and identifies and compares the major functions in which *oh yeah* is used in the 2010s. Furthermore, this analysis will verify if *oh yeah* is being used in innovated ways or if the usage

is the same as seen in the COHA data. Finally, section 5.7 offers the final discussion and conclusions.

5.2 Theoretical background of *oh yeah*

Oh yeah is a two-word DM that, like many other DMs, can operate in a variety of functions. In the etymology section of the Oxford English Dictionary Online (OED), *oh yeah* is labeled as the combination of the interjection *oh* and the adverb *yeah.* It is considered colloquial and its definition is as follows:

> oh yeah, int. (and adj.): As an exclamation or interrogative: expressing incredulity, disbelief, scepticism, or interest; 'really?' Also as adj.

Though this definition as a face threatening argument challenging interjection does indeed explain some of non-dialogical uses found in the COHA (see 4 below), it is far from comprehensive and it overlooks the DM functions that will be investigated in this chapter.

(4) This vile NAMBLA group was formed in 1978 and calls for the "empowerment" of youth in the sexual area. It says it does not engage in any activities that violate the law. **Oh yeah**? What about the fact that NAMBLA was involved in funding an orphanage in Thailand that allowed grown men to rape and molest the children who lived there? (COHA Non-Fiction: The No Spin Zone 2001)

This multifunctionality of *oh yeah* as an interjection (4), adverb where it marks an affirmative response and DM (see 1 -3 above) demonstrates the difficulty that one encounters when trying to define its characteristics and functions. In the literature of linguistic studies, *oh yeah* is often only mentioned in relation to *oh*. For example, several studies mention its existence as merely a common variant of the discourse marking functions of *oh* (e.g. Schiffrin 1987: 91; Biber *et al*. 2000:1077; O'Keeffe & Adolphs 2008). When discussing it as a variant of *oh,* however, it is essential to differentiate *oh* as an interjection and as a DM. The interjection *oh* is considered a complete utterance that "encodes an entire basic message typically involving the speaker's

emotional state" (Fraser 1990: 391). Similarly, Schourup (1985: 25-6) discussed *oh* as an 'evincive interjection' that indicates that thinking is now occurring or has just now occurred. In particular, Schourup indicates three sub-types of the 'evincive interjection *oh'.* The first type indicates that the speaker has just become aware of something, a sudden strong emotion, or that the speaker should perform some speech act. The second sub-type indicates that the speaker has paused to make a decision or choice between equally available alternatives. The third sub-type is close to the second sub-type, but it indicates only casualness (in a response marker function), and it could easily be replaced by *well*.[51] By contrast, when *oh*, or even *yeah* for that matter, has a text-organizing and/or interpersonal function, it functions as a DM (Aijmer 2002: 98). As a DM, *oh* has overlaps with other DMs such as *well* and *now*. These overlaps occur when it is placed in the left-hand position of the sentence/utterance, and it has a loose attachment to the preceding information given in the discourse (Aijmer 2002: 98-9). It is for this reason that *oh* and *oh yeah* are frequently labeled as utterance launchers (Biber *et al*. 1999:1076), reception markers (Jucker & Smith 1998), or information management markers (Schiffrin 1987).

Because *oh* and *oh yeah* are considered to be highly emotive and interpersonal in nature, most of the past studies that address (*oh*) *yeah* concentrate on oral speech data where it is used as a reception/information managing marker. For example, in their discussion of utterance launchers, Biber *et al.* (1999: 1076) show that *oh yeah* often marks the initial part of a speaker taking his/her turn in response to information presented by their collocutor. Furthermore, they argue that *oh*, *oh yeah* and other combinations have highly conventionalized functions that introduce or respond to utterances in a way that retains some of its core (or original) interjection characteristics that expresses some degree of surprise, unexpectedness, or emotive arousal

[51] In this discussion of *oh* as an 'evincive interjection' Schourup (1985) follows closely James' (1974) interpretation of *oh.*

(1999: 1083). In the same vein, Aijmer (2002) investigated the high multifunctionality of *oh* and found that it was most frequently used in oral conversation and served a variety of functions that varied depending on its collocates, position, prosody, discourse type and position. However, in general Aijmer (2002: 151) found that *oh* has an attention-getting or intensifying function that serves to emphasize the upcoming message. It is precisely these forms of *oh yeah* in these attention-getting or intensifying functions that have been innovated and changed into journalistic prose (1)-(3).

Another important study that investigates the DM *oh* and comments on how it is used to highlight information is Schiffrin's (1987) chapter on *oh*. Schiffrin (1987: 100-1) globally designates *oh* to be a marker of information management that marks shifts in the speaker's orientation (objective and subjective) to information as the flow of information occurs during discourse. Though this work does not discuss *oh yeah*[52] separately, it is briefly mentioned when discussing *oh* and *oh yeh* as backchannel signals that often alternate with other signals of hearer attention such as *yeh* and *mmhmm* (*ibid*. 94). In this same work, there is much discussion of *oh* serving the purpose of displaying one's subjective orientation or 'evaluation' of the information being presented in the discourse *(ibid*.: 95). Furthermore, it is also shown that *oh* can display the speaker/hearer alignment toward their own discourse or the discourse of others (*ibid*.: 100).

In another study that observes the DM uses of (*oh*) *yeah*, Jucker & Smith (1998) investigate a variety of DMs that include what they refer to as 'reception markers'. Despite not discussing it in the analysis as such, this study lists *oh yeah* as a reception marker in their table of DMs that were used in their conversational data (1998: 176). Furthermore, and more importantly, this study analyzes the similarities and differences between the two separate DMs *oh* and *yeah* respectively. For example, in an investigation of the three most common reception markers in

[52] In Schiffrin (1987) *oh yeah* is written as *oh yeh.*

their data (*yeah*, *oh* and *really*), they find that they represent a continuum in the way in which speakers mark the ease of integration of new information (1998: 182). *Yeah*, for example, marks information that fits easily with preexisting assumptions; *oh* marks information that requires some extra processing in order to be integrated; and *really* is used to indicate that the hearer needs further confirmation before the information can be integrated. Interestingly, despite stating that these two DMs have specialized functions vis-a-vis the ease of processing new information, no explanation is given that explains the function or reason for the usage of the combined reception marker *oh yeah.* Despite not discussing *oh yeah* in detail, this study does show that both *oh* and *yeah* serve reception marking functions. Such similarities could explain why these two separate DMs have fused together to form an independent DM that fulfills various functions in the realms of reception/information management.

Shifting our focus now to the occurrences of *oh yeah* in written texts, it is clear that the prototypical functions of *oh yeah* can be located at the informal and more spontaneous speech end of the spectrum. Ironically, when Schourup (1999: 234) argues that DMs are not as allied to speech as many claim, he argues that this unjustified association of DMs with orality exists because "early work on DMs focused predominantly on conversational items like *well* and *oh*". In the case of *oh yeah,* even its individual components are regarded as highly conversational (see OED definition above). As discussed above, *oh* as an interjection saliently marks the speakers surprise or emotive arousal[53], while its DM functions are strongly connected to the speaker's orientation to the flow of information. *Yeah* is also highly colloquial. It is labeled in the OED as a colloquial variant of the adverb 'yes'. Furthermore, in the literature mentioned above, *yeah* is found to have similar speaker orientation DM functions as *oh* (e.g. Jucker &

[53] See Taavitsainen (1998) for a study of the interjection *oh* used in 18th and 19th century writing where written *oh* is found to function differently in writing than in dialogue sequences. These specialized written functions include a peak-marking function (intensifying emotions or making a change in the tone), a foregrounding function that directs the reader's expectations and, to a lesser extent, a turn-taking function between speaking characters.

Smith 1998). Therefore, a question that arises is why we find such examples in writing? It was observed Rühlemann & Hilpert's (2017) study of *well*, some of its dialogical functions such as word-searching or hedging were adapted in order to mark important lexical items saliently. Furthermore, *well* was also used to saliently mark wordplay for the purposes of marking the author's (inter)subjective stance. Similar results are found in chapters 3 and 4 of this work. The question to be addressed here is, will *oh yeah* behave similarly?

This chapter will also use the framework of Traugott & Dasher (2002: 174-6) to discuss the (inter)subjectification of this DM as we have with *well* and *like.* However, to my knowledge, there has yet to be any research on the process of (inter)subjectification on *oh yeah*. As with the other two DM in this thesis, *oh yeah* develops meanings that express the speaker's attitude or viewpoint (subjectivity) that then followed by use of *oh yeah* intersubjectively. An observation of the individual components of *oh yeah* and its functions shows that it is likely that (*oh) yeah* has originated as an adverbial meaning *yes,* though perhaps with a stronger emotional component due to the addition of *oh*. From there, the usage of *oh yeah* would extend to cases where it has textual (or cohesion making) functions. This would include many of the text-organizing functions that were discussed above where *oh yeah* is used to signal information management as it is often used in speech. Following Traugott's cline of (inter)subjectivity, (see 5 below), these textual functions would be the precursors to the subjective area of the spectrum (cf. Traugott 2010: 34, 1989; Traugott & Dasher 2002: 40).

(5) non-/less subjective – subjective – intersubjective

In the middle of this cline would be functions that mark the speaker/author's attitude or viewpoint. One example that we have already looked at that would illustrate such a subjective function would be example (1) where the author rhetorically highlights the last and most important aspect of his/her list of nouns explaining a professional football team. At the far-right

end of this cline would be intersubjective functions. These functions mark the speaker/author's attention to the addressee's self-image or potential reactions. Example (3) illustrates such a use particularly well. What is also curious is that this intersubjective use of *oh yeah* is offset from the rest of the text by the use of parentheses. These meta-textual comments of oh yeah are used to both comment on the information that the reader has come to read and, more importantly, engage with the readership's potential/imagined reaction to said information. In these parenthetical comments, the author attempts to break the fourth wall and engage with and seek-alignment directly with his/her reader(s). Example (6) below displays another example of these parenthetical meta-textual comments:

(6) With Black Friday, Small Business Saturday and Cyber Monday (**oh yeah**, and Giving Tuesday... after we already blew all our bucks) quickly approaching, it's time to make those Christmas lists. (NOW: Arkansas Democrat Gazette 2017)

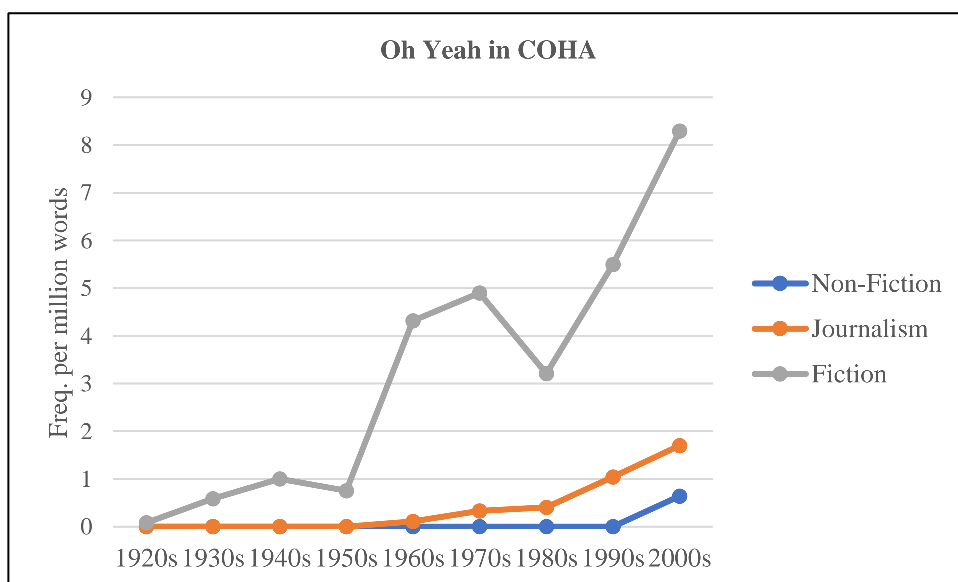
In line with the fact that the intersubjective functions of a DM develop last, these meta-textual examples will be discussed primarily in the analysis of the NOW data. This is done because it is only in the NOW data that we find multiple examples of intersubjective uses. These functions will be shown to not only be rare, but also can be considered the latest development in the other non-dialogical uses of *oh yeah*.

5.3 Trends of *oh yeah* in the COHA data

This section will discuss the frequency trends in the COHA data of *oh yeah.* Like in the other case studies of this work, all concordances of *oh yeah* were extracted from the COHA using a special search string designed to weed out its non-DM uses.[54] The occurrences of *oh yeah* were then analyzed and annotated for position, function and for the type of text that they were taken from. As in all other cases, any examples that were not DM uses such as the adverbial

[54] |(|.|,|:|?|!|... oh yeah .|,|:|?|!|...|)|

affirmation variants were removed from the data after the manual editing process. This process was conducted separately for each of the three sub-genres found in COHA: non-fiction, journalism and fiction. Figure 5.1 shows the overall frequency usage per million words of *oh yeah* across the three sub-genres. Table 5.1 displays the percentages of non-dialogical uses of *oh yeah* in each of the three sub-genres. In comparison with other DMs in the COHA in this study, *oh yeah* is by far the least utilized in writing. Furthermore, when one excludes the data from the fiction sub-genre, it is also clear in this data set that *oh yeah* was virtually unused in non-fiction and journalism until well into the second half of the 20th century (in non-fiction it is not used until the 2000s). Looking at its usage across these three sub-genres, we see that, as in the other studies, it is used significantly more in fiction texts than in the other sub-genres of the COHA. However, as we have also seen in the previous chapters, the uses in fiction are by far more dialogical in nature, because the usage of *oh yeah* comes primarily from fictional quote/dialogue sequences which are constructed to imitate speech.



Genres	Kendall's τ	P (two tailed)
Non-Fiction	NA	NA
Journalism	1	2.00E-02
Fiction	0.83	9.00E-04

Figure 5.1: Diachronic Frequencies for COHA Data of *oh yeah*

Despite *oh yeah* being used less frequently in writing overall than the other DMs analyzed in this study, its proportion of non-dialogical uses are fairly high in journalism and 100% in non-fiction. This can be seen in Table 5.1 below.

Genre	Concordances	Total	%
Non-Fiction	2	2	100%
Journalism	15	40	38%
Fiction	27	381	7%

Table 5.1: Non-Dialogical *Oh Yeah* Percentage for sub-corpora of COHA

As already stated, one unsurprising observation is that the *oh yeah* data from the fiction sub-corpus displays both the most overall uses of the DM in question while also displaying the lowest proportion of uses outside of quotations or dialogues.

The other two sub-genres act differently from the fiction data. As Figure 5.1 shows, *oh yeah* first appears in the journalism data in the 1960s. With the non-fiction data, *oh yeah* is not used as a DM until the 2000s. This diachronic progression in journalistic writing is further analyzed using the Variability-based Neighbor Clustering (VNC) algorithm. The data is grouped together between three stages, though one could have also reasonably made only two by combining the 1990s with the 1960s -1980s (see Figure 5.2 below). In any case, the VNC dendrogram shows that there is a big increase in period 3 in the 2000s. This clustering finds further support if one considers that 50% of all the occurrences of *oh yeah* in the journalism corpus occur in the 2000s. Furthermore, if we combine the data from the 2000s with the 1990s then the percentage is 80%. The timing of the surge at the tail-end of the 20th and beginning of the 21st century is similar to the situations with *well* and *like* in the previous chapters and gives further support to the observations of Rühlemann and Hilpert (2017) involving *well* in the TIME corpus. However, similar to *like*, *oh yeah* enters into writing significantly later than *well* does.

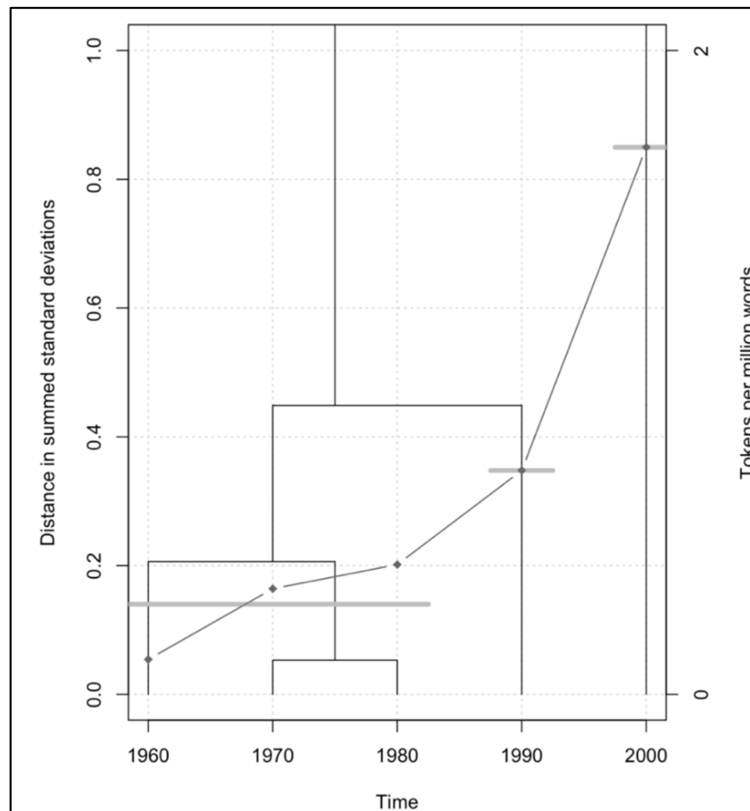


Figure 5.2: VNC dendrogram with overlaid line plot for frequencies of DM *oh yeah* for periods1 (1960 1980), 2 (1990) and 3 (2000) in Journalism and Non-Fiction data

5.4 Qualitative analysis of non-dialogical *oh yeah* in the COHA data (1960s-2000s)

This analysis has demonstrated so far that starting in the 1990s, and even more so in the 2000s, there is a surge of change in the usage of *oh yeah* in journalistic writing. Some similar examples, though very rare, were found for non-fiction writing. This section will address the type(s) of articles *oh yeah* is used in and in which functions this increase has been occurring. This section will begin by looking at the functions in which non-dialogical *oh yeah* is being used. It has already been discussed in section 5.2 above that *oh yeah* has developed functions in writing that show the author's (inter)subjective stance. In the case of *oh yeah*, instead of a gradual development of the subjective functions, as was seen particularly with *well,* what happens is that there is a sudden appearance of one subjective construction in particular that dominates its usage from the 1990s. This single function that suddenly appears in the COHA data is the end of the list marker that was discussed in (1).

