Hikaru Hotta* and Martin Hilpert

English comparative constructions at different levels of schematicity: what is the role of adjective-specific variability?

https://doi.org/10.1515/cog-2025-0025
Received February 13, 2025; accepted September 2, 2025; published online September 26, 2025

Abstract: This study analyzes the English comparative alternation (morphological *prouder* and periphrastic *more proud*) with regard to adjective-specific variability. Substantial empirical evidence indicates that speakers redundantly represent both abstract schemas (e.g., ADJ + *-er*) and their specific instantiations (e.g., *prouder*) as symbolic units, which together constitute a network of constructions at different levels of schematicity. Against this background, we offer a re-analysis of data from (Hilpert, Martin. 2008. The English comparative - Language structure and use. *English Language & Linguistics* 12(3). 395–417), which yields new results on the basis of a mixed-effects (multi-level) model that incorporates individual adjectives as random effects. The results demonstrate robust lexical idiosyncrasies in this alternation. First, many adjectives exhibit tendencies that deviate from the general conditioning factors incorporated as fixed effects. Furthermore, frequency of adjectives in comparative forms positively correlates with the degree of skewness. In other words, the more an adjective is used in comparative forms, the stronger its preference becomes, regardless of the direction of the preference. These findings support the idea that competition between linguistic forms occurs across different levels of schematicity: at the schematic level, the competition between variants is governed by probabilistic factors, while at the lexically specific level, the conventionalization of concrete forms can override these factors.

Keywords: English comparatives; mixed-effects regression; schematicity; variation; conventionalization; frequency effects

***Corresponding author: Hikaru Hotta**, Institute of English Studies, University of Neuchâtel, Neuchatel, Switzerland, E-mail: hikaru.hotta@unine.ch. https://orcid.org/0009-0000-1873-0229
Martin Hilpert, Institute of English Studies, University of Neuchâtel, Neuchatel, Switzerland

Open Access. © 2025 the author(s), published by De Gruyter. This work is licensed under the Creative Commons Attribution 4.0 International License.

1 Introduction

This paper examines adjective-specific variability in the English comparative alternation between the morphological variant (e.g., *prouder*) and the periphrastic variant (e.g., *more proud*).[1] In corpus-based studies of English alternations, such as the dative alternation (Bresnan et al. 2007), the particle placement alternation (Gries 2003), or the genitive alternation (Szmrecsanyi 2010), the alternating constructions are typically abstract or schematic, containing open slots for lexical elements.

For the comparative alternation, the alternating schematic variants are [ADJ-*er*] and [*more* ADJ], where adjectives capable of adopting both variants fill the ADJ slots. Previous research (Hilpert 2008; Mondorf 2009, amongst others, cf. Section 2.2) has identified a range of factors that influence the choice between these comparative schemas, which can be seen as parts of a constructional network. In a usage-based approach to language, that constructional network is thought to encompass not only schematic constructions but also concrete, lexically specific constructions. Usage-based Construction Grammar (e.g., Croft 2022; Diessel 2023) defines constructions as conventional pairings of form and meaning across all levels of schematicity, from argument structure constructions (Goldberg 1995) to words or multi-word units (Dąbrowska 2009). In morphology, abstract patterns are schemas such as $[V_{TR} - \textit{able}]_{adj}$, in which the suffix ‘able’ is added to a transitive verb in order to derive adjectives such as *acceptable, affordable*, or *doable* (Booij 2013). Importantly, both individual complex words and their parent schemas can be specified in the grammar, avoiding the erroneous notion that linguistic constructs that are fully regular and transparent need not be mentally represented (i.e., *the rule/list fallacy*, Langacker 1987: 29). This results in what are often referred to as redundant representations of language (Hilpert 2021; Langacker 2000), since both analytical and holistic processing can produce the same output. However, it is important to note that the term ‘redundant’ here does not imply ‘unnecessary’. Schematic constructions enable the productive potential of language, while concrete constructions support fluent, conventional utterances, including those that are tailored to particular social groups, registers, or even an individual’s specific patterns of language usage (Biber and Conrad 2019; Pawley and Syder 1983; Szmrecsanyi 2005; Wray 2002).

Supporting this idea, psycholinguistic evidence shows that speakers engage in holistic processing for highly frequent regularly inflected forms, such as *walked* or

---

1 The two variants may be termed ‘synthetic’ and ‘analytic’, a terminology argued to be preferable in Mondorf (2009: 1). In this paper, we prioritize consistency with Hilpert (2008), from which our dataset is drawn.

ladies (Katz et al. 1991; Sereno and Jongman 1997). Frequency effects have also been observed for multi-word sequences in English (Arnon and Snider 2010). However, for multi-unit phrasal patterns, findings about frequency effects remain inconclusive (Hotta 2025; Jolsvai et al. 2020). Nonetheless, there is no denying that constructional networks encompass a substantial range of lexically specific patterns, as extensive empirical evidence from both corpus-based and experimental studies demonstrates (e.g., Ferreira and Patson 2007; Pijpops and Van de Velde 2016).