For the COHA data, most of the non-dialogical examples fit more into the subjective area of the cline with only one example being highly intersubjectified. Established intersubjective constructions such as those that resemble using *oh yeah* as a response marker (example 2) or marking the beginning of a meta-textual comment (example 3) will be discussed further in 5.6. The subjective uses of *oh yeah* in writing, which are utilized in the COHA data, include examples such as (1) above and (7) below which saliently mark the end of a list where the most obvious or important item is given last. It does so in a fashion that resembles the convention that it is an afterthought even though it is, in fact, the 'hot' item of the list. Coincidentally, this function is very similar to a sub-function of WCM *well* that also marked the end of a list in the same fashion. These end of list markers will henceforth be called 'faux spontaneous additives' because they expand on the dialogical function of *oh yeah* where it marks the speaker spontaneously remembering something.[55] This is by far the most common usage of non-dialogical *oh yeah* in the COHA. Other functions are rarer and mimic more literally the oral dialogical uses of *oh yeah*. One of these functions marks a 'faux quote' in which the imaginary speaker is remembering something (see example 8). Another usage of *oh yeah* is the function where *oh yeah* acts as a challenge interjection like example (4) above and (9) below. In example (9), *oh yeah* is used to rhetorically challenge the evaluation of information presented just to the right of it in the text. In example (10), *oh yeah* is used rhetorically by the author to respond to the reader's potential/imagined retort to what has been written in the article. In fact, this is the only example where *oh yeah* is intersubjectified. In this particular example, *oh yeah* is used as a response marker to an imagined question by the readership.

(7) The Three Great Sins Against Children -- those that invoke the law, bring down the wrath of the community and the hounds of child protective services agencies -- have an unwritten hierarchy: Physical abuse and sexual abuse come first, and then, **oh yeah**, there's that neglect thing. (COHA Magazines: Mother Jones 1999)

[55] In the COHA data, all 10 faux spontaneous additives are end of list markers. In the NOW data, this category expands and has other sub-subjections.

(8) If you walked around downtown Washington this spring, you'd have found virtually no sign that the Wizards were in the playoffs -- no Wizards pennants in store windows, no overheard conversations in coffee shops about last night's winning three-pointer. When I talked to friends -- most of them, like me, transplants from other cities -- about the Wizards, the typical reaction was "**Oh yeah**, they're in the playoffs, aren't they?" They, not we. (COHA Magazine: Washington Monthly 2006)

(9) Ski resorts put a lot of energy into trying to convince you and me that skiing is a "family sport." **Oh yeah**? When I start toting up the costs for, say, a family of four, the figures flabbergast me, and I think of skiing as an expensive sport for a family of one. Many resorts do offer free skiing and/or lodging for kids six and under, which helps, but what about skiers older than six? (COHA Magazines: Skiing 1994)

(10) Because what they almost never show in the magazines is pictures of people out there actually buying the food. **Oh yeah**, I know. What about those charming photos of pretty young chefs choosing fresh strawberries and placing them in a wicker basket at the open-air farmers' market? Get real, girlfriend. I mean, it would be nice and all. But since when can you buy Pampers and tuna fish at the farmers' market? (COHA Newspapers: Washington Post 1994)

The breakdown of the categories of these different functions of non-dialogical *oh yeah* can be observed in Table 5.2. What is immediately obvious is that the faux spontaneous additive function (that marks the end of a list) of *oh yeah* is the clear front-runner. In fact, it accounts for 67% of the COHA non-dialogical data. It is also the most conventionalized of all the functions as it follows a very predictable construction. It is to the left of the final lexical element of a list and surrounded by punctuation. Considering that this end of list marker has a fixed structure and that it is the single most used construction in the *oh yeah* COHA data, it is reminiscent of pred *well*. For example, like pred *well* in the 1990s, it appears suddenly as a major function, and its usage continues to grow into the 2000s. The apparent sudden appearance and fast growth of *oh yeah* in this function is reminiscent of how slang terms enter the lexicon (Adams 2009: 15). It will therefore be important to verify if it loses its dominance like pred *well* did in the 2010s or if this function of *oh yeah* can maintain its place. The other three functions of *oh yeah* are much more rarely used and do not appear to have developed into fixed constructions like this faux spontaneous additive end of list function. Overall, the analysis of

the uses of *oh yeah* in written functions demonstrates the usefulness and increased conventionalization of utilizing *oh yeah* to display the author's subjective stance.

Function	Frequency	Percentage
Faux Spontaneous Additives	10	67%
Faux Utterance Launchers	3	20%
Challenging Interjections	1	6.5%
Intersubjectified Comment	1	6.5%

Table 5.2: Function distribution of non-dialogical *oh yeah* in journalism COHA data

When these examples from journalistic prose are compared with the types of oral uses that they resemble, it becomes clearer why these uses have been expanded into writing to serve subjective and sometimes (inter)subjective functions. For example, the faux spontaneous additive (or end of list) usage of *oh yeah* that was discussed in (7) bears a striking resemblance to how one uses *oh yeah* when they suddenly remember something (11) or want to make it look like they just remembered something as (12) below demonstrates. These examples come from segments of reconstructed speech from a newspaper article.

(11) A guy came out of the restaurant and spotted Eddy Folson. He gave his name as Tommy Condon and reminded Folson that they had worked together up at Seabrook about two summers ago. "**Oh yeah**," Folson remembered. "How you been doing? " "Eehh," Condon said with a shrug of his shoulders. "You working? "Folson asked him. "Once in a while," Tommy Condon said. "There' s not much around, though. How about you?" (COHA: Boston Globe 1982)

(12) The couple went into the restaurant and ordered a pizza to be delivered to Cohen's home. And then Cohen's father-in-law asked for something else, something that Cohen believes he wouldn't have thought of in his younger years." He said, "'**Oh yeah**, **and one more thing**,'" Cohen said. "'I'd like you to deliver us with it.'" (COHA: Atlanta Journal Constitution 2007)

When one compares (11) and (12) to the end of list faux spontaneous additive example in (7), it is obvious that (7) and (12) have more in common with each other due to the fact that in both cases, instead of marking the cognitive process of remembering something, *oh yeah* is used in

these cases rhetorically. In (7) it is marking what the author considers the most important item of a list, and in (12) it is used to introduce the usual request that the man is making to the pizza delivery person. Interestingly, (12) includes the addition of "and one more thing" to *oh yeah*. Though *oh yeah and one more thing* could easily be used in the end of list written function as well, it does not appear in the COHA or even the NOW data that will be discussed in 5.6.

As the function's name indicates, the faux utterance launchers are also extensions of the oral uses of *oh yeah* that serve to start an utterance, particularly in cases where they are responding to a question. An example of this in reconstructed dialogue from the COHA data can be seen in (13). As discussed in the previous discussion in 5.2, this usage of *oh yeah* to mark the beginning of an utterance also carries with it elements of surprise or emotive arousal. In this particular case, the speaker is using *oh yeah* in a hedging function *(oh yeah I meant to tell you*) that downplays the fact that he has been caught by his superior about an action he took concerning 200 children. The one intersubjectified example from the COHA (10) is also an expansion of this form of utterance launcher/response marker.

(13) About that time, the late Archbishop Coleman Carroll called Walsh, asking if he knew why "someone from the State Department is trying to reach you about taking in 200 children." "**Oh yeah**, I meant to tell you about that, "stammered Walsh." I have made a commitment in our behalf," thinking to himself that "saving 200 children from communism is worth my career." (COHA: Washington Post 1998).

Examples (11) – (13) demonstrate the origins from which the written uses of *oh yeah* expanded. In journalistic prose, *oh yeah* is used rhetorically to mark the stance of the authors. However, the ways in which *oh yeah* is used to do this resemble the textual (coherence) uses of *oh yeah* from oral speech where *oh (yeah)* is used to mark shifts in the speaker's orientation (objective and subjective) to information as the flow of information occurs during discourse (cf. Schiffrin 1987). When used in journalistic prose, however, *oh yeah* is used primarily to mark the stance

of the author, while it also highlights essential linguistic items which the author deems important or as meriting special attention. Therefore, as was seen with *well* and *like* in the previous two chapters, the non-dialogical uses of *oh yeah* work at the level where they are used by authors to pretend that they are suddenly remembering something in a list or they are marking their emotive arousal when responding to a question for the purposes of highlighting their subjective stance. Moreover, they are also using *oh yeah* in a way that indexes informal conversation between close collocutors in a strategy to get the reader to align with them in this evaluation (e.g. second and third order indexicality as discussed in 2.2).

Having analyzed the uses of *oh yeah* in journalistic prose and having compared it to the oral uses from which the written functions have expanded, it is not time to address in what type of articles *oh yeah* is used in journalistic prose. Table 5.3 demonstrates that non-dialogical *oh yeah* is used primarily in the informal categories of journalism.

Article Type	Frequency	%
Entertainment	4	27%
Lifestyle	4	27%
Sports	3	20%
Other	2	13%
Parenting	1	7%
Culture/Culinary	1	7%

Table 5.3: Types of articles where *oh yeah* is used in journalism COHA data

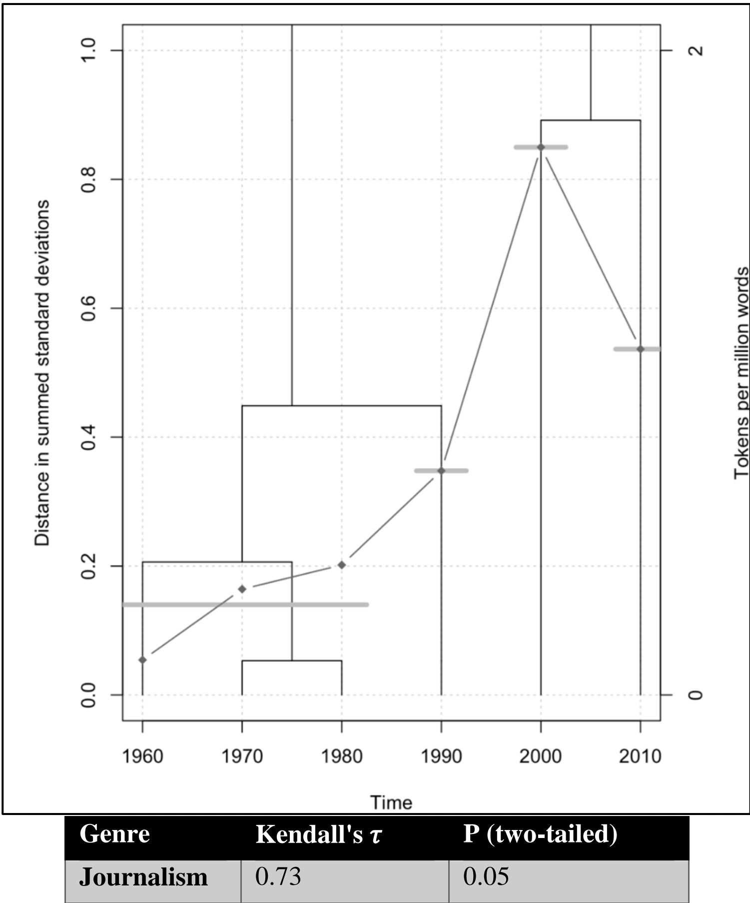
Furthermore, Table 5.3 shows that the top three categories for journalism that included examples of non-dialogical *oh yeah* were entertainment, lifestyle and sports articles. Unsurprisingly, these are the areas of journalism that the previous chapters have also highlighted as being the most prone to show (inter)subjective uses of other saliently dialogical DMs. It is also worth noting that the top three categories, entertainment, lifestyle, and sports, make up 74% of all uses. This also shows that *oh yeah* is much more adapted to informal

journalism than *well* and even *like* for that matter in the COHA data. We can therefore tentatively point to the usage of *oh yeah* having the salient function of being used as an anti-formality device for the purposes of creating a sensation of intimacy between author and the readership, though this will have to be verified in the 2010s NOW data. Furthermore, the two examples from non-fiction also come from informal sources as well. One comes from a book by Bill O'Reilly, a political commentator on a previously popular political talk show program on Fox News. The other example comes from a book on parenting, a category that also appears on the list of Table 5.3 for the journalism data.

The analysis of the COHA data for *oh yeah* once again indicates a sudden rise in the usage of a DM from the 1990s and furthermore into the 2000s. This coincidentally matches what we have seen with the development of (inter)subjective word-choosing marker uses of *well* and *like* during the same period. Though the written functions are not completely confined to functions that highlight lexical items, the single largest function, the faux spontaneous additive, does work as an end of list marker. The similar trajectories of these three DMs could potentially be explained in terms of their slight overlaps of terms among their prototypical dialogical uses, which then expanded into the non-dialogical ones. These overlaps include initiating turns, signaling the speaker's response/evaluation of the ongoing conversation and marking self-correction. Though overall *oh yeah* enters into usage later and is clearly used less frequently in the COHA data than *well* and *like* are, it enters into writing and increases in frequency in a very similar way as has already been seen with *well* and *like*. Moreover, like *well,* from the 1990s the use of *oh yeah* is dominated by one particular construction, the end of list faux spontaneous additive. If this dominance will continue into the 2010s or if this particular construction will lose it dominance like pred *well* is to be determined.

5.5 Trends of *oh yeah* in the News on the Web Corpus (NOW)

The dataset for *oh yeah* in the NOW ranges from 2010 through mid-2018 while the frequency of *oh yeah* usage as a DM was calculated to represent its frequency per million words as was done with the COHA data. Interestingly, as can be observed in Figure 5.3, *oh yeah* is actually used less often per million words in the 2010s NOW data than it was in the COHA 2000s data. Despite this drop in frequency during the 2010s, *oh yeah* is still used more per 1 million words than it was in the 1990s COHA data. The reason(s) for this are for the moment unclear. One possibility is that the usage of *oh yeah* simply plateaued in the 2000s. Only further investigation of *oh yeah* in the decades to come will reveal what is happening.



Genre	Kendall's τ	P (two-tailed)
Journalism	0.73	0.05

Figure 5.3: VNC dendrogram for Journalism data with NOW 2010s data added with overlaid line plot for frequencies of DM *oh yeah* for periods1 (1960-1980s), 2 (1990-2000s) and 3 (2010s) with Kendall's τ score

Despite the drop in the 2010s, however, something interesting is happening with the development of *oh yeah*, as it is used in journalistic prose. As attested in Table 5.4, there is a

rise in the percentage of the non-dialogical use of *oh yeah* in the 2010s data. In the 2010 NOW data, this percentage is up to 58% of the total usage, which is 18% higher than it was in the previous decade in the COHA data. This indicates that though *oh yeah* may not be continuing its steady rise in frequency of usage overall, it is, in fact, solidifying its non-dialogical functions in journalistic writing.

Decade	% Non-Dialogical Functions
1960	0%
1970	0%
1980	50%
1990	42%
2000	40%
2010	58%

Table 5.4: Non-Dialogical *Oh Yeah* Percentage for journalism data 1960-2000s (COHA) and 2010s (NOW)

As was the case with the NOW investigation of *like*, there is a sizable number of examples of non-dialogical *oh yeah* that come from the comment sections of the news articles. As these comments can come from just about anybody, the case could very convincingly be made that these examples should not be included in this analysis of the language practices of journalism. If we were to revise Table 5.4 to exclude these examples from our dataset completely, the percentage of non-dialogical *oh yeah* would only drop slightly to 53%, which does not take much away from the argument that non-dialogical *oh yeah* is clearly gaining ground in journalistic prose. In terms of the overall influence of these comments in terms of data frequency in Figure 5.3, 118 of the total examples of *oh yeah* (roughly 11% of the total data) came from these comment sections. For these reasons as well as others mentioned in previous chapters, the decision was made to include comments in this analysis as well.

5.6 Qualitative analysis of non-dialogical *oh yeah* in the NOW data (2010-2018)

The same subjective functions of non-dialogical *oh yeah* that were found in the COHA data are also present in the NOW data with the addition of there being the development of meta-textual function. Before contrasting the functions of *oh yeah* in the NOW data with those of the COHA data, a general explication of the properties of *oh yeah* in NOW data will be given.

As the analysis of the COHA data displayed, the non-dialogical uses of *oh yeah*, in particular the end of list faux spontaneous additive, sometimes appear sentence-initially and sometimes sentence medially. Therefore, two separate collexeme analyses were conducted. Table 5.5 shows the top twenty collocates located at 1R while Table 5.6 does so for all collocates at 1L for *oh yeah* based on their collostructional strength based on the Fisher Yates exact test.

Unsurprisingly, the two tables demonstrate that *and* is a strongly correlated collocate to appear either immediately to the right of left of *oh yeah*. This suggests that *oh yeah* continues to be used at the end of hypothetical lists as a *faux spontaneous additive* in the contemporary NOW data. The co-DM *well* is also found at number 14 on Table 5.5. As already observed in the previous case studies, it is not uncommon for DMs to co-occur in close proximity or in sequences (e.g. Haselow 2019; Lohmann & Koops 2016). Upon further analysis, however, it was found that these DM sequences only occurred in dialogical sequences and not in journalistic prose. Number 16 in Table 5.5 and number 11 in Table 5.6 is *because* which, upon inspection in the corpus data, was found to occur in cases where *oh yeah* functioned as a faux spontaneous additive marker or where *because* was used as a real or faux utterance launcher in both quotations and non-dialogical situations. Other collocates in Table 5.5 include the adverbs such as *definitely* and *absolutely*, which hints at the dialogical uses of *oh yeah* where it is used to mark one's affirmation to previously stated information. Similarly, Table 5.6 includes collocates such as the verbs *thought* and *say* and *said* that indicate common oral utterance

launchers such as '*thought oh yeah'* and '*said oh yeah'* which are also common in oral speech. As this is an investigation of the non-dialogical uses of *oh yeah* as a DM in writing, we will leave these adverbs and utterance verbs out of the following analysis.

	COLLEX	ASSOC	COLL.STR.FYE	SIGNIF
1	I	attr	126.15928	*****
2	AND	attr	80.48529	*****
3	HE	attr	44.67061	*****
4	IT	attr	28.66818	*****
5	ABSOLUTELY	attr	25.98791	*****
6	WE	attr	25.5253	*****
7	THEY	attr	21.55295	*****
8	DEFINITELY	attr	20.26556	*****
9	THERE	attr	19.9534	*****
10	THAT	attr	19.69558	*****
11	YOU	attr	12.48085	*****
12	SHE	attr	11.29173	*****
13	SURE	attr	8.03359	*****
14	WELL	attr	4.69871	****
15	MY	attr	4.57658	****
16	BECAUSE	attr	4.36507	****
17	RUN	attr	4.02234	****
18	THIS	attr	3.92726	***
19	LET	attr	3.89992	***
20	ONE	attr	2.82883	**

Table 5.5: Collexeme Analysis of NOW *oh yeah* 1R

	COLLEX	ASSOC	COLL.STR.FYE	SIGNIF
1	AND	attr	105.60345	*****
2	LIKE	attr	13.3985	*****
3	SUPREMACIST	attr	13.0002	*****
4	SAY	attr	11.59162	*****
5	THOUGHT	attr	8.50117	*****
6	SEASONS	attr	6.92239	*****
7	HIM	attr	6.13537	*****
8	AGAIN	attr	6.09973	*****
9	OUT	attr	2.46398	**
10	ME	attr	2.04481	**
11	BECAUSE	attr	1.63417	*
12	UP	attr	1.12317	ns
13	TIME	attr	1.07903	ns
14	IT	attr	0.74323	ns
15	YOU	attr	0.43635	ns
16	WE	attr	0.39874	ns
17	SAID	attr	0.36348	ns
18	THAT	rep	0.488	ns
19	ON	rep	1.46641	*
20	'	rep	6.47583	*****

Table 5.6: Collexeme Analysis of NOW *oh yeah* 1L

What these collexeme analyses do for the discussion at hand, however, is give further credence to the conventionalized use of *oh yeah* as a highly subjective marker that mimics the act of spontaneously highlighting important information (the faux spontaneous additive function). The usage of these spontaneous additive functions in journalistic prose is, of course, not actually spontaneous. An author of an article, or any form of writing for that matter, has the luxury of being able to prepare and edit their work before publishing it. This is an option that someone who is speaking orally does not have. Of the sub-functions of faux spontaneous additives in the NOW data, the most frequent and the most conventionalized form is the same end of a list marker that was observed in the COHA data. As before, this end of list construction features *and* occurring either immediately to the left or right of *oh yeah*, which is then followed by the last item of a list. In the NOW data, however, though this end of list marker is still, by far, the most frequently used single usage of non-dialogical *oh yeah*, it is also apparent that things are

more varied than was seen with the COHA data. Table 5.7 shows the breakdown of the major functions of non-dialogical *oh yeah*. All of the major categories will be discussed in the following sub-sections.

Functions	Frequency	%
Faux Spontaneous Additive	429	70%
Faux Utterance Launchers	121	20%
Meta-Textual Comments	37	6%
Other	23	4%

Table 5.7: Functional Categories of non-dialogical *oh yeah*

5.6.1 Faux spontaneous additive *oh yeah*

This discussion of the major functions of non-dialogical *oh* yeah will begin with the most frequently occurring. As Table 5.7 indicates, by far the most frequent function involves the faux spontaneous additive function. When looking at the totality of the spontaneous additive markers, which almost exclusively have either *and* immediately to its right or left (see Table 5.8), 52% of them are end of list markers (see example 14). Once again, this end of list function is the most frequently used individual construction. However, making up 36% of all non-dialogical uses of *oh yeah*, it no longer constitutes the vast majority of the data. Interestingly, similar usages can be found in social media dialogue like example (15) which is a Facebook post from a non-journalist. In this post, the person lists the reasons why people should not feel outrage over the arrest of the founder/director of WikiLeaks, Julian Assange. All of the author's arguments for why people should not view Julian Assange as a hero are listed. At the very end of the post, as if he had almost forgotten, the author gives what he views is the most uncontroversial reason for not giving Julian Assange praise. Namely, he ends with the assertion that he is "most likely a sex offender" on top of everything else.

Combination	Frequency	%
And, oh yeah	156	26%
Oh yeah, and	133	22%
Combined	289	48%

Table 5.8: Total Frequencies of Oh yeah collocating with and its percentages of the total non-dialogical data

(14) If you need a refresher, Rhaegar is Dany's older brother, the son of the Mad King Aerys and, **oh yeah**, Jon Snow's daddy. (NOW: Huffington Post 2016)

(15) Don't let Julian Assange's image as a romantic hero defying bullies fool you. He not only released cables that exposed U.S. military operations, he exposed ones that humiliated and undermined American diplomats seeking to strengthen alliances and avoid conflict. His website, Wikileaks, has become a clearinghouse and accomplice for Russian intelligence, who released embarrassing info about Hilary Clinton though Wikileaks in order to undermine her campaign and get Trump elected. He's not protecting democracy, he's helping the world's most reactionary regimes undermine democracy. He's not a warrior for transparency, he's an anarchist who believes in no secrets at all. You can make a case for stronger transparency without supporting this dude, and you should. **Oh yeah** and he's most likely a sex offender too. (Facebook 2019)

In the other 48% of the faux spontaneous additives, *oh yeah* functions in other similar subjective functions. In these cases, it is also utilized to display an essential piece of information in their text as if it were an afterthought or something that they had nearly forgotten. Example (16) does exactly this. In this example, the author mimics remembering something and intentionally interrupts the normal flow of the sentence. This shares many of the characteristics with the end of list construction. Furthermore, Tables 5.5 and 5.6 also showed that *because* is among the words that is most attracted to *oh yeah*.

(16) Also, there was a nice moment when Hayden pep-talked the ghost of the Black Dahlia (Mena Suvari) because, **OH YEAH**, the Black Dahlia was in this! (NOW: TV.com 2011)

In the discussion of an episode of the first season of the television series American Horror Story, the author expresses his/her subjective attitude of the episode's dense plot by utilizing *oh yeah* to stress that a character based on the real life murder victim, the Black Dahlia, makes an

appearance in the episode. The author does so by using *oh yeah* in a way that resembles the dialogical information management function of *oh yeah*. This is done to convey to the reader that there is yet another huge twist in the episode's plot. In the future, it will be important to verify if *because oh yeah* becomes a more dominant construction like the end of list marker or not.

5.6.2 Faux utterance launcher *oh yeah*

The second most frequent function for non-dialogical *oh yeah* is globally labeled as the faux utterance launcher function. Within this category, there are two main executions. The more frequent of the two involves the author setting him/herself up with a rhetorical question to which *oh yeah* marks the author's response. Example (17) is one example of this. It is an interesting example because it also occurs near an example of pred-*well* and because it is used to express the author's perspective on human eating habits. This situation in which these dialogical DMs are used in written functions in the near vicinity of each other is something that has already been observed in other examples..

Similarly, example (18) displays the author/reviewer's attitude about a practice of the Nintendo company by responding sarcastically to his/her own question about how many times it was necessary to buy the same game in order to play it on the new gaming systems.

(17) Every meal looks like my undoing. Salt looks like lye. Oil looks, well, really oily. We eat that why? **Oh yeah**, because it's a thick golden pleasure-delivery system. (NOW: GQ Magazine 2013)

(18) You have like a billion amazing, beloved games in your back catalogue. The best of them are still good enough to stand alongside any new release. Yet it is still so difficult to buy most of them, and when we do buy one, we get no guarantee that it'll work on any other Nintendo systems we own. How many times have I bought Earthbound in the past year? **Oh yeah**, twice. I've bought it twice. (NOW: Kotaku 2016)

These two examples, along with the rest of the rhetorical response markers, follow the same basic structure. The author poses the question to him/herself, and then responds to said question with *oh yeah* appearing at the left periphery of the response. In many ways, this resembles what has been seen with *oh (yeah*) in past studies that used oral data. In these studies, *oh yeah* is labeled an utterance launcher, which the speaker uses to take their turn in response to the question posed by their collocutor (e.g. Biber *et al.* 1999: 1076). The difference between examples from oral conversation and the usage of *oh yeah* in (17) and (18) above, however, is that these two examples come from monologue texts where no other interlocutor is present. Furthermore, unlike in spontaneous speech, these authors are clearly capable and even required (these two examples come from reputable magazines) to revise their articles before they are published. Therefore, this indicates that certain conventions are being utilized in monologue journalism articles by authors in order to indicate their stance to their readership in a way that is engaging and interpersonal like oral speech. Furthermore, they can also rhetorically highlight lexical items like in (14) where it is used to flag the upcoming metaphorical wordplay that likens oils in food to "a thick golden pleasure-delivery system". In (15), it is marking the adverbial "twice" as in he/she had to rebuy a favorite video game two times within the last year to keep up with the new Nintendo gaming systems.

Other faux utterance launchers, though considerably less frequent and less conventionalized as the rhetorical response marker *oh yeah,* involve uses like example (19) below:

(19) Flying in the 60s. What Air Passengers Did Before In-Flight Entertainment. **Oh yeah**, this looks a lot like my last flight, minus the woman with the hacking cough in front of me, and the screaming kids behind me, and the mystery stains on my tray table. (NOW: Minneapolis Star Tribune 2015)

Here as well, the author starts the new sentence with the use of *oh yeah* that resembles a spontaneous reaction to the sentence that precedes it. Furthermore, it intensifies his/her sarcastic

comment that follows, which lists all the things that make the photos of flying in the 1960s different from the experience one has in commercial airliners today.

Globally, *oh yeah* is utilized as a faux utterance launcher in a fashion that resembles the spontaneous conversation conventions of *oh yeah* as an utterance launcher that marks the speaker's management to the information previously mentioned (e.g. Biber *et al*. 1999: 1076; Jucker & Smith 1998). However, the examples from the NOW data come from monologue texts where the turn taking system of oral discourse is not in play. Therefore, what these usages in journalism display is clearly not a simple adoption of an oral discourse convention of language into writing. Rather, what we see is the innovation of a saliently oral convention of language that is used to mimic the oral convention, but at the same time adapts it to mark one's rhetorical stance in writing. Some of these uses appear to be developing into specific subjective functions that also can serve the double function of highlighting special lexical items like we have seen with the end of list marker.

5.6.3 Meta-textual comments

So far, subjective uses of *oh yeah* in the NOW data have been discussed. In addition to these rhetorically subjective uses, however, there are also examples of *oh yeah* in intersubjective meta-textual comments. As was the case in the COHA data, these uses are relatively rare compared to the others. Though these meta-textual uses are still significantly less frequent than the other functions, the larger NOW corpus provides us with more examples, which allows for a better analysis of how authors utilize *oh yeah* in this way. These uses are also the most intersubjective of any other comments discussed in this chapter because they are offset from the main text and because they appear to address the imagined reactions of the readership.

Examples of these uses of *oh yeah* are constructed as indicated in (20) below. *Oh yeah* is preceded by a parenthesis and then followed by a comment which is then followed by the closing parenthesis.

(20) "First of all, Miss Thing," she began (**oh yeah**, we're not kidding), "does your man not have a mouth of his own with which to speak?"(NOW: E! Online 2016)

The example above is particularly salient, because the author of the article is using this offset comment to interrupt a quote from a video posted online from a celebrity in which she attacks another celebrity for speaking on behalf of her husband. It is also clear that the author is displaying his/her personal attitude toward this very confrontational video post, in particular the utilization of the term 'Miss Thing'[56]. The fact that the comment is offset from the regular text also indicates that the author is doing so in order to make it clear that he/she is addressing the reader(s) directly. The convention of offsetting comments in parentheticals in order to mimic directly seeking alignment with the reader(s) has already been seen in online writing such as in the 'sarcastic much?' investigated by Hilpert & Bourgeois (2020)[57]. Looking at an example of this 'sarcastic much?' construction shows a striking resemblance to (20) above.

(21) Still I kept at the classic literature because it was important to me that others respect my intelligence (**damaged by high school much? :-P**). Naturally when I decided to write a novel, it came out as literary fiction. (*ibid.*)

Examples (20) and (21) come from very different text types. Example (20) is from a celebrity news website while (21) was taken from a personal blog. However, despite coming from

[56] Urban Dictionary gives the following definition for the utilization of Miss Thing: "A sarcastic term for haughty, puffed-up women (especially women of color) who think that they are really something when, in fact, they are not". Urban Dictionary can be consulted online at: <https://www.urbandictionary.com/define.php?term=Miss%20Thing&utm_source=search-action >

[57] For more discussion of the 'sarcastic much?' construction see also Adams (2014) and Gutzmann & Henderson (2019).

different text-types, there are some key similarities between these two intersubjective constructions. First, both of these texts are monologue texts, meaning that they do not have an immediate audience in their surroundings nor an implied message board from where the author expects an immediate reply. For what purpose, therefore, would an author use such dialogically styled constructions? Furthermore, why are there similar intersubjective phenomena found in the writing of a personal blog and a published media source? It is important to examine this issue because the two text types are usually viewed as having significant differences in formality and purposes. For example, while digital news sources are still primarily following the conventions of print based news, which is concerned with transmitting information (e.g. Biber & Conrad 2009: 118), social media is used in a much more interpersonal fashion that allows for both informing and contributing back and forth through message boards as allowed on popular social media sources such as Twitter, Facebook, Instagram and personal blogs (Zappavigna 2012: 2). One observation about the usage of meta-textual *oh yeah* in journalism, as seen in example (20) above and example (3) which is reproduced below for convenience in (22), is that these examples come from the informal branches of journalism. In particular, these examples come from sources or sections of news sources that are dedicated to reporting entertainment news such as celebrity gossip columns and television show recaps.

(22) After the death of Oberyn Martell last season, Cersei is concerned about her daughter Myrcella (**oh yeah,** remember her? More on her below) who has been living in Dorne since season two. (NOW: USA Today 2017)

Based on these observations, one could speculate that these meta-textual comments are coming from web-only based sources since their nature is so strongly interpersonal. However, though meta-textual comments are found in the majority of cases in articles from web-only sources, Table 5.9 shows that a sizable amount of these uses comes from traditional news sources.

Source	Frequency	%
Web Only	20	53%
Print & Web	12	32%
Television & Web	4	10%
Radio & Web	2	5%

Table 5.9: genre distribution of meta-textual *oh yeah*

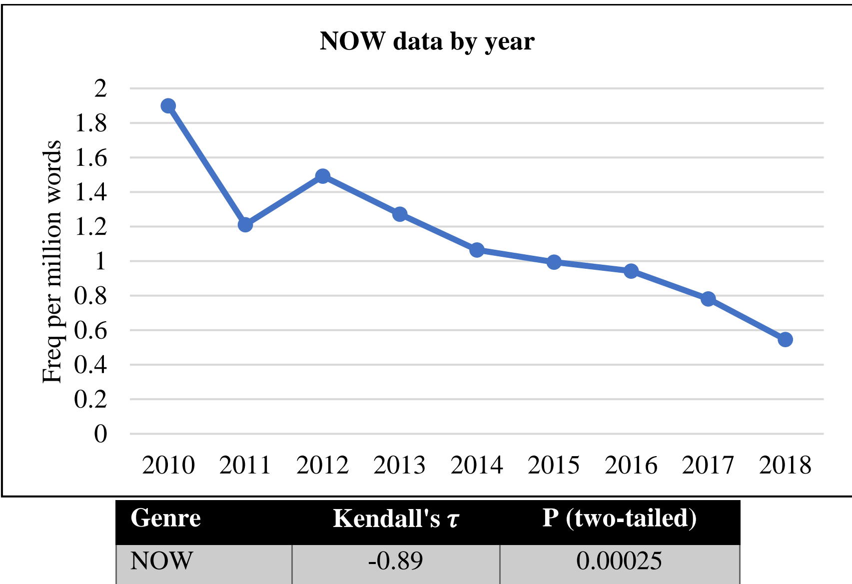
The breakdown of the media sources of meta-textual *oh yeah* indeed shows that 53% of the meta-textual uses of *oh yeah* come from "web-only" news sources. However, the second highest source of meta-textual comments at 32% is traditional news, which continues to appear both digitally and in paper-form. This includes sources such as TIME Magazine, USA Today, The New York Post as well as various local newspapers. The other 15% involve the websites of television stations such as CNN and E! or radio stations such as affiliates of NPR (National Public Radio).

The analysis of the examples of meta-textual *oh yeah* and the breakdown of the news sources that they come from bring us back to the important question about why these uses of non-dialogical *oh yeah* only come about from the 1990s? Another is why the use of non-dialogical *oh yeah* has seen the further conventionalization of subjective and intersubjective functions after the beginning of the 21st century? These questions will be addressed in the next section.

5.6.4 Discussion of non-dialogical NOW data

In the analysis of the NOW data, many of the observations of the COHA data have been confirmed, but there have also been some noticeable differences as well. One noticeable trend that was not continued from the COHA data to the NOW data was a continued rise in frequency per million words. In the COHA data all indications were there that the usage of *oh yeah* was steadily rising in frequency from its tentative start in the 1960s. In the 2010s, however, this steady increase in frequency stops and then declines. This is more apparent if one observes the

yearly decline in the frequency per million words from 2010 to June 2018 (see Figure 5.4). In 2010, interestingly, the frequency of *oh yeah* per million words is near 2 which is approximately the same as the frequency of use per million words for the 2000s COHA data. However, after 2010 the frequency of *oh yeah* steadily drops. This is especially the case after 2012. Despite this decline, however, it is important to mention that the decrease is not so drastic that it drops back down to levels found in the early period before the rise in usage of non-dialogical *oh yeah*, namely the 1960s-1980s.



Genre	Kendall's τ	P (two-tailed)
NOW	-0.89	0.00025

Figure 5.4: Diachronic Frequencies NOW Data *oh yeah* per individual year

Despite this downward trend in the overall frequency of *oh yeah,* however, it has also been shown that the percentage of non-dialogical *oh yeah* compared to its overall usage is actually much higher than what was seen in the COHA data. For the 2010s data overall, the percentage of non-dialogical *oh yeah* in the NOW data is 58%. However, Table 5.10 shows that this percentage got as high as 70% in 2016 and is at 60% for the first half of 2018.

Year	% Non-Dialogical Functions
2010	54%
2011	48%
2012	44%
2013	66%
2014	58%
2015	56%
2016	70%
2017	62%
2018	60%

Table 5.10: Non-Dialogical *Oh Yeah* Percentage for NOW data

Furthermore, in terms of the functions of non-dialogical *oh yeah*, the end of list sub-function of the faux spontaneous additives was, by far, the most utilized single function in both the COHA and NOW data. In particular, *oh yeah* is used as an end of list marker in 67% of all non-dialogical uses of *oh yeah* in the COHA data, and 37% of all non-dialogical uses in the NOW data. Though no longer constituting the majority of all uses, this function of *oh yeah* has maintained its dominance. Furthermore, it is still the most used construction involving *oh yeah* in the NOW. This construction shares parallels with the pred *well* construction discussed in chapter 2 in that it is highly conventionalized and that it is a uniquely written function of a DM that appears suddenly at the end of the 20th century. However, the end of list *oh yeah* does lose ground in the contemporary NOW data. Therefore, like pred *well*, the end of list marker *oh yeah* construction is like a slang term that suddenly comes about in the 1990s, but unlike pred *well* it does not lose its notoriety to the same extent. Though it is common for slang terms to fall out of fashion just as quickly as they appear, it is also the case that some slang terms become mainstream and are maintained in the lexicon over long periods of time . Such an observation was made in Zappavigna (2013:137) in reference to the use of *geek* in microblogging language. She attributed the persistence and 'dilution' of its meaning from pejorative only to a marker of affiliation to it entering the mainstream, a process whereby Moore (2004) claims that the slang of certain subgroups only becomes more 'widespread'. In discussing this, Moore (2004: 61) claims that:

> Basic slang is probably used as much within mainstream society as among subcultures, but its usage still conveys a somewhat subversive message.