The goal of this study is to explore the role of lexically specific constructions in the English comparative alternation. If speakers represent individual comparative forms, such as *wealthier* and *more mature,* what does this imply for the aggregate variation in English comparatives? We address this question by re-analyzing data from Hilpert (2008) through a mixed-effects regression model, which allows us to examine the idiosyncratic variability of individual adjectives, specified as random effects. Additionally, we demonstrate that the frequency of adjectives in comparative forms correlates with the strength of bias for one variant over the other in the alternation, on top of what is explained by general predictors. We argue that these findings support the view that competition occurs at multiple levels of schematicity. At the schematic level, the competition between variants is governed by probabilistic factors, while at the lexically specific level, the conventionalization of concrete forms can override these factors.

The remainder of this paper is structured as follows: Section 2 discusses the theoretical background, focusing on the notion of competition in language and reviewing research on the English comparative alternation. Section 3 introduces the data and methodology. In Section 4, we present the results of a mixed-effects logistic regression analysis. Section 5 offers a discussion that interprets the results in terms of variability and competition at different levels of schematicity. Our conclusions are presented in Section 6.

2 Background

2.1 Competition at multiple levels of schematicity

Building on the growing body of empirical research on lexically specific constructions, several studies have investigated lexical specificity in linguistic alternations. Perek (2014) uses collostructional analysis (Stefanowitsch and Gries 2003) to examine the English *at*-alternation, demonstrating that the construction expresses distinct meanings depending on the semantic class of the verb. Examples (1) and (2) below, taken from Pijpops et al. (2021: 490), effectively illustrate this contrast. In (1), with the “cutting” verb *chip*, the *at*-variant conveys a repeated action in which contact with the object is made. In contrast, when the variant involves a “striking” verb such as *slap*, as in (2), it conveys an attempted action where contact with the object is not necessarily made.

(1) *Sam chipped at the rock* (Broccias 2001: 77)

(2) *He slapped at it with his other hand, but it was beneath his thumbnail before he could get at it.* (Perek 2015: 134)

Pijpops et al. (2021) further develop the approach introduced by Perek (2014), employing collostructional analysis at multiple levels of schematicity and utilizing distributional semantics to investigate the Dutch *naar*-alternation. Crucially, their findings do not support systematic variation in the *naar*-alternation without lexical specifications, possibly due to verb-specific *naar*-alternations that exhibit distinct meaning differences. In other words, since lower-level alternations exhibit distinct distributions that are independent of each other, this dialutes clear semantic distinction at the most schematic level.

Both Perek (2014) and Pijpops et al. (2021) highlight the importance of taking into account the multilevel structure of the constructional network in alternation research. While systematic variation can emerge at any level of schematicity, lexemes that fill the open slots of constructions can cluster based on the meaning expressed, giving rise to lower-level generalizations. Alternations at lower levels tend to capture more tangible meaning differences, which arguably align more closely with the distinctions to which speakers are attuned. The alternations at the schematic level in these studies (English *at*-alternation and Dutch *naar*-alternation without any lexical specifications) may emerge as conglomerates of their lower-level alternations, without directly addressing a functional need of the speaker.

Most studies on alternations focus on cases where speakers of a language can choose between two (or more) variants (see Pijpops 2020 for a discussion on the notion of alternations). However, many lexical elements that occupy the slots of alternating constructions exhibit a bias toward one variant over the other. A method that can be used to reveal lexical bias in alternations is distinctive collexeme analysis (Gries and Stefanowitsch 2004). For instance, in the case of the English dative alternation, the verb *tell* exhibits a substantial preference for the ditransitive construction over the *to*-dative (128_{DITR}:2_{TO}), whereas *send* (64_{DITR}:113_{TO}) has a more balanced distribution (Gries and Stefanowitsch 2004: 106). In the dative alternation, lexical bias is so pervasive that the verbs can actually be used to predict the correct variant with well above chance accuracy (Gries and Stefanowitsch 2004: 119). The strong lexical bias in the dative alternation is corroborated by many other studies (e.g., Lehmann and Schneider 2012; Röthlisberger et al. 2017).

There are various ways to interpret lexical bias in alternations. Gries and Stefanowitsch (2004) argue that distinctive collexemes represent semantic differences between the two variants of the alternation. For example, considering that the central meaning of the ditransitive construction involves the successful transfer of an object from an agent to a recipient (Goldberg 1995), the preference of *tell* for the ditransitive may be explained as a metaphorical extension of communication as transfer (Gries and Stefanowitsch 2004: 106) However, quantitative corpus-based studies generally suggest that the determinants of alternations are inherently probabilistic, encompassing, for instance, phonological, syntactic, semantic, and discourse-level factors (e.g., Bresnan et al. 2007; Mondorf 2009; Röthlisberger et al. 2017). Can these conditioning factors account for the strong bias of *tell* in the ditransitive? An even more extreme example is the verb *explain*, which almost invariably occurs with the *to*-dative construction. For verbs like *explain*, detailed memory traces of their usage in the prepositional dative, where their usage in the ditransitive might have been expected, may preempt speakers from using the ditransitive construction for the verb (Boyd and Goldberg 2011; Goldberg 2011). In other words, for highly biased lexical elements, the probabilistic constraints of the parent schemas may be less relevant to speakers when learning to use the conventional variant.