In the particular case of the end of list marker function, its survival in the dominant position most likely does not have much to do with the dilution of any subversive message. However, it does appear that its development into a written function has helped encourage the usage of *oh yeah* in other more generalized written functions that also resemble a speaker suddenly remembering something. For example, this explains the appearance of the *because oh yeah* construction that arrives in the 2010s. Furthermore, this highly conventionalized end of list function of *oh yeah* in journalistic writing is primarily the reason why *oh yeah* co-occurs so frequently with the conjunction *and*. In the non-dialogical uses of *oh yeah* in the NOW data, the collocations '*and*, *oh yeah'* or '*oh yeah, and*' was just shy of occurring in 50% of all the uses of *oh yeah* in journalistic prose.

In addition to these end of list markers, *oh yeah* has developed other sub-functions of the faux spontaneous additive category that includes what appear to be 'staged remembering markers'. Similar to the end of list markers, these are used by the author in a way that presents information he/she finds important as if it were nearly forgotten or an afterthought. These uses also frequently have the collocate *and* occurring immediately to the left or right of *oh yeah*. In the end, the 2010s saw a high level of innovation in terms of the faux spontaneous additives. Instead of only being composed of the end of list markers, as was the case before, the innovation of the convention of marking oneself suddenly remembering something takes the form of other constructions. The faux spontaneous additives also account for the vast majority of all uses of non-dialogical *oh yeah* at 70% to be precise.

Furthermore, there were other indications that the 2010s saw the expanding of the non-dialogical functions of *oh yeah*. This includes what I have termed the faux utterance launcher

functions. The biggest part of this category occurs in contexts in which the author poses a rhetorical question, and then continues with *oh yeah* as the response launcher that precedes the actual response. These faux utterance launchers are often phrased sarcastically in order to display the author's strong subjective attitude toward the content.

Finally, though constituting a small percentage of the overall usage of non-dialogical *oh yeah* in the NOW data, there were the intersubjective meta-textual comment functions. These were offset from the main text by parenthesis. The stylistic effect of these parenthetical comments is that they appear to be addressing the potential reactions or questions of the reader(s) individually in alignment-seeking functions. This is done in an attempt from the author to react to or acknowledge their readership's potential reaction. These meta-textual comments are arguably the most salient of all the non-dialogical uses because of being separated so explicitly from the main body of the text, not to mention their action of attempting to break the fourth wall.

Another important factor to consider is the appearance of *oh yeah* near other DM constructions involving the DM *well*. DMs occurring in sequences or in the near vicinity of each other have been a common observation in the past two case studies. For example, while *well* and *like* are used most often next to each other in DM sequences, *well* and *oh yeah* are found to occur near each other, but in separate yet similar functions, such as in the following example that was also discussed in chapter 3. This example includes pred *well* followed by the end of list marker of *oh yeah* in the very next sentence.

Why we can't wait to hear Eminem's latest

Encore (Nov. 16)

By EW Staff | September 18, 2004 at 04:00 AM EDT

Why we can't wait to hear Eminem's latest

He's given white rap street cred, powered a Hollywood flick, and turned a so-so crew of Detroit MCs (D12) into chart-topping stars. Now America's favorite speed-rapping, profanity-spewing, button-pushing Caucasian continues his incredible streak with a fourth solo album — and the product is under tighter security than the Bush twins. But according to sources in the Shady secret service, *Encore* will sound a lot like, well, Eminem: crude wit, intricate rhymes, and once again a disc produced by Marshall himself and Dr. Dre. Oh yeah, one more thing: It's gonna be huge.

Image 5.1: Extract from Entertainment Weekly magazine article of Eminem album 2004

When looking at these non-dialogical uses of *oh yeah* in journalism, it is also important to consider what type of articles that they come from. As already discussed in the analysis of the COHA data, there are definitely sub-genres of journalism where *oh yeah* is used more than others. Table 5.11 breaks down the types of articles in which *oh yeah* was used in journalistic prose. As can be seen, the top four categories make up for 79% of all the data. Furthermore, the top four categories are also those that have already been shown to contain more colloquial styles of language than other more formal forms of journalistic writing. However, even though the types of articles in which non-dialogical *oh yeah* is used is not surprising based on what was observed in chapters 3 and 4, it is surprising how much of a percentage these functions take up. This demonstrates that non-dialogical *oh yeah* is the most niched into the informal sub-genres of journalism than the other two investigated in this thesis. What is curious when looking at these top four categories is that *oh yeah* was not used the most in the comment sections, in which anyone can write down a reaction to an article much in the same way as the letter to the editor section of a newspapers or magazines. Instead it appears that *oh yeah* is still used the most in articles on entertainment and pop culture, which includes movie reviews, television show reviews, celebrity gossip and video game reviews. This preference for *oh yeah* to be used in articles involving entertainment, sports and columns/op-eds confirms the observations made

in the COHA data analysis. This is not a convention that is extended to the more formal types of journalistic articles. Furthermore, *oh yeah* was used essentially just as much in sports articles as they were in the comment sections. The findings here that confirm that *oh yeah* is also found primarily in articles pertaining to entertainment news, sports and columns/op-eds also suggest that the authors of these types of articles constitute a niched discourse community of specific authors that distinguish themselves from other journalists by using a more colloquial style of language.

Categories of Journalism	Frequency	%
Entertainment	213	35%
Comment Section	118	19%
Sports	115	19%
Columns/Op-eds	34	6%
Traditional News	32	5%
Lifestyle	26	4%
Business/Finance	22	4%
Tech/ and other Specialized Technology	22	4%
Culture/Restaurants	13	2%
Womanhood/Parenting	7	1%
Fashion	4	0.5%
Other	4	0.5%

Table 5.11: Types of articles in NOW data where non-dialogical *oh yeah* is used

The next characteristic of *oh yeah* that must be checked is if it has a clear preference to be used in web-only based news or traditional media. Table 5.12 below shows that *oh yeah* is used the most in web-only news sources, though not to the extent where it is overwhelmingly used more than traditional print and web sources. As with all other chapters, these two categories are trailed by examples coming from articles from television channels and radio stations that also feature written articles on their websites.

Source	Frequency	%
Web only	278	46%
Print & Web	235	38%
Television & Web	90	15%
Radio & Web	7	1%

Table 5.12: Journalism source distribution for non-dialogical *oh yeah* in NOW

The combined evidence of Tables 5.11 and 5.12 indicate that the non-dialogical uses of *oh yeah* are niched in the informal areas of journalism and that they are used the most in online-only news sources. However, though the preference for *oh yeah* to be used in entertainment news, sports news and columns/op-eds is much higher than other article types, its preference to come from online sources only is not. This co-usage of *oh yeah* in articles coming from traditional news sources and web-only ones is part of an overall trend that has been noticed with *well* and *like* as well, though to different extents. It is important to recall that the journalism sources of the COHA are all from print based sources (many of the COHA journalism sources would now fit into the print and web category of the NOW), and these sources were already using the written functions of all three of these DMs before these web only news sources were widely available. However, many of the web only sources are specialized in specific themes, such as entertainment, fashion or sports news, which may explain why web-only sources have consistently been the category with the highest percentage of usage in the NOW data.

The analysis of *oh yeah,* and also *well* and *like* in the previous chapters, has demonstrated that there have been some major developments in the usage of these DMs in (inter)subjective non-dialogical functions in American journalism, particularly from the 1990s until the present time. The fact that these changes start to occur around the time that the internet becomes mainstream is too much to overlook. Furthermore, the fact that the uses of non-dialogical *oh yeah* are more likely to occur in web-only based news sources also point to the language used online as a significant influence of change. However, the internet is not the sole influence of this change

in how *oh yeah* is used in writing. This issue will be addressed in detail in the following chapter in 6.6.

5.7 Conclusions

Through the analysis of data from the COHA and NOW, this chapter has demonstrated that *oh yeah* has gone through remarkable change in how it is used non-dialogically. Despite being a prototypically dialogical DM, *oh yeah* has progressively been used in written functions that are highly (inter)subjective in a way that resembles their information management functions; yet is specifically suited to the needs to journalistic writing. It has therefore been argued that *oh yeah*, like the other two DMs investigated in this work, is used to indicate their rhetorical stance toward use of a certain item or phrase in a way that imitates its oral function of marking the author's reflexive cognitive metalinguistic process of organizing information. Moreover, the use of this DM is useful in taking stance because it indexes casual discourse, which further allows the authors to seek alignment with the readership (cf. Silverstein 2003; Du Bois 2007; Jaffe 2009) In particular, it continues to highlight lexical items or the clause that follows. However, at the same time, it also carries the author's subjective stance or can mark the author's intersubjective attention to the readership (if not both). *Oh yeah* is used the most to mark rhetorically the last and most important or obvious lexical item in a list. This convention makes up the majority of the COHA data. It is also the most frequent function found in the NOW data accounting for roughly 36% of the total non-dialogical uses. Coincidentally, the end of list function of *oh yeah* has the most in common with the other word-choosing functions of *well* and *like*. *Oh yeah* is also used in a subjective function that resembles *oh yeah* as a response marker. However, in these cases, the authors are responding to their own rhetorical questions, not questions posed by another. Though much rarer than the subjective uses in both the COHA and NOW data, *oh yeah* can also be used intersubjectively in a way that engages with the audience meta-textually, often in parenthetical comments. This is similar to the meta-textual

comments that were observed in the *well* and *like* chapters. Furthermore, similar meta-textual comments have also been observed in social-media writing. If one considers that these uses are only attested in journalistic writing since the 1990s and that they appear most often in articles pertaining to certain informal article types, then it would seem very plausible that these intersubjective uses were partially influenced by the emergence of online language, which was starting to gain prominence in the 1990s.

One thing that separates *oh yeah* from *well* and *like* is the fact that it is often overlooked as a DM in its own right. However, the fact that the uniquely written functions of *oh yeah* are those which are used to express the author's subjective stance, or even to perform alignment-seeking comments, indicates that their usage in informal journalistic articles is similar to what has been observed with *well* and *like.* In these cases, the author rhetorically employs *oh yeah* in writing in order to discuss popular themes such as entertainment, sports and their personal opinions in columns/op-eds in a way that seeks alignment with the readership.

The emergence of these (inter)subjective reader-alignment seeking functions of *oh yeah*, along with the findings of the other two DMs in this thesis, have some significant implications for studying the colloquialization of written English, because they are more than just the increased usage of a colloquial element into writing. They are also the extension of a colloquial element of language into new contexts that are unique to writing. *Oh yeah*, like *well*, has even developed a highly conventionalized function (the end of list marker) that appeared seemingly out of nowhere in the 1990s, much like a slang term (cf. Adams 2009), and continues to be the single most used function today.

Furthermore, what we have in the case with *oh yeah*, is the utilization of a highly dialogical DM in a way that resembles how it is used orally. In writing, however, the author uses *oh yeah*

to mark his/her own attitudes toward the information presented in their own monolog texts. Furthermore, in some cases, the author even engages with their readership by attempting to anticipate how the readership might react to their texts. These developments mirror what we have seen with *well* and *like*. Like these other two DMs, we see the development of genre-specific functions (Aijmer 2013, 2015) that also coincidentally come about around the same time as the advent of internet based writing like e-mail in the 1990s (cf. Baron 2000; Crystal 2001; Barton & Lee 2013). Furthermore, *oh yeah* usage in these functions expands further in the 2000s, which is also the same time that newer forms of social media become more prominent in everyday life (*ibid*.). Therefore, this chapter has demonstrated that *oh yeah* has not only crossed into the informal sub-genres of journalism, but it has also developed genre-specific written functions that continue to push the development of *oh yeah* into being used to take up an intersubjective stance (cf. Du Bois 2007; Jaffe 2009).

Chapter 6:
Discussion of common observations from the three case studies

6.1 What do these three DMs in writing have in common?

The previous three chapters have discussed the individual diachronic changes of the DMs *well*, *like* and *oh yeah* in journalistic writing. They have also examined the developments of these DMs in functions that work as word-choosing markers in journalistic texts. These word-choosing markers resemble how these DMs are used in their dialogical textual functions of word-searching, self-corrections or information management, but instead encode their (inter)subjective stances toward the lexical items/information being highlighted. Furthermore, the 2010s also saw the emergence of meta-textual comments, which occur almost exclusively in parentheticals with the DM in question occurring immediately to the right of the first parenthesis. Like the written word-choosing functions that came before them, the use of these DMs in meta-textual comments specialize in highlighting lexical items, especially in the cases of *well* and *like*. In addition to this, these meta-textual comments are also among the most intersubjective of all the written functions. In particular, not only are these comments separated from the main text with the use of parenthesis, but they almost exclusively serve alignment-seeking functions, which share many characteristics with other constructions that are often used in online discourse.

This chapter addresses what these three DMs have in common in journalistic writing. Section 6.2 compares the results of the case studies on *well*, *like* and *oh yeah* and discusses some of the major observations that brings these three DMs together. Section 6.3 discusses the development of the written functions of the three DMs and traces from where the new functions have their origins. Section 6.3 compares the sequencing characteristics of *well*, *like* and *oh yeah* and expands on observations that were made in the previous three chapters. Section 6.5 compares the characteristics of these three DMs and gives an overview as to how this all relates to other phenomena of colloquialization. Section 6.6 addresses why these three DMs have all gone

through similar changes at approximately the same time (i.e. from the 1990s). Section 6.7 closes this chapter and 6.8 offers the final conclusions of this thesis.

6.2 Comparing the functionalities of *well, like* and *oh yeah* in journalistic prose

When we compare the three DMs *well*, *like* and *oh yeah*, it is clear that *well* is the DM that starts developing these functions the earliest. Furthermore, *well* is special in this list because the bulk of this change occurs in what has normally been its overlooked or less natural position of use, namely the sentence-medial position. It is also the DM that shows the most pronounced development and creativity in terms of expanding its functions into writing, and it is the DM that is used the most in these specifically written functions. *Like* occupies an intermediate position. *Oh yeah* is by far the least used, though its written functions appear to be amongst the most conventionalized of the three. In light of these observations, I argue that *well* is the DM that started developing specifically written functions first, in particular pred *well* which showed signs of becoming a truly written function in the 1980s. This practice of highlighting lexical choices appear to also be used with the DMs *like* and *oh yeah*, which are used to both highlight lexical items and express (inter)subjective stance.

Particular functions of *oh yeah* that have overlaps with pred *well* and WCM *well* include the faux spontaneous additive *oh yeah*. Moreover, the most common variant of the faux spontaneous additives, the end of list marker, has a near equivalent function with WCM *well* where the last item of a list is also preceded by *well*. In the end of list marker functions for *well* and *oh yeah*, not only is the final lexical item of a list highlighted by the use of a DM, but the author's subjective stance toward the item in question is also encoded into the DM. Examples (1) and (2) illustrate both *well* and *oh yeah* serving in these functions.

(1) At the end of Hothouse, Kachka describes the sleek new offices that FSG moved to in 2007. There's blonde wood and light, an open floor plan and, **well**, space—the usual trappings of your average contemporary workspace, many of which were missing from FSG's longtime tattered home above Union Square. (NOW: The Awl 2014)

(2) But it's a reminder that, while the Clinton presidency did see some bipartisan achievements (budget surpluses and welfare reform, for example), the context of these achievements is also important. It was a time of brutal political warfare, government shutdowns and, **oh yeah**, impeachment. (NOW: U.S. News & World Report 2014)

To say that *well* and *oh yeah* function in exactly the same way subjectively, however, is going too far. *Oh yeah* still maintains superficial links to its oral functions that indicate information management in such a manner as if something has suddenly been remembered. Naturally, the function that it serves in writing is very different: namely, it is used to highlight that the next item is important. The function of *well*, on the other hand, has a resemblance of marking the upcoming lexical item as a dispreferred in the sense that the usage of the noun *space* is too obvious to be worth stating, though the context given in the following clause explains why mentioning space is so important. If one were to replace *well* with *oh yeah* in (1), a similar meaning would still result, though the element of a faux remembrance would be stronger as opposed to marking a dispreferred element. If one were to replace *oh yeah* with *well* in (2), the result would be similar.

However, when we investigate other examples the differences in their subjective inferences becomes clearer. Furthermore, in many cases the replacement *well* by *oh yeah* or vice versa is not possible at all without significantly changing the meaning. For example, in examples (3) and (4) below, *oh yeah* cannot be replaced by *well* and still carry a similar meaning.

(3) The 23-year-old hit .330/.460/.649, with 42 home runs, over 654 plate appearances. He led baseball in runs scored, home runs, on-base percentage and slugging percentage en route to winning his first MVP award. **Oh yeah**, and he's only 23, did I mention that? (Yahoo Sports 2015)

(4) The 23-year-old hit .330/.460/.649, with 42 home runs, over 654 plate appearances. He led baseball in runs scored, home runs, on-base percentage and slugging percentage en route to winning his first MVP award. **Well**, and he's only 23, did I mention that?

For example, the implicature of "Well, and he's only 23…" is not as strong in conveying the message that the person's age is important when considering his accomplishments. A similar point could be made with example (5) that was discussed in chapter 5 with *oh yeah*, which is reproduced below. When we replace *oh yeah* with *well* the meaning of the sentence changes significantly. In (6) *well* implies more of a hedging sense to the inclusion of "a fire-stoking measure of outrage" that is not present when *oh yeah* is used.

(5) The NFL ultimately might remember the 2002 Raiders as champions, but for now, this is a crack team built on experience, poise and confidence. **Oh yeah**, and a fire-stoking measure of outrage (COHA Newspapers: San Francisco Chronicle 2003).

(6) The NFL ultimately might remember the 2002 Raiders as champions, but for now, this is a crack team built on experience, poise and confidence. **Well**, and a fire-stoking measure of outrage.

Even in this case in which the element being highlighted is a lexical element only as in (5), the replacement of *oh yeah* with *well*, as demonstrated in (6) shows that the implicatures of these two DMs, even in similar constructions, carry different meanings. The usage of *oh yeah* in these contexts specifically carries the author's subjective stance that the next element is the strongest element in the list that is given. *Well*, on the other hand, more often encodes a sort of hedging function in these cases in which the implication is that the author is highlighting the usage of this lexical item ironically in some sense. The result of replacing *oh yeah* with *well* in this case appears awkward and non-idiomatic.

There are also lexical highlighting functions of *like* that share overlaps with pred *well* and WCM *well.* This is particularly the case with the focuser hedge function in particular. Examples (7) and (8) highlight such similarities.

(7) They're just dead. And they'd like you to be dead too. That's, **like**, it. (NOW: NJ Pen 2019)

(8) You may have heard that idea expressed at, say, the regular old Ironman triathlon, which is normally considered the Mount Everest of the sport: a 2.4-mile swim, then a 112-mile bike ride, and then a marathon (26.2 miles) in the heat of Hawaii. If you could finish it, you were, **well**, an Ironman (or woman) (NOW: The Week Magazine 2016)

Example (7) above shares the exact same structure as pred *well*, an example of which is provided in (8). We have the copular verb *be* at the right side of the DM in question and the DM is followed by the subject predicate. Though there are some obvious overlaps in these two functions, there are also differences in terms of the implicatures of *like* and *well* in these two examples. The DMs *like* and *well* could arguably be switched in (7) and (8) without causing a significant difference in the meanings of the sentences overall. However, other characteristics appear to differentiate how these DMs are used even in similar functions. For example, in (8) *well* is used to highlight the noun *Ironman*. As was already discussed in chapter 3, *well* is often used in the pred *well* construction to not only highlight lexical items such as nouns or adjectives, but it was also used when the lexical item being highlighted was a term that was being repeated. Though there are some similar cases with *like,* such as (9) below, these cases were not nearly as prominent.

(9) Before you go off and watch it, I should say, Black Mirror is dark. **Like**, really dark. (NOW: The Verge 2014)

6.3 Comparing the developments of the written functions of *well*, *like* and *oh yeah*

Another factor that separates the new (inter)subjective written functions of *like* and *oh yeah* from *well* is how they enter into journalistic writing from the 1990s in the COHA data. With *well* there was a period of gradual development over the 20th century where the written functions of *well* appeared sporadically before they suddenly conventionalized at the end of the 1980s and 1990s. With *like* and *oh yeah*, this development is less evident in the COHA. For *well*, this

development is exemplified by (10) and (11) below. Example (10) exemplifies how *well* is used to mark a word search sequence between the copular verb and the subject predicate. Example (11) is an example of pred *well* as used in journalistic prose.

(10) You remember Modesty Blaisel And The Day the Fish Came Out? **Well**, now there's a raving new one called The Touchables, and it is absolutely not to be believed it's so-so, **well**, divine. (COHA: TIME 1968)

(11) Now, hopefully, he realizes that being boring is, **well**, boring, and goes back to what he does best: being awesome. (NOW: PopMatters 2010)

As argued in chapter 3, written pred *well* is used extremely sparsely until the 1990s when it suddenly becomes the most used sentence medial function in the journalistic COHA data. Due to the colloquial nature of *well*, it is clear that the textual use of *well* in which it is used to mark word-searching or self-correction is the ancestor of pred *well* (a similar point is made in regard to WCM *well*) and not the other way around. Why precisely authors started using *well* in the pred *well* (and WCM *well)* is not clear. However, based on these observations, it is apparent that pred and WCM *well* demonstrate a meaning shift of medial *well* from the textual speech functions to medial *well* in writing that indicates (inter)subjective and rhetorical stance. This shift follows Traugott's left to right shift from non/less subjective to (inter)subjective (cf. Traugott & Dasher 2002; Traugott 2010). What is unique about this shift, however, is that *well* has already gone through (inter)subjectification in oral usage. Traugott & Dasher (2002: 175-6) found that *well* was first anchored in the speech of others and then the narrator before it preempted to the speaker/writer's perspective, followed by developing meanings with strong orientation to the hearer/reader's face (cf. Traugott & Dasher 2002: 175-6).

Furthermore, as discussed in chapter 3, many of the interpersonal uses of *well* that index the speaker/author's perspective and or orientation to the hearer/reader's face occur utterance initially in speech, while the sentence medial functions of *well* in speech are generally serving

the textual (cohesion) functions. In writing, it appears that the use of *well* in the sentence medial positions have developed new rhetorical interpersonal meanings. This appears to be an independent process of (inter)subjectification that is specific to its developments in writing. In the examples discussed here, however, these DMs are used in journalistic writing in a way that encodes the author(s)' stance towards lexical items in ways that resemble the more salient ways in which these DMs are used in spontaneous speech. Similar observations were made in Rühlemann & Hilpert's (2017) study of *well* in magazine articles and Tottie's (2017) study on *uh*, *um* and *er* in journalistic prose. One possible solution for such developments could be that (inter)subjectification is not occurring as said by Narrog "in a social vacuum or (only) in an individual's cognition but also in speaker-hearer interaction in social contexts in specific communities and cultures" (2010:422).[58] For the purposes of this study, we are not talking about speaker-hearer interaction but instead the author-reader relationship. This path of (inter)subjectification is different from past looks at this process because the use of these DMs in writing, at least at the beginning, was clearly rhetorical. Furthermore, their sudden development in the 1990s and growth afterwards demonstrates that something particular has happened in journalistic writing since that time. This is different from what has been observed in many studies of (inter)subjectification, since the changes involved in these past studies often span centuries.

In the cases with *like* and *oh yeah*, a similar trajectory is likely to have taken place. However, as chapters 4 and 5 indicate, the processes by which these two DMs developed their written meanings appear to be much more sudden than was the case with *well.* Examples (12) and (13) show how *oh yeah* is used in transcribed speech (12) and in journalistic prose (13). In (12), *oh yeah* is used in the typical utterance launcher function. In this case, it both marks the speaker's

[58] A similar point is made in Traugott (2007: 296-297).

response to information presented by his/her collocutor as well as the author's remembrance of the person being discussed. In contrast, *oh yeah* is used in (13) rhetorically in journalistic prose. In this example, the author gives a list of activities that a certain reality television star has done since leaving the show. Then the author uses *oh yeah, one more thing* in order to present a final entry to the list, which is also the most gossip-worthy element, the fact that he is dating someone. Should *oh yeah* in (13) be the written outgrowth of uses in spontaneous speech as seen in (12), then it can be argued that the written functions of *oh yeah* have followed a similar trajectory as *well.* The journalistic prose functions resemble certain textual (coherence) based functions of *oh yeah,* particularly those involving information management such as remembering and/or responding to information present by others in the conversation.

(12) He said' **Oh yeah**! I liked him in "The Real McCoys". (COHA: Saturday Evening Post 1972)

(13) Since then, the burly runner-up has moved on. He's hired a personal trainer, begun a diet and started using Propecia to combat hair loss. **Oh yeah, one more thing**. "I'm dating someone now," he says. (COHA: People Magazine 2003)

A similar case could be argued to have occurred with *like*. This can be exemplified in the comparison of *like* as used in transcribed speech and journalistic prose. Example (14) shows *like* used as in a word-searching situation, while (15) shows *like* used in a focuser hedge function that induces the author's evaluation of a character in a television program.

(14) "It's a little unfair, people saying, How come she didn't stay 21? You know, **like** ... how come she didn't stay a virgin?" Still, Wedding producer Jerry Zucker says that, for Roberts, growing up in public has had a plus side. (COHA: People Magazine 1997)

(15) 12. Ahmad Maybe he's not the nicest guy around, what with his whole plot to kill Kublai and his whole family and whatnot. But, **like,** look: You'd probably be like that too, if you accidentally had sex with your mom. Honestly, his restraint should be applauded. (NOW: The Wrap 2016)

Overall, all three of these DMs demonstrate a shift in meaning when they are imported into journalistic prose. These changes are different from other meaning shifts of DMs due to the intentionality of their inclusion into certain journalistic article types. However, they also curiously follow the less subjective to (inter)subjective shift in change as they become more popular in journalistic prose. Moreover, once the written functions of these DMs, which specialized in (inter)subjective word choosing, became more established in writing and journalistic writing in particular, it led to further developments such as those that we observe with the meta-textual parenthetical uses in the 2010s. Another feature that links the functions together is the fact that they can occur in sequences or in near proximity to each other within the same texts. These issues will be discussed in more detail in the following section.

6.4 DM sequencing with *well*, *like* and *oh yeah*

DM sequencing is a common characteristic of DMs. DM sequencing is the co-occurrence of two-part and sometimes even longer sequences (e.g. *oh well* or '*you know*, *like*, *well*'). Though sequencing is the term adopted in this study, past approaches to this phenomenon have also called it 'DM clustering' (e.g. Maschler 1994; Aijmer 2004) and 'stacking' (e.g. Brizuela *et al.* 1999). The reasons why the phenomenon occurs is a matter of debate. Furthermore, Aijmer (2002, 2004) even makes the distinction between conventionalized collocations (such as *oh well*, *well you know*) and loosely co-occurring markers. This distinction is made using the criterion of reversibility, stating that the more conventionalized a particular combination is, the more fixed its internal order. When discussing the latter, she states that they have "a floor-holding function while the speaker is thinking of a new argument (2002: 31). When discussing what Aijmer (2002) calls DM 'collocations', or common DM combinations, the most widespread generalization in past studies of this phenomenon is that when DMs co-occur, the two or more DMs in question have related and/or functional overlaps. It is for this reason that Aijmer (2002: 31) states that *well* and *by the way* cannot co-occur while *oh* and *by the way* can.

In her argumentation for this, she references Murray (1979: 731), who argues that while *oh* can precede *by the way*, *well* cannot because *well* is inconsistent with the introduction of an unanticipated topic. A different approach to DM sequencing argues that DMs co-occur not because their functions overlap or are similar, but because two, or more, 'disparate' functions become simultaneously relevant at particular points in the discourse (Lohmann & Koops 2016: 432). In other words, the functions of the DMs being sequenced are seen as complementary[59].

Using corpus data, Lohmann & Koops (2016) found conflicting trends in terms of the functions of DM sequences. For example, despite having calculated an optimal sequencing hierarchy for 11 common DMs as can be seen in (10) below (see Koops & Lohmann 2015 for details of that calculation), they found that that not all combinations that were predicted by that hierarchy were frequently found in their corpus data.

(16) Optimal sequencing hierarchy for DM sequences formed for 11 common DMs
oh > well > and > so > or >but > because > then > you know > now > I mean

Furthermore, they also found sequences that violated the hierarchy, such as '*you know because*'. Finally, though the general finding in their directed graph based on the co-occurrence of data of the same 11 DMs (see figure 6.1) was that almost all sequences were unidirectional, there was one major exception to this rule, the DM *you know*.

[59] See Maschler (1994) for an investigation of Modern Hebrew DM sequences in conversation where she argues that the use of multiple DMs reflects that several transitions are executed at the same point.

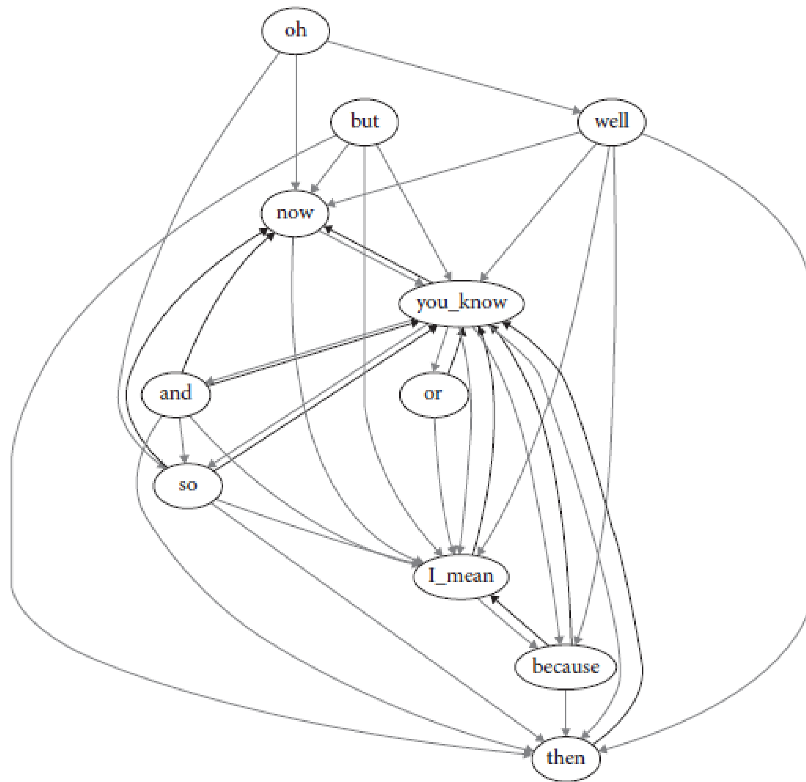
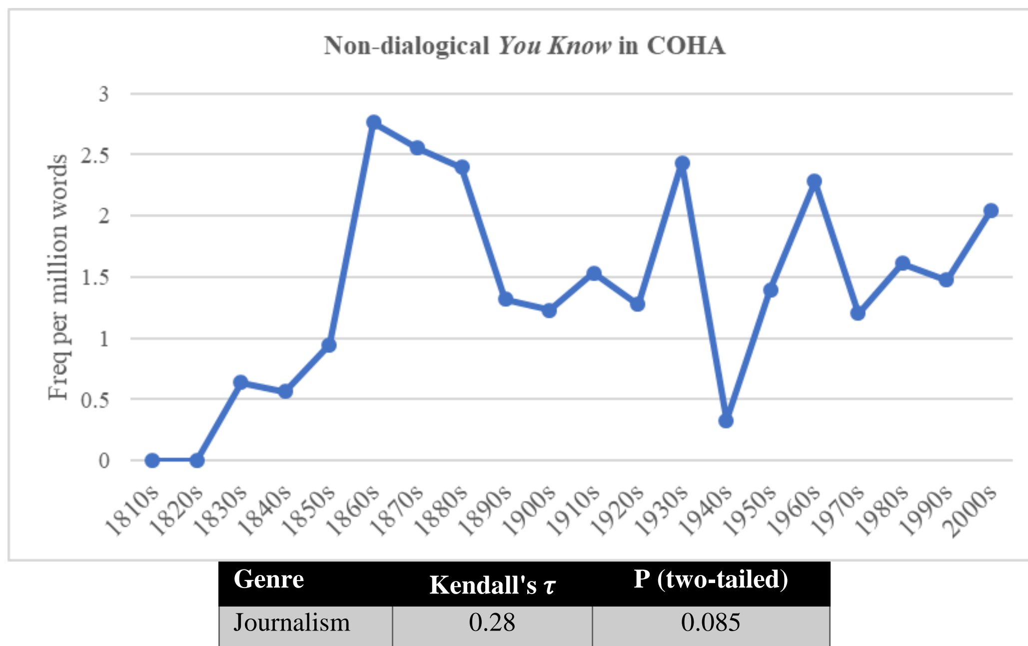


Figure 6.1: Directed graph based on co-occurrence data of 11 DMs (taken from Lohmann & Koops 2016: 439)

You know is unique in this list because it has many bidirectional connections, and it has what they call ‘extreme collocating behavior’ (Lohmann & Koops 2016: 440). Therefore, these “extreme collocating behaviors” of *you know* provide an explanation as to why this DM is a common part of DM sequences involving *well* and *like* in the written contexts. At the same time, it explains why it can be used in these sequences while not developing much along the lines of specific written functions as we have seen with the other three DMs investigated in this thesis. Furthermore, unlike the three DMs investigated in this work, there is little indication of *you know* being used increasingly in non-dialogical contexts (see Figure 6.2) nor of it developing journalism writing-specific functions.



Genre	Kendall's τ	P (two-tailed)
Journalism	0.28	0.085

Figure 6.2: Diachronic Frequencies of COHA journalism data for non-dialogical *you know*

Coming back to how *you know* is used in sequences with *well* and *like*, examples (11) - (13) demonstrate *you know* occurring in sequences with these two and sometimes other DM and interjections as well.

(17) WHICH LEAVES US WITH a final but very important question. When will the ESC promise pay off? When can we expect something more from them than arcane articles in medical journals (though repeatedly portrayed as miracle breakthroughs in the New York Times)? **You know**, **like**, **well**, **actually** making sick people better? (COHA Magazines: American Spectator: 2007)

(18) We'd always heard that the rooms at the Algonquin were, **uh**, **well**, **you know**, small, and Room 512 was no exception. (NOW: Town Topics 2013)

(19) But the fact he's having to run away from something (**like**, **you know**, a giant zombie army) is just A BIT worrying. (NOW: BuzzFeed News 2017)

In terms of the purposes of these DM sequences, they also resemble how they would be used in oral speech although they have been adapted to writing. In example (11), *you know*, *like* and *well*, are all used as faux utterance launchers (e.g. Biber *et al.* 1999: 1077) where the author

responds to his own self-posed question – marking the response first with *you know* and then following that up with *like* and *well*. Coincidentally, all three of these DM are working as hedges, which resemble the author downplaying a strong and negative stance that the author takes in the following sentence, namely that the only true measurement of success in doing a medical research is to make sick people better. Example (12) also includes more than two items in sequence, and it is reminiscent of a speaker using a hedging face-saving device that softens the blow of the criticism that he/she is about to make. In example (12), it is the assertion that the rooms in a certain New York City hotel are small. In example (13), which is a sequence of two DMs only, *like* and *you know*, the DM sequence also resembles a hedging strategy of down-playing the intensity of what the character in the television program is running away from – in this case a giant zombie army. *Like* and *you know* are also accompanied by an additional downplaying hedging strategy, where the author adds that having to run away from a giant zombie army is "just a BIT worrying". Though all three of these examples are clearly highly subjective, the oral functions that they resemble could also have the secondary function of being floor-holding continuation signal devices, which speakers often use while searching for the right terminology or when performing self-correction etc. However, in the case of writing, what they are actually doing is marking the stance that the terminology that they are using is important and that the readership should take note of it. The usage of these DMs in written contexts with evolved written functions has already received much attention in the previous three chapters. The argument made here is that this convention carries on in the same way when DM sequences are used in writing as well.

The remainder of this discussion on DM sequences will analyze other combinations that are common with the three DMs that have been the focus of this study. Though the sequence hierarchy and a two-dimensional graph of the nature of DM sequencing from Lohmann & Koops (2016) provide rich insights about the restrictions and constraints of DM usage in

English, they admit themselves that their findings do not resolve the motivations behind DM sequencing. For example, they state that though their findings point to a certain 'grammar' of DM sequencing that speakers adhere to, "(s)o far it is not clear what the scope of these motivations is, or whether there is a division of labor between them" (2016: 441). This lack of clarity is apparent in their data, because "the empirical results indicate a number of different functional motivations underlying both the co-occurrence of DMs and their ordering" (*ibid.*)

The factor that separates the DM sequences in this study from all of the others mentioned so far is that the DM sequences investigated here come from writing and not from spontaneous speech. Therefore, as has been the case for the entirety of this study, the analysis of the journalistic data must be analyzed through a double lens. For example, it must balance how these uses of colloquial DMs are similar to and resemble their usages in oral speech, and how they are used and in ways that are particular to writing.

Unsurprisingly, DM sequences were more numerous in the dialogical areas of the COHA and NOW data, meaning that they were more numerous in the interview and oral transcription parts of the news articles. Overall, most of the DM sequences found in this investigation include typical combinations that have been well attested in other investigations of DM sequences. With *well*, the most typical DM sequences in which it occurred were from the quote *well* function. This includes '*oh*, *well*' (see example 20), '*well*, *well,* (*well)*' and '*yeah well*'. This was found in both real quotes and in fabricated ones.

(20) He realizes, "**Oh**, **well**, looks like this body's doomed." (NOW: IGN 2013)

‘*Like*, *well’*, ‘*like*, *oh*’ and ‘*like*, *yeah*’ also often occur in sequences together in the quotative *like* function, which, as in example (20) above, is often used in quotations both real and fabricated (see example 21).

(21) “When I heard that was the reason, **like, ‘Oh**, he’s just an old man and he doesn’t know any better and he’s just being harmless and playful and it’s just where his arm falls … I just burst into uncontrollable sobbing,” Corrigan said. (NOW: New York Post 2017)

The only DM sequencing that occurs with *oh yeah* (as opposed to the fact that it is a fused two-word DM) comes from the dialogical portion of the data. There are no examples of *oh yeah* occurring with another DM in any of the non-dialogical data sets. The closest we see to *oh yeah* occurring in a sequence is with *and*, but in these examples it is clearly acting as a conjunction as it most often occurs at the end of a list of items. Example (22) provides an illustration from a radio-show transcription, where it co-occurs with both *well* and *you know*.

(22) PLACE: **Oh yeah**, **well, you know**, it's funny, Dylan, you know, he freely admits to being a Woody Guthrie jukebox when he first came to New York City. (NOW: NPR 2012)

When we restrict our attention to examples from the non-dialogical data only, however, the occurrences of DM sequences become noticeably low. Table 5.1 gives the raw numbers and percentage for how many DM sequences (for the purposes of this study DM sequences can also involve interjections as well) occur in the non-dialogical NOW data for each DM.

DM	Sequences	% of non-dialogical
well	34	5%
like	22	10%
oh yeah	0	0

Table 6.1 DM sequences for non-dialogical *well*, *like* and *oh yeah* in NOW

As Table 6.1 shows, the presence of DM sequences among the non-dialogical functions is relatively low, and even non-existent with *oh yeah*. When these DMs are used in sequences, however, they are more salient than when they are used on their own. In example (23) below, which involves both *like* and *well*, these two DMs are used to both artificially inject a sense of spontaneity to the text while also serving as a way to mark the author's own subjective stance.

(23) We'll say this about Gilbert Arenas's Twitter account, which was taken down some time today: it wasn't boring, like LeBron James's worthless feed, and it wasn't used for incessant self-promotion, **like**, **well**, every pro athlete ever. (NOW: Deadspin 2011)

On the surface, this example appears to work with the assumption that DM sequences work together or have two functions that work in a complementary fashion. For example, in this case, *like* is functioning here in its exemplifying function while *well* hedges the author's strong stance on the Twitter feeds of all pro athletes. However, it is also worth mentioning that both *like* and *well* are also used in a variety of similar functions involving word-searching and hedging, which has also been discussed at length in this thesis. Therefore, it is also possible to analyze this DM sequence from the perspective that they co-occur because of the overlaps that exist in their overall functions.[60] Furthermore, it is important to keep in mind that these DMs are not entirely serving the same functions as they do in oral speech, since their usage in writing, in and of itself, is used as a rhetorical strategy to express their (inter)subjective stance while also making their texts more engaging and interpersonal.

A more obvious example where these two DMs work together to serve essentially the same function is found in (24), which was already presented in the chapter on *like*.

[60] See Hutchinson (2003)'s analysis of formal written DM sequences where they have a tendency to co-occur with related markers and can thus be exploited for the automatic classification of DMs into larger classes.

(24) Forget the smirking 1960s comic-book character and the blissful family-of-women '90s sitcom starring Melissa Joan Hart. In 2018, teens have nothing to smile about, so Kiernan Shipka's sorceress lives in two traumatizing worlds. At regular high school, one friend gets bullied; another protests a book-banning. Now, it's not all bad — her boyfriend loves comic books; the best baes always will — but Sabrina double-lifes in the Church of Night, a coven that worships, **um**, **like**, **well**, the Devil. The hoof-horned Dark Lord demands that young witches sign their name in his book on their 16th birthday. It's a gender-abasing Faustian bargain. He gives her power if she gives him everything. (NOW: Entertainment Weekly 2018)

In this example, both *well* and *like* are used in ways that resemble how both of these DMs are used to hedge potentially offensive items (such as the Devil). In the specific case of this article, however, they are used to a humorous effect. In this case, they are used to intersubjectively engage with the readership over their potential surprise that a popular fictional character now worships Satan in the reboot of the television series.

Similar to what was observed in Lohmann & Koops (2016), examples (17) and (18) show that there can be multiple functional motivations for DM sequences. However, what is most important in this analysis of dialogical DM usage in journalistic texts is not why these DMs are combined in this way, but that these uses are imitating oral practices of using DM sequences. Therefore, it will be important to monitor how DM sequences are used in journalistic writing in future studies. Will such sequencing usage increase, or will they continue to only sparingly be used in journalistic writing?

6.5 DMs in relation to other colloquial phenomena

Punctuation surrounding the use of dialogical DMs is another aspect that warrants discussion. Past works such as Rühlemann & Hilpert discuss how DMs and inserts are offset from the "neighboring co-text by topographical means" (2017: 109; see also Tottie 2017). The need to offset DMs from the neighboring text with the aid of punctuation is undoubtedly due to their syntactical independence from the sentence structure (such their tendency to occur sentence initially, finally or between clausal/syntactic borders) as well as the prosodic features that

surround DMs such as pauses and stress in spontaneous speech (cf. Aijmer 2013: 27-8). In other words, when these dialogical DMs are used in writing, the authors are also imitating the way in which DMs are used in speech by means of punctuation. Therefore, when pred *well* is used sentence medially in journalistic prose, for example, it is offset from the surrounding text by commas in order to simulate the pauses that would surround the word-searching *well* function that pred *well* resembles (compare example 25 from journalistic prose with 26 from transcribed speech). Therefore, the utilization of commas to offset DMs from the surrounding text is also adding to the overall colloquial-like writing style for the articles in which these dialogical DMs are used.

(25) The wildlings tasted, **well**, wilder. (NOW: The Atlantic 2010)

(26) "Determining that if there is any minimal threat whatsoever we're simply going to hold them forever is, **well,** quite frankly, un-American.[…]" (NOW: Huffington Post 2013)

Besides being surrounded by commas, the NOW data also displayed a more recent phenomenon where the three DMs under investigation were offset from the neighboring text by means of round parenthesis (see 27). Coincidentally, Leech *et al*. (2009: 246) noted an immense increase of round and square brackets in their corpus data that compared data from the 1960s and 1990s. Though Leech *et al.* 2009: 246) state that brackets which surround parenthetical comments are "considered more typical of serious written style", the examples discussed in this thesis would certainly not fit such a description. Instead, the uses of parenthesis in these cases is further offsetting these DMs and the comments that follow from the main text.

(27) I won't talk about dreams anymore. (**Well,** maybe briefly.)(NOW: Kotaku 2011)

As addressed in chapters 3-5, these uses of parenthesis to offset the DM and the proceeding comment differs from what was observed with the other cases in that they serve a more explicit

alignment seeking function than seen with the other uses that are offset by commas. In these cases, the DM is not only offset from the syntactic structure of the surrounding text in a way that imitates word searches or self-corrections, but it is separated from the main text as a whole and it imitates the author addressing the readership directly. Therefore, when the authors use dialogical DMs in these parenthetical contexts, they are not only taking advantage of the conversational flair that these DMs represent, but they are further indexing (through the use of common positive politeness strategies[61]) their self-identification as being 'close' to their readership. This is done through the use of common spontaneous speech items that are characteristic of discourse between close peers (See discussion of the third level of indexicality in Silverstein 2003).

On a global level, the uses of *well*, *like* and *oh yeah* in journalistic prose give the articles in which they appear a more casual, close and personal style as opposed to one that is distant and impersonal. When discussing colloquialization, Leech *et al.* (2009: 247) state that of the various factors[62] in which writing is becoming more oral-like, an informal style that includes interpersonal linguistic elements "is the one that is most centrally involved in colloquialization". In their discussion of how an informal style is facilitative to the inclusion of colloquial elements in writing, Leech *et al.* (2009: 247-8) give the following example in which the first and second person pronouns, *I* and *you* are utilized by the author to set up "a personal rapport with the reader":

(28) Don't worry. Editors don't want school compositions or essays. All you need do is tell it like it is, write as though you were talking to your neighbour over the garden fence. The more simply you write, the better it is. [. . .] The basic rule is to write concisely and simply. Don't waste words; don't use a long word where a short one will do and don't try to be "literary". Write the way you talk. Use varied, but fairly short, sentence-lengths: and keep the paragraphs short – maybe 50–60 words, three or four sentences only.

---

[61] See Brown & Levinson (1987) for a discussion of politeness theory

[62] The other factors listed are its avoidance of elaboration or specification of content; its shared addresser – addressee context and its interactive nature.

Other ways in which (28) displays its informal colloquialness is, of course, the use of the negation contraction *don't*. In observing how the three DMs in this study are used in journalistic prose, it is clear that this is yet another way to inject colloquial elements into texts which in turn also give them a more interpersonal and friendly rapport.

However, the use of dialogical DMs in journalistic prose is also different from what is seen in (28), because the colloquial elements that are used, pronouns and negation contractions, are not functioning in ways that are specific to the written medium. On the other hand, the usage of the three DMs in written prose is different from how they are used in speech. The fact that they have already been proven to develop specialized text-type functions in part explains why these DMs were able to adapt and be used in journalistic prose in special ways. Furthermore, the fact that they were able to be used in a variety of functions before transferring into journalistic prose may also be a factor that explains their development in journalistic texts.

For future study, it would be indispensable to investigate how these three and other dialogical DMs were used in a variety of what Barton & Lee (2013) call 'vernacular literacies' in the decades leading to the 1990s. Vernacular literacies include all forms of writing that are voluntary and self-generated (Barton & Lee 2013: 138). These casual forms of writing, such as personal journals, letters, e-mail and blogs were not covered in this thesis, which narrowed its focus to published texts. However, it is plausible that it was in these genres of writing that these DMs first acquired their written functions before they appeared seemingly out of nowhere in journalistic texts in the 1990s. Moreover, the increasing importance and availability of new vernacular literacies online (e-mail, blogging, microblogging etc.) could also be a factor that is further influencing the continued colloquialization of published texts in journalism and beyond.

Another question that must be addressed is how these uses of dialogical DMs relate to other metadiscursive comments such as general parentheticals or 'scare quotes'. It has already been discussed how brackets were found to be significantly on the rise in British and American English at the end of the 20th century (Leech *et al*. 2009: 246). Moreover, commas, which can also serve to offset parentheticals, were found to be stable across the second half of the 20th century (*ibid*.). As discussed in chapters 3-5, the most typical articles in which the written functions of *well*, *like* and *oh yeah* appeared were those involving entertainment news, sports news and columns. Example (29) from an entertainment news article demonstrates some of the other types of parentheticals that appear in this contemporary journalistic writing.

(29) It all could have worked better if the past two seasons had felt less like headlong rushes toward predetermined outcomes, at the expense of character and story believability. **(Whatever that means in a dragon epic.)** I might have even accepted King Bran the Broken and his "everything happens for a reason" rhetoric if the show had **just … nah**, **actually**, I probably **wouldn't** have. But so many of the things that drove fans loudly crazy this season most likely **wouldn't** have if **they'd** been given more room to breathe. **(More on this in a minute.)**[63]

The two parenthetical comments in (29), like the parenthetical comments that are introduced by DMs, serve rhetorically intersubjective functions. While also utilizing what appear to be spontaneous additions to their writing[64], these parentheticals index the writer's attention and alignment to the readership whom the author includes in the analysis of the finale of the television program, 'Game of Thrones'. In addition to these two parenthetical comments, this paragraph also includes other colloquial features such as verb and negation contractions, and the sequence *just ... nah, actually* which is used rhetorically to negate that author's acceptance of the plot twist in which the character Bran becomes king. Overall, the language employed in this article excerpt, which appeared in The New York Times (a high prestige newspaper), is

[63] This article comes from the New York times and can be consulted in its entirety via the link below: <https://www.nytimes.com/2019/05/19/arts/television/game-of-thrones-series-finale-recap.html>
[64] For example, note the use of "more of this in a minute" which is a conversational phrase used in speech where the length of conversation is often calculated by time instead of pages.

highly informal in the sense that the author uses many casual language features. This includes colloquial elements that have been addressed in past colloquialization studies (contractions, increased use of brackets) as well as others that have not as of yet received much attention (such as the sequence *just ... nah, actually* that imitate the spontaneity of spontaneous speech in which pauses and reformulations occur).

In another excerpt from the same article, the author also uses a scare quote. Scare quotes involve the use of quotation markers around a particular word or phrase that the author wishes to use ironically. Scare quotes can also be used because the author wishes to disassociate him/herself from the word or phrase being quoted. In (30) the author uses a scare quote in order to note the irony of the use of the word 'liberating' by the character Daenerys Targaryen when discussing a group of people who she had just butchered.

(30) The mix of highborn and low was meaningful, and combined with the depictions of the Targaryen Regime — the Nurembergish rally, the tyrannical doublespeak about "liberating" people who had just been butchered — it unsubtly hammered home the show's main themes: Power corrupts. Working together is our only hope.

This practice also exists in spontaneous speech where one signifies the quotation marks either with their hands or by saying, "quote/end quote" before and after the word or phrase that one wishes to offset. Scare quotes are similar to the uses of *well*, *like* and *oh yeah* in that they are used to highlight a particular word/phrase as well as indicate their affective stance towards its usage. While scare quotes typically indicate the author's desire to distance his/herself from the usage of the word or phrase, there is also a strong intersubjective element to their usage. For example, the author positions not only his/her stance toward to use of the term in question, but also attempts to get the reader to align themselves with the same stance (cf. Du Bois 2007).

The changes involving the DMs *well*, *like* and *oh yeah* fit in with the usage of other colloquial elements of language in writing. Their usage gives the texts the feeling of the spontaneity of spontaneous speech. However, what they actually do is index casualness and closeness for the purposes of marking their subjective stance in a fashion that also attempts to get the reader to align with him/her. When all of these elements are used together in written texts, the overall impression is that they are much more colloquial than other forms of writing, such as academic articles or textbooks. However, it would be incorrect to assume that these articles read like transcribed speech. Instead, what is occurring is that certain salient linguistic features that seem oral-like are increasingly used in writing for certain rhetorical purposes that make the written text feel more close, personal and easy to process. One more open question that remains to be answered is whether these new written uses of the three DMs in question are now being transferred to spontaneous speech. This is a phenomenon that would be worth investigating in future studies.

6.6 Why is there a surge of these DM in written functions starting in the 1990s?

Another important question to approach however, is why these three DMs developed specifically written functions at precisely the same time? One interesting coincidence about the development of these written functions is the fact that they all gain ground around the 1990s, which is also the same time of the mass adoption of e-mail and other forms of online-based language. When discussing the language of the internet, many regard online writing as much more interpersonal and interactive than traditional writing, especially in domains of social media (e.g. Zappavigna 2012). However, despite claims that online writing is essentially 'written speech' (e.g. Elmer-Dewitt 1994), Crystal (2001: 47) stresses that there are constraints that computer mediated communication puts on users. This includes a lack of being able to use prosody, facial-expressions, gestures and conventions of body posture. All of these limitations make the literal interpretation of online language as 'written speech' problematic. Instead of

comparing online language, or Netspeak as Crystal calls it, directly to that of oral speech, Crystal states that "Netspeak is better seen as written language which has been pulled some way in the direction of speech than as spoken language which has been written down" (2001: 47). Furthermore, Crystal also goes even farther by stating that Netspeak is, "something fundamentally different from both writing and speech, as traditionally understood" (2001: 238) In a similar vein, Baron (2000: 258-9) argues that e-mail should be regarded as a mixed contact system that currently has a hybrid character that is part speech and part writing.

The evidence from this study would instead suggest an alternative way of looking at what has become internet writing. In particular, though many of the features that could be identified as being speech-like in journalistic writing come about at the same time as internet writing began to become popular, it is also certainly the case that these functions were not entering into journalistic texts only after these functions became popular online. Instead, these DMs appear to have developed journalistic written functions after an initial wave of change towards a more colloquial written style already occurred (cf. Leech *et al.* 2009). This would therefore suggest that the spread of these colloquial elements of language is not a creation of the internet, but rather a continuation of the spread of increasingly colloquial practices of writing that are mostly used in informal sub-genres of writing. For example Barton & Lee (2013) argue that internet based 'vernacular literacies', particularly from sources such as Facebook, Flickr and instant messaging, has helped further spread the conventions and circulate informal and more colloquial-like writing conventions (Barton & Lee 2013: 138). Therefore, internet language is likely an encourager of these changes, but not the reason for them. Similar findings have been observed with *uh*, *um* and *er* in journalistic texts, particularly in the sentence medial position, from the 1960s, but particularly from the 1990s in Tottie (2017). Tottie points out that though this particular function of word-searching is not a very frequent one in spoken language, it is a salient one (2017: 124). This has many parallels with how pred *well* and WCM *well* developed,

which have been argued to be the functions that have opened the door for the other similar word highlighting functions of *like* and *oh yeah*.

Furthermore, what is occurring in online platforms and in informal journalistic texts is not the same. For example, the language styles used by regular social-media or comment-board users come from unpublished texts with considerably fewer restrictions or expectations put on what they write about. Journalistic texts, even those from the informal genres of journalism, have editors and guidelines that must generally be followed. However, the analysis of these new functions also shows that the new functions of these three DMs are not used freely across all sub-genres of journalism. Instead, they appear to be gaining ground in the more informal areas of journalism such as entertainment, sports and column/opinion articles. As the themes touched upon in these articles tend to aim more at entertaining or engaging with their readers than other more formal sub-genres of journalism, it should not be surprising that the authors choose to use language that is more colloquial and interpersonal-like. Though it has been established that the language used in social media and journalism are very different from each other, it is also plausible that this constant exposure to online language, which Crystal (2001: 48) identifies as having “speech + writing + electronically mediated properties”, is encouraging many people, journalistic authors included, to increasingly use oral-like items of language in their writing. This includes the use of dialogical DMs in ways that resemble orality yet do so in a way the is specifically suited for writing. This is all also consistent with Aijmer’s (2013, 2015) analysis on how DMs do not have a fixed meaning but a ‘meaning potential’ that allows them to adapt and develop specialized functions for different conversational and written text types.

In this investigation of the examples of the written functions of the three DMs investigated in this thesis from the COHA and NOW data, it appears likely that certain journalism sub-genres, especially those involving entertainment, sports, and op-eds/columns, are more linked to the

online social sphere than the more formal areas of journalism. This would not be the first time that online communication has been a facilitator and/or the breeding ground for the spread of vernacular words and/or constructions already underway in speech. One such case that has already been brought up is the 'sarcastic much?' construction. Though the 'sarcastic much?' construction was originally a dialogical construction that expressed the speaker's critical attitude towards an "antecedent" (i.e. an element in the context) it has been found to have gone through changes in terms of intersubjectification in online writing that have extended the construction from confrontational uses only, which have been documented in the literature (Adams 2014, Gutzmann & Henderson 2019), to alignment-seeking functions that occur in parenthetical comments in online writing, some of which have no antecedent at all (see Hilpert & Bourgeois 2020).

Another case involves the 'because x' construction that encompasses 'because + Adj', 'because + N' and 'because + particle'. Bergs (2018) demonstrates in his article on the 'because x' construction that despite the fact that this is not necessarily a recent construction by any means, "(p)resent-day English seems to have seen a spread of the micro-construction(s) and the development of the meso-construction, partly because of increased type frequency, but also because of prominent and popular singular constructs such as *because science* or *because reasons* that serve as attractors in a kind of snowcloning effect", which were made prominent especially due to their appearances online (2018: 57).

What both of these 'sarcastic much!' and 'because x' constructions show is that linguistic changes that occur with the help of exposure in online communication need not originate in online contexts, nor be confined to them. Furthermore, the analysis of these two constructions also shows that once certain 'popular singular constructs' are frequently used and spread to the mainstream via internet usage, which is increasingly an important fashion for new

constructions, slang terms etc. to spread, they can serve as influencers/attracters for snowcloning effects for new constructions (Bergs 2018). Furthermore, it can lead to the generalization of a given construction's meaning which allows it to take on (inter)subjective stances. This was seen, for example, with the 'sarcastic much?' construction. When it comes to DMs used in written functions, it appears that the salient and fixed written functions of the DMs such as pred *well*, the end of list function of *oh yeah* and to a lesser extent the focuser hedge function of *like* began as stylized constructions that mimicked orality that then became fixed and conventionalized written constructions. Once these constructions became more widely accepted in the 2000s and 2010s, they then opened the door for new similar yet unique conventions involving these DMs to come about such as the meta-textual comments. These new functions are able to be formed because DMs have a number of easily identifiable (or conventionalized) meanings that are also able to be "adjusted to the demands of a new situation" (Aijmer 2015:202). Therefore, similarly to what Traugott (2018: 43) says when discussing the expansion of the functions of *after all*, the new functions of DMs "are conceptualized as the outcome of speakers partially matching the expressions to extant abstract constructional schemas and subschemas". These processes have clearly been discussed most when dealing with changes in oral speech. However, similar processes could be at play with the encouragement of what is occurring in certain written genres, particularly those that are informal. Therefore, the changes in how DMs are used in journalistic writing are influenced by many different factors. This includes, of course, the changes that have already been observed in terms of colloquialization. This shift in the writing style of journalism is further pushed at the end of the twentieth century thanks to the vernacular characteristics of the online writing style that are increasingly popular and interconnected in certain informal areas of journalism, like, for instance, entertainment news.

The conclusions drawn above differ from the ideas presented by Baron (2000) and Crystal (2001) who stress that Netspeak is something different from the written language that came before it. The analysis above argues that digital discourse should not be seen as the main cause of the changes in writing from the 1990s, but it should instead be seen as an additional facilitator of the process of colloquialization that has already been proven to be underway. Furthermore, the analysis here sides more with the notion that the language of the internet should be seen as a recent expansion[65] of vernacular literacies (cf. Barton & Lee 2013). While vernacular literacies were once confined to small readerships (sometimes the author was also the only reader as in the case of private journals), the writing style of social media is readily available to wide audiences. Moreover, over the past couple decades that have seen this explosion of social media platforms, the language used across these platforms remains influenced by the writing styles (genres) that have come before them. For example, Thurlow & Mroczek (2011: xxviii) state that "*Facebook* profiles bear a strong formal and functional resemblance to personal home pages, and that interactivity, user comments, and online collaboration existed *before* the so-called Web 2.0". In the case of the development of these new functions of the DMs *well*, *like* and *oh yeah* in journalistic writing, their vernacular-like functions could have very likely benefitted from the arrival and expanding popularity of digital discourse. For instance, one reason why entertainment news was always the top category for the three DMs in question could likely be that this area of news is increasingly interconnected with social media and the language of online platforms (see also discussion in Rühlemann & Hilpert 2017: 131-2).

[65] See also Jones, Schieffeln & Smith 2011 for a discussion on how people (groups) can engage in "adapting new communicative resources for particular cultural ends" in which the new media practices develop out of or in concert with previous media practices.

6.7 Discussion

This chapter has demonstrated that the reasons for the change that has been observed with *well*, *like* and *oh yeah* are complex. For example, the influencers of this transition include factors that involve changes to the writing style in journalism in general, the developments of online writing as well as factors that have to do with the development of DMs specifically. However, when comparing the results of the three case studies together, it was determined that *well* was the first to begin showing signs of developing written specific functions. When these DMs are used in these new functions, they specialize in highlighting word choice in a way that resembles the oral textual functions of searching for the right term, performing a self-correction, or managing information. However, when these DMs are used in this manner in writing, they perform different stance-taking functions. In particular, they encode the author's (inter)subjective stance toward the lexical item in question.

Moreover, this chapter has also shown that in addition to these three DMs developing specific written functions, *well* and *like* are also capable of being used together or with other DMs/interjections in DM sequences in non-dialogical contexts. One common DM that occurs in DM sequences with *well* and *like* is *you know*, which ironically has not enjoyed similar developments in journalistic writing as the other three have. However, such uses of DM sequences are relatively rare in the non-dialogical data, and they are confined to the DMs *well* and *like*.

In terms of why these DMs developed written functions when they did, there are a number of factors to consider. Along with the processes of colloquialization, which became significant since the second half of the 20th century, these developments were also likely helped along by the arrival of digital discourse, which has increased in importance over the past few decades in

terms of the amount of users as well as the explosion of digital platforms that are utilized by the public (cf. Thurlow & Mroczek 2011).

In future studies, it will be necessary to continue to monitor the three DMs investigated in this thesis. Furthermore, it will also be essential to see if these three DMs continue to expand their functionalities in journalistic writing as we have seen over the past few decades. It will also be necessary to monitor the situation to see if other dialogical DMs begin to show signs of developing written specific functions, which could very likely be the case since it has already been demonstrated that various DMs are capable of expanding their functions to suit the needs of different text types.

6.8 Final conclusions

This thesis has shown which dialogical DMs are progressively used in journalistic texts in specialized written functions. Furthermore, it has also has also brought to light some of the ways in which DMs are able to develop text-type (genre) specific functions that retain some of the essential key formal features of the prototypical spontaneous speech functions that they emulate. These changes to the three DMs indicate, as already stated by Aijmer (2013: 27), that DMs are flexible and "have the capacity to be used in new senses in new linguistic contexts and situations". Furthermore, this thesis has also discovered that through this process of entering into new contexts, certain textual functions of DMs can become further (inter)subjectified while the original textual functions that these uses resemble become weakened. Coincidentally, this weakening of the textual functions is what allows these DMs to take on meanings that are relevant to the written medium, particularly when it comes to indicating one's stance toward particular lexical items within a text.

In terms of the contribution that this study makes to the study of colloquialization, this thesis has demonstrated how dialogical DMs are used in ways that appear like characteristics of colloquial spontaneous speech, but are in fact used in a way that is specific to the constraints of the written medium. The colloquial-like effect, however, is accentuated by the use of punctuation that accompany the use of these dialogical DMs in journalistic articles. The use of punctuation around the DMs t offsets them from the neighboring text. The usage of these DMs in certain types of articles indicates that not all areas of journalism are becoming more colloquial in the exact same way(s) nor to the same extent. For example, in spontaneous speech the three DMs investigated in this thesis are used most often in informal and friendly discourse (cf. Jucker & Smith 1998; Dailey-O'Cain 2000 Aijmer 2013: 27). Coincidentally, these three DMs were found in article types that contain many other indicators of an informal style such as parentheticals, contractions, the use of other interjections and DMs (in sequences or otherwise), person pronouns and scare quotes. The article types that most often included these features were those discussion entertainment news, sports and columns/op-eds. Moreover, the fact that the DMs *well*, *like* and *oh yeah* were found the most in these areas of journalism also suggests that the authors are profiting from the use of these dialogical DMs in order to be rhetorically anti-formal. The usage of these DMs and many of the other colloquial features discussed in 6.5 also allow the authors of these articles to distinguish themselves from more formal journalists as a specific discourse community within the larger genre of journalism (cf. Watts 1999; Swales 2016).

Another factor that has been discussed in detail in this study is the timing of these developments, which take off from the 1990s and continue to this day. Though past studies of colloquialization (e.g. Mair 2006; Leech et al. 2009) have discussed in length the changes to more colloquial writing style in various written genres, including journalism, since the second half of the 20th century, few have tackled shifts in these same genres since the drastic changes that have

occurred since the 1990s. It is important to track changes that have occurred during this later period of time because of the significant digitalization that has occurred since. Though the stance taken in this thesis is that the changes were already underway before digital discourse became mainstream, these same changes most likely accelerated thanks to the rapid development and expansion of various digital discourses that also use a writing style that is highly informal and colloquial.

For future studies, it will be important to see how *well*, *like* and *oh yeah* continue to evolve in journalistic texts in the decades to come. Furthermore, whether other dialogical DMs will follow these three is also an issue that requires future attention. In terms of the future study of colloquialization in general, it will be essential to observe what new colloquial elements find their way into journalistic writing, particularly those pertaining to entertainment news, sports news and columns. Another important aspect of colloquialization that requires more attention in future studies is the threshold of acceptability when it comes to colloquial elements that are used in written texts. For example, Leech *et al.* (2009: 248) observed how some elements of language, such as the semi-modals (*have*) *got to*/*gotta* or even (*had*) *better*, failed to be used significantly in their written corpus data. They hypothesized that these constructions were perhaps "too colloquial, as yet, to be acceptable in written texts" (*ibid.*). In future studies it would be useful to re-evaluate what elements are currently too colloquial to be utilized regularly in published texts and if this limit has changed over the past decades.

References

Adams, Michael. 2009. *Slang: The people's poetry*. Oxford: Oxford University Press.

—. 2014. *Slang in new media. A case study*. In Coleman, Julie (ed.), *Global English slang. Methodologies and perspectives*, 175-86. London: Routledge.

Aijmer, Karin. 2002. *English Discourse Particles: Evidence from a corpus*. Amsterdam/Philadelphia: John Benjamins.

—. 2004. Pragmatic Markers in spoken interlanguage. *Nordic Journal of English Studies* 3.173-190.

—. 2013. *Understanding pragmatic markers: a variational pragmatic approach*. Edinburgh: Edinburgh University Press.

—. 2015. Pragmatic markers. In Rühlemann, Christoph & Aijmer, Karin (eds.), *Corpus Pragmatics: A Handbook*, 195-218. Cambridge: Cambridge University Press.

Aijmer, Karin & Simon-Vandenbergen, Anne-Marie. 2004. A model and a methodology for the study of pragmatic markers: the semantic field of expectation. *Journal of Pragmatics* 36. 1781-1805.

Allen, Keith & Burridge, Kate. 1991. *Euphemism and Dysphemism: Language Used as Shield and Weapon*. New York: Oxford University Press.

Allwood, Jens, Joakim Nivre & Ahlsén, Elisabeth. 1990. Speech Management—on the Non-written Life of Speech. *Nordic Journal of Linguistics* 13. 3-48.

Andersen, Gisle. 1997. They gave us these yeah, and they like wanna see like how we talk and all that: The Use of like and Other Discourse Markers in London Teenage Speech. In Kostinas, Ulla-Britt, Stenström, Anna-Brita & Karlsson, Anna-Malin (eds.), *Ungdomsspråk i Norden*, 82-95. Stockholm: Institutionen för Nordiska Språk, Stockholms Univ.

Ariel, Mira. 1994. Pragmatic operators. In Asher, Ron E. (ed.), *The Encyclopedia of language and linguistics*. Oxford: Pergamon Press.

Baron, Naomi S. 2000. *Alphabet to email: How written English evolved and where it's heading*. London: Routledge.

Barton, David & Lee, Carmen. 2013. *Language online: Investigating digital texts and practices*. London: Routledge.

Beeching, Kate. 2016. *Pragmatic markers in British English: Meaning in social interaction*. Cambridge: Cambridge University Press.

Benvenist, Emile. 1971. Subjectivity in language. *Problems in general linguistics*. Trans by Mary Elizabeth Meek, Coral Gables FL. University of Miami Press. Original publication 1958.

Bergs, Alexander. 2018. Because science! Notes on a variable conjunction. In Seoane, Elena, Acuña-Fariña, Carlos, & Palacios-Martínez, Ignacio (eds.), *Subordination in English: Synchronic and diachronic perspectives,* 43-60. Berlin/Boston: De Gruyter Mouton.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

—. 1995. *Dimensions of Register Variation: A Cross-Linguistic Comparison*. Cambridge: Cambridge University Press.

Biber, Douglas & Conrad, Susan. 2009. *Register, Genre, and Style.* Cambridge: Cambridge University Press.

Biber, Douglas & Finegan, Edward. 1989. Drift and evolution of English style: a history of three genres. *Language* 65. 487-517.

Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward. 1999. *Longman grammar of spoken and written English*. Essex: Pearson Education Limited.

Blakemore, Diane. 2002. *Relevance and Linguistic Meaning: The Semantics and Pragmatics of Discourse Markers*. Cambridge: Cambridge University Press.

Blyth, Carl, Jr., Recktenwald, Sigrid & Wang, Jenny. 1990. I'm like, 'say what?!': A New Quotative in American Oral Narrative. *American Speech* 65. 215-227.

Bolinger, Dwight. 1989. *Intonation and its uses: Melody and grammar in discourse*. Stanford: Stanford University Press.

Bréal, Michel. 1964. *Semantics: Studies in the Science of meaning*. Trans by Mrs. Henry Cust. New York: Dover. Original publication 1900.

Breban, Tine. 2006. Grammaticalization and subjectification of the English adjectives of general comparison. In Athanasiadou, Angeliki, Canakis, Costas & Cornillie, Bert (eds.), *Subjectification: Various Paths to Subjectivity*, 241-277. Berlin/New York: De Gruyter Mouton.

Brinton Laurel, J. 1996. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. Berlin/New York: Mouton de Gruyter.

Brizuela, Maquela, Andersen, Elaine & Stallings, Lynne. 1999. Discourse markers and indicators of register. *Hispania* 82. 128-141.

Buchstaller, Isabelle. 2006. Social stereotypes, personality traits and regional perception displaced: Attitudes towards the 'new' quotatives in the U.K.1. *Journal of Sociolinguistics* 10. 362-381.

Clift, Rebecca. 2001. Meaning in Interaction: The Case of Actually. *Language* 77. 245-291.

Coleman, Julie. 2014. Introduction: Understanding slang in a global context. In Coleman Julie (ed.), *Global English Slang: Methodologies and Perspectives*, 1-9. London/New York: Routledge.

Crystal, David. 2001. *Language and the Internet*. Cambridge: Cambridge University Press.

—.2006. *The Cambridge Encyclopedia of the English Language*. Cambridge: Cambridge University Press.

Cuyckens, Hubert, Davidse, Kristin & Vandelanotte, Lieven. 2010. Introduction. In Cuyckens, Hubert, Davidse, Kristin & Vandelanotte, Lieven (eds.), *Subjectification, intersubjectification and grammaticalization*, 1-26. Berlin/New York: de Gruyter Mouton.

D'Arcy, Alexandra. 2006. Lexical Replacement and the Like(s). *American Speech* 81. 339-357.

—. 2007. Like and Language Ideology: Disentangling fact from fiction. *American Speech* 82. 386-419.

Dailey-O'Cain, Jennifer. 2000. The Sociolinguistic Distribution of and Attitudes Toward Focuser like and Quotative like. *Journal of Sociolinguistics* 4. 60-80.

De Fina, Anna. 1997. An analysis of Spanish bien as a marker of classroom management in teacher-student interaction. *Journal of Pragmatics* 28. 337-354.

de Klerk, Vivian. 2005. Procedural meanings of well in a corpus of Xhosa English. *Journal of Pragmatics* 37. 1183-205.

de Smet, Hendrik & Verstraete, Jean-Christophe. 2006. Coming to terms with subjectivity. *Cognitive Linguistics* 17. 365-392.

Defour, Tine. 2010. The semantic-pragmatic development of well from the viewpoint of (inter)subjectification. In Davidse, Kristin, Vandelanotte, Lieven & Cuyckens Hubert (eds.), *Subjectification, intersubjectification and grammaticalization*, 155-195. Berlin/New York: De Gruyter Mounton.

Diessel, Holger. 2017. Usage-based linguistics. In. Aronoff, Mark (ed.), *Oxford Research Encyclopedia of Linguistics.* New York: Oxford University Press.

"Discourse Markers. (so, right, okay)" *Cambridge Dictionary Online: English Grammar Today*. Cambridge: Cambridge University Press. <https://dictionary.cambridge.org/grammar/british-grammar/discourse markers/discourse-markers-so-right-okay> (accessed 20 March 2018).

Du Bois, John. 2007. The Stance Triangle. In Englebretson, Robert (ed.), *Stancetaking in Discourse: Subjectivity, evaluation, interaction*, 139-182. Amsterdam/Philadelphia: John Benjamins

Eckert, Penelope & McConnell-Ginet, Sally. 1992. Communities of practice: Where language, gender, and power all live. In Hall, Kira, Buchholtz, Mary & Moonwomon, Birch (eds.), *Locating Power, Proceedings of the 1992 Berkeley Women and Language Conference*, 89-99. Berkeley: Berkeley Women and Language Group.

—. 1999. New generalizations and explanations in language and gender research. *Language in Society* 28. 185-201.

Elmer-Dewit, Philip. 1994. Bards of the Internet: If E-mail represents the renaissance of prose, why is so much of it so awful? *Time*, 4 July, 66-67.

Even-Zohar, Itamar. 1982. The emergence of speech organisers in a renovated language: The case of Hebrew void pargmatic connectives. In Enkvist, Nils E. (ed.), *Impromptu Speech: A symposium on problems in the linguistics study of impromptu speech*, 179-193. Abo: Abo Academi Forskiningsinstitut.

Finell, Anne. 1989. Well now and then. *Journal of Pragmatics* 13. 653-56.

Fairclough, Norman. 2003. *Analysing discourse: Textual analysis for social research*. London: Routledge.

Fischer, Kerstin. 2006. Frames, constructions, and invariant meanings: the functional polysemy of discourse particles. In Fischer, Kerstin (ed.), *Approaches to Discourse Particles*, 427-448. Amsterdam: Elsevier.

Fraser, Bruce. 1988. Types of English discourse markers. *Acta Linguistica Hungarica* 38. 19-33.

—. 1990. An approach to discourse markers. *Journal of Pragmatics* 14. 383-398.

—. 2009. An Account of Discourse Markers. *International Review of Pragmatics* 1. 293-320.

Fried, Mirjam & Östman, Jan-Ola . 2005. Construction Grammar and spoken language: The case of pragmatic particles. *Journal of Pragmatics* 37. 1752-1778.

Fuller, Janet M. 2003a. The influence of speaker roles on discourse marker use. *Journal of Pragmatics* 35. 23-45.

—. 2003b. Use of the discourse marker like in interviews. *Journal of Sociolinguistics* 7. 365-377.

Greasley, Peter. 1994. An investigation into the use of the particle well: Commentaries on a game of snooker. *Journal of Pragmatics* 22. 477-494.

Gries, Stefan Th. 2013. Elementary statistical testing with R. In Schlüter, Julia & Krug, Manfred (eds.), *Research Methods in Language Variation and Change*, 361-381. Cambridge: Cambridge University Press.

Gries, Stefan Th. & Hilpert, Martin. 2008. The identification of stages in diachronic data: variability-based neighbor clustering. *Corpora* 3.59-81.

—. 2012. Variability-based neighbor clustering: A bottom-up approach to periodization in historical linguistics. In Nevalainen, Terttu & Traugott, Elizabeth C. (eds.), *The Oxford Handbook of the History of English*, 134-44. Oxford: Oxford University Press.

Gumperz, John J. 1996. The linguistic and cultural relativity of inference. In Gumperz, John J. & Levinson, Steven C. (eds.), *Rethinking Linguistic Relativity*, 374-406. Cambridge: Cambridge University Press.

Gutzmann, Daniel & Robert Henderson. 2019. Expressive updates, much? *Language* 95. 107-135.

Haddican, Bill & Daniel E. Johnson. 2015. Change in the syntax and semantics of be like quotatives. In Biberauer, Theresa & Walkden, George (eds.), *Syntax over Time: lexical, morphological and information-structural Interactions*, 54–71. Oxford: Oxford University Press.

Halliday, Michael A. K. 1994. *An introduction to functional grammar*. London: Edward Arnold.

Haselow, Alexander. 2019. Discourse marker sequences: Insights into the serial order of communicative tasks in real-time turn production. *Journal of Pragmatics* 146. 1-18.

Hilpert, Martin. 2012. Diachronic collostructional analysis: how to use it and how to deal with confounding factors. In Allan, Kathryn & Robinson, Justyna A. (eds.), *Current Methods in Historical Semantics*, 130-60. Berlin/Boston: De Gruyter Mouton.

—. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax* Cambridge: Cambridge University Press.

Hilpert, Martin & Bourgeois, Samuel. 2020. Intersubjectification in constructional change: From confrontation to solidarity in the "Sarcastic much?" construction. In Timponi Torrent, Tiago, Edison da Silvas Matos, Ely, & Sathler Sigiliana Natália (eds.), *Construction Grammar across Borders*, 96–120. Amsterdam/Philadelphia: John Benjamins.

Hutchinson, Ben. 2003. Automatic classification of discourse markers on the basis of their co-occurrences. In Steds, Manfred & Zeevat, Henk, (eds.), *Proceedings of the ESSLLI Workshop: The Meaning and implementation of discourse particles*, 1-8. Nancy.

Hundt, Marianne & Mair, Christian. 1999. "Agile" and "Uptight" Genres: The Corpus-Based Approach to Language Change in Progress. International Journal of Corpus Linguistics 4 (1999): 221-242.

Jaffe, Alexandra. 2009. Introduction: The sociolinguistics of stance. In Jaffe, Alexandra (ed.) *Stance: Sociolinguistic perspectives*, 3-28. Oxford: Oxford University Press.

James, Deborah. 1974. *The Syntax and semantics of some English interjections*. Ann Arbor, Michigan: University of Michigan Papers.

Jaworski, Adam & Thurlow, Crispin. 2009. Taking an elitist stance: Ideology and the discursive production. In Jaffe, Alexandra (ed.) *Stance: Sociolinguistic perspectives*, 195-226. Oxford: Oxford University Press.

Johnstone, Barbara. 2010. Locating Language in Identidy. In Llamos, Carmen & Watt, Dominic (eds.), *Language and Identities*, 29-36. Edinburgh: Edinburgh University Press.

Jones, Graham M., Bambi B. Schieffelin & Rachel E. Smith. 2011. When friends who talk together stalk together: Online gossip as metacommunication. In Thurlow, Crispen & Mroczek, Kristine (eds.), *Digital discourse: language in the new media*, 26-47. Oxford: Oxford University Press.

Jucker, Andreas. 1997. The discourse marker well in the history of English. *English Literature and Linguistics* 1. 91-110.

Jucker, Andreas H. 1993. The discourse marker well: A relevance-theoretical account. *Journal of Pragmatics* 19. 435-452.

—. 2000. English historical pragmatics: Problems of data and methodology. In Di Martino, Gabriella & Lima, Maria (eds.), *English diachronic pragmatics*, 17- 55. Napoli: CUEN.

Jucker, Andreas H & Smith, Sara W. 1998. And people just you know like ‘wow’. Discourse Markers as negotiating strategies. In Jucker, Andreas H. & Ziv, Yael (eds.), *Discourse Markers. Description and Theory*, 171-201. Amsterdam/Philadelphia: John Benjamins.

Jucker, Andreas H. & Ziv, Yael. 1998. Discourse Markers: Introduction. In Jucker, Andreas & Ziv, Yael (eds.), *Discourse markers: Descriptions and theory*, 1-12. Amsterdam/Philadelphia: John Benjamins.

Kaltenböck, Gunther, Heine, Bernd & Kuteva, Tanie. 2011. On thetical grammar. *Studies in Language* 4. 852-897.

Koch, Peter & Oesterreicher, Wulf. 1985. Sprache der Nähe - Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch 36*. 15-43.

—. 1990. *Gesprochene Sprache in der Romania: Französisch, Italienische, Spanisch.* Tübingen: Max Niemeyer.

Koops, Christian & Arne Lohmann. 2015. A quantitative approach to the grammaticalization of discourse markers: Evidence from their sequencing behaviour. *International Journal of Corpus Linguistics* 20. 232-259.

Kress, Gunther. 1995. The social production of language: History and structures of domination. Discourse in society: systematic functional perspectives. In Fries, Peter H. & Gregory, Michael (eds.) *Meaning and choise in Language: Studies for Michael Halliday*, 115-40. Westport, CT: Ablex Publishing.

Kytö, Merja. 2010. Data in historical pragmatics. In Jucker, Andreas & Taavitsainen, Irma (eds.), *Historical Pragmatics*, 33-68. Berlin/New York: De Gruyter Mouton.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford: Blackwell.

Lakoff, Robin. 1973. Questionable answers and answerable questions. In Kachru Braj B., Lees, Robert B., Malkiel, Yakov & Saporta, Sol (eds.), *Issues in Linguistics: Papers in Honor of Henry and Rene Kahane*, 453-467. Urbana: University of Illinois Press.

Lam, Phoenix 2006. Well but that's the effect of it: The use of well as a discourse particle in talk shows. *Sprache und Datenverarbeitung (International Journal for Language Data Processing)* 30. 99-108.

Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1, 5-38.

—.1999. Losing control: grammaticalization, subjectification, and transparency. In Blank, Andreas & Koch, Peter (eds.), *Historical Semantics and Cognition*, 147-175. Berlin/New York: de Gruyter Mouton.

—.2008. *Cognitive Grammar: A basic introduction*. Oxford: Oxford University Press.

Leech, Geoffrey, Hundt, Marianne, Mair, Christian & Smith, Nicolas. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.

Leech, Geoffrey, Smith, Nicolas. 2009. Change and constancy in linguistic change: how grammatical usage in written English evolved in the period 1931-1991. In Renouf, Antoinette. & Kehoe, Andrew (eds.), 173-200. Amsterdam/New York: Brill | Rodopi.

Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.

Lohmann, Arne & Christian Koops. 2016. Aspects of discourse marker sequencing: Empirical challenges and theoretical implications. In Kaltenböck, Gunther, Keizer, Evelien & Lohmann, Arne (eds.), *Outside the Clause: Form and function of extra-clausal constituents*, 417–446. Amsterdam/Philadelphia: John Benjamins.

López-Couso, María José. 2010. Subjectification and intersubjectification. In Jucker, Andreas H. &Taavitsainen, Irma (eds.), *Historical Pragmatics*, 127-163. Berlin/New York: De Gruyter Mouton.

Mair, Christian. 1997. Parallel corpora: a real time approach to the study of language change in progress. In Ljung, Magnus (ed.), *Corpus-based studies in English*, 195-209. Amsterdam: Rodopi.

—. 2006. *Twentieth-Century English: History, Variation and Standardization*. Cambridge: Cambridge University Press.

—. 2009. Corpus linguistics meets sociolinguistics: Studying educated spoken usage in Jamaica on the basis of the International Corpus of English. In Hoffmann, Thomas & Siebers, Lucia (eds.), *World Englishes – Problems, Properties and Prospects: Selected papers from the 13th IAWE Conference*, 39-60. Amsterdam/Philadelphia: John Benjamins.

Maschler, Yael. 1994. Metalanguaging and discourse markers in bilingual conversation. *Language and Society* 23. 325-366.

Meehan, Teresa. 1991. It's Like, 'What's Happening in the Evolution of Like?': A Theory of Grammaticalization. *Kansas Working Papers in Linguistics* 16. 37-51.

Moore, Robert L. 2004. We're Cool, Mom and Dad Are Swell: Basic Slang and Generational Shifts in Values. *American Speech*. 79. 59-86.

Mortier, Liesbeth & Degand, Liesbeth. 2009. Adversative discourse markers in contrast: The need for a combined corpus approach. *International Journal of Corpus Linguistics* 14. 338-366.

Müller, Simone. 2005. *Discourse Markers in Native and Non-native English Discourse.* Amsterdam/Philadelphia: John Benjamins.

Murray, Dinah. 1979. Well. *Linguistic Inquiry* 10. 727-732.

Narrog, Heiko. 2010. (Inter)subjectification in the domain of modality and mood: Concepts and cross-linguistic realities. In Cuyckens, Hubert, Davidse, Kristin Vandelanotte, Lieven (eds.), *Subjectification, intersubjectification and grammaticalization*, 385-429. Berlin/New York: de Gruyter Mouton.

Norén, Kerstin & Linell, Per. 2007. Meaning potentials and the interaction between lexis and contexts: An empirical substantiation. *Pragmatics* 17. 387-416.

Norrick, Neal R. 2001. Discourse markers in oral narrative. *Journal of Pragmatics* 33.849-878.

O'Keeffe, Anne & Adolphs, Svenja. 2008. Response tokens in British and Irish discourse. In Schneider, Klaus & Barron, Anne (eds.), *Variational Pragmatics: A focus on regional varieties in pluricenntric languages*, 69-98. Amsterdam/Philadelphia: John Benjamins.

Ochs, Elinor. 1996. Linguistic resources for socializing humanity. In Gumperz, John J & Levinson, Stephen C. (eds.), *Rethinking Linguistic Relativity*, 407-437. Cambridge: Cambridge University Press.

Oh, Sun-Young. 2000. Actually and in fact in American English: a data-based analysis. *English Language and Linguistics* 4. 243-268.

Pogner, Karl-Heinz. 2005. Discourse communities and communities of practice. *EGOS Colloquium* working paper no 80.

"oh yeah, adv." *OED Online*. Oxford University Press (oed.com/view/Entry/231470) (Accessed 2019-10-06)

Redeker, Gisela. 2006. Discourse markers as attentional cues at discourse transitions. In Fischer, Kerstin (ed.), *Approaches to discourse particles*, 339-358. Amsterdam: Elsevier.

Romaine, Suzanne & Deborah Lange. 1991. The Use of like as a Marker of Reported Speech and Thought: A Case of Grammaticalization in Progress. *American Speech* 66. 227-279.

Rosch, Eleanor. 1978. Principles of categorization. In Rosch, Eleanor & Lloyd, Barbara, B. (eds.), *Cognition and Categorization*. Hillsdale, Mi: Lawrence Erlbaum.

Rühlemann, Christoph & Hilpert, Martin. 2017. Colloquialization in journalistic writing: The case of inserts with a focus on well. *Journal of Historical Pragmatics* 18. 104-135.

Rühlemann, Christoph & Aijmer, Karin. 2015. Introduction. In Aijmer, Karin & Rühlemann, Christoph (eds.), *Corpus Pragmatics: A handbook*, 1-26. Cambridge: Cambridge University Press.

Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge Univ. Press.

Schourup, Lawrence. 1985. *Common discourse particles in English conversation*. New York: Garland Publishing Inc.

—. 1999. Discourse markers. *Lingua* 107. 227-265.

—. 2001. Rethinking well. *Journal of Pragmatics* 33. 1025-1060.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language & Communication* 23. 193-229.

Smith, Sara W. & Jucker, Andreas H. 2000. Actually and other markers of an apparent discrepancy between propositional attitudes of conversational partners. In Andersen, Gisle & Fretheim, Thorstein (eds.), *Pragmatic Markers and Propositional Attitude*, 205-237. Amsterdam/Philadelphia: John Benjamins.

Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell.

Stefanowitsch, Anatol & Gries, Stefan Th. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8. 209-243.

Svartvik, Jan. 1980. Well in conversation. In Greenbaum, Sidney, Leech, Geoffrey & Svartvic Jan (eds.), *Studies in English Linguistics for Randolph Quirk*, 167-77. London: Longman.

Svartvic, Jan & Stenström, Anna-Brita. 1985. Words, words, words: The rest is silence? In Bäckman, Sven & Kjellmer, Goran (eds.), *Papers on Language and Literature: Presented to Alvar Ellegard and Erik Frykman*, 342-353. Gothenburg: Gothenburg Studies in English.

Swales, John M. 1990. *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.

—. 2016. Reflections on the concept of discourse community. *ASp* 69. 1-12.

Taavitsainen, Irma. 1998. Emphatic language and romantic prose: Changing functions of interjections in a sociocultural perspective. *European Journal of English Studies* 2. 195-214.

Talbot, Mary M. 1994. Relevance. In Asher, Ron E. (ed.), *Encyclopedia of Language and Linguistics*, 3524-3527. Oxford: Pergamon.

Taylor, John R. 2001. Linguistics: Prototype Theory. In Smelser, Neil J., Baltes, Paul B. (eds.), *International Encyclopedia of the Social & Behavioral Sciences*, 8954-8957. Oxford: Pergamon.

Thurlow, Crispin & Kristine Mroczek. 2011. Introduction: Fresh perpectives on new media sociolinguistics. In Thurlow, Crispin & Kristine Mroczek (eds.), *Digital discourse: Language in the New Media*, xviv-xliv. Oxford: Oxford Universtiy Press.

Tottie, Gunnel. 2017. From pause to word: *uh*, *um* and *er* in written American English. *Language and Linguistics* 23. 105-130.

Traugott, Elizabeth C. 1989. On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change. *Language* 65. 31-55.

—.1997. The role of the development of discourse markers in a theory of grammaticalization. *Paper Presented at ICHL XII, Manchester 1995.*

—. 2003. From subjectification to intersubjectification. In Hickey, Raymond (ed.), *Motives for Language Chang*e, 124-140. Cambridge: Cambridge University Press.

—. 2010. (Inter)subjectivity and (inter)subjectification: A reassessment. In Davidse, Kristin, Vandelanotte, Lieven & Cuyckens, Hubert (eds.), *Subjectification, Intersubjectification and Grammaticalization*, 29-74. Berlin/Boston: De Gruyter Mouton.

—. 2018. Modeling Language Change with Constructional Networks. In Bordería, Salvador P. & Lamas, Óscar L. (eds.), *Beyond Grammaticalization and Discourse Markers: New Issues in the Study of Language Change*, 17-50. Leiden: Brill.

Traugott, Elizabeth C. & Dasher, Richard B. 2002. *Regularity in Semantic Change*. Cambridge: Cambridge University Press.

Tree, Jean E. Fox. 2006. Placing like in telling stories. *Discourse Studies* 8. 723-743.

Underhill, Robert. 1988. Like is, like, a focus. *American Speech* 76. 234-238.

Verschueren, Jef. 1999. *Understanding pragmatics*. London: Arnold Publishers.

Watts, Richard J. 1988. A relevance-theoretic approach to commentary pragmatic markers: The case of “Actually”, “Really” and “Basically”. *Acta Linguistica Hungarica* 38. 235-60.

—. 1989. Taking the pitcher to the ‘well’: Native speakers' perception of their use of discourse markers in conversation. *Journal of Pragmatics* 13. 203-37.

— 1999. The social construction of Standard English: Grammar writers as a 'discourse community'. In Bex, Tony & Watts, Richard J. (eds.), *Standard English: The widening debate*, 40-68. London and New York: Routledge.

Westin, Ingrid. 2002. *Language change in English newspaper editorials*. Amsterdam: Radophi.

Westin, Ingrid & Christer Geisler. 2002. A multi-dimensional study of diachronic variation in British newspaper editorials. *ICAME Journal* 26.133-52.

Zappavigna, Michele. 2012. *Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web*. London/New York: Continuum.

Corpora

Alexander, Mark. & Davies, Mark. (2015). *Hansard Corpus 1803-2005*. (http://www.hansard-corpus.org.)(Accessed 2020-01-22)

Davies, Mark. 2007. *TIME magazine corpus: 100 million words, 1920s-2000s*. (https://www.english- corpora.org/time/) (accessed 22 January 2020)

Davies, Mark. 2010. *The corpus of historical American English (COHA): 400 million words, 1810-2009*. (https://www.english-corpora.org/coha/) (accessed 22 January 2020)

Davies, Mark. 2016. *Corpus of news on the web (NOW): 3+ billion words from 20 countries, updated every day*. (https://www.english-corpora.org/now/) (accessed 22 January 2020)

Dictionaries & Writing References

OED Online. 2019. Oxford University Press. <http://www.oed.com/viewdictionaryentry/Entry/11125> (accessed 23 January 2020).

Cambridge Dictionary Online. 2019. Cambridge University Press. < https://dictionary.cambridge.org/> (accessed 23 January 2020)

Urban Dictionary. 2019. <https://www.urbandictionary.com>

Software/R packages

Flach, Susanne. 2017. An R Implementation for the Family of Collostructional Methods Version: 0.1.0. [Computer Software]. Berlin, Germany. (http://userpage.fuberlin.de/~flach/corpling/collostructions/) (accessed 23 January 2020).

Gries, Stefan Th. & Hilpert, Martin. 2008. The identification of stages in diachronic data: variability-based neighbor clustering. *Corpora* 3.59-81.