To understand lexical bias in alternations, it is useful to remind ourselves of the multi-level nature of the constructional network. At the schematic level, distinct lexical items can indeed reflect meaning differences between alternating constructions. Accordingly, the verb *bring* is strongly distinctive for the prepositional dative, as it conveys the meaning of caused motion, rather than successful transfer. As Perek (2014) and Pijpops et al. (2021) have demonstrated, they may be meso-schematic alternations that emerge from clusters of lexemes. Even further down the vertical line of the constructional network, since combinations of schematic constructions and individual lexical items can themselves become entrenched and conventionalized, their entrenchment may override the conditioning factors that govern the alternation between the schematic variants. For the verb *explain*, speakers simply discard the fact that it is semantically compatible with the ditransitive, and they use it exclusively in the prepositional dative. Through progressive conventionalization, one variant may thus ultimately drive out the other (De Smet et al. 2018). This means that lexical bias in alternations can result either from the compatibility between the lexical meaning of a verb and the constructional meaning of an abstract schema, or it can result from the conventionalization of a lexically specific construction (Diessel 2016). As Goldberg (2011: 151) notes in reference to alternations, “While much of language is semantically and historically motivated, there remain pockets of idiosyncrasy that speakers must learn.” One such idiosyncrasy is skewness in distribution, which is often not entirely predictable from the parent schemas.

In morphology, situations where multiple variants fulfill the same grammatical function are referred to as *overabundance*. Thornton (2019) provides a comprehensive overview of this phenomenon, presenting examples of overabundant cells across a diverse range of languages. Similar to syntactic variation, most cases of overabundance are conditioned by a variety of language-internal factors (such as syntactic or semantic variables) and language-external factors (such as social or stylistic variables). However, there is also a significant area where the choice appears to be "free", with no apparent factors determining the selection. If two morphological variants are truly in free variation for a given lexeme, Thornton (2019) predicts that they should occur with roughly equal frequencies. Corroborating this prediction, Bermel and Knittl (2012) report on frequency proportions of competing variants in Czech nominal declension (see Tables 4 and 5 in Bermel and Knittl 2012). In the paradigm of "masculine hard inanimate nouns", genitive singular forms exhibit variation between the endings – *u* and – *a*, while locative singular forms vary between the endings – *u* and – *ě*. They note that, although only small groups of nouns in the paradigm show variation in the first place, within these groups, many nouns exhibit roughly equal frequencies for both variants. Therefore, there are instances where one variant does not appear to be driving out the competing variant in their dataset.

In the case of English comparatives from Hilpert's (2008) data, drawn from the British National Corpus, adjectives such as *angry* (*angrier*/*more angry*), *choosy* (*choosier*/*more choosy*), and *risky* (*riskier*/*more risky*) exhibit a relatively balanced distribution, although most adjectives tend to favor one variant over the other. The conditioning factors influencing the choice of English comparative forms have been extensively studied in the literature, which will be reviewed in the next section.

2.2 The English comparative alternation

English has two ways of expressing the comparative degree: by adding the suffix – *er* or by placing the degree marker *more* before the adjective to form the periphrastic variant. The strongest determinant of the alternation is the length of the adjective (e.g., Hilpert 2008); monosyllabic adjectives overwhelmingly take the morphological form, while adjectives with three or more syllables almost invariably take the periphrastic form. Disyllabic adjectives generally exhibit more robust variation, though some of them, such as *happy,* are extremely skewed (Hilpert 2008, 414–415). However, previous studies have uncovered a complex set of factors influencing the alternation, with no condition rising to the level of a categorical rule. Other factors include the final elements of adjectives and syntactic patterns, such as attributive or predicative usage and the presence of following complements. Mondorf (2003, 2009) argues that these factors can be attributed to considerations in processing effort. Specifically, speakers tend to prefer the more explicit periphrastic variant in cognitively complex environments (Rohdenburg 1996), such as when followed by a complement, as it offers a processing advantage.

From the extensive body of literature on the English comparatives, Hilpert (2008) and Mondorf (2009) provide comprehensive overviews of the conditioning factors. This section focuses on studies that identify idiosyncratic patterns of alternation specific to individual adjectives or clusters of adjectives, aligning with the goals of the present study.

Idiosyncratic patterns within the alternation have been sporadically noted in previous studies. For example, it has been reported that adjectives ending with the sound /i/ do not show a consistent tendency (Mondorf 2009; Chua 2018), despite findings suggesting that this group leans toward the morphological variant (Kytö and Romaine 1997). The variation can be partially explained by separating adjectives ending in /li/ from the broader group, as they have been found to favor the periphrastic variant (Hilpert 2008; Lindquist 1998). However, even within the /li/ group, the tendency appears to vary (Watanabe and Iyeiri 2020). For example, in Hilpert's (2008) data, *costly* (82_{MORE}:$19_{\text{-ER}}$), *deadly* (30:9), and *likely* (3,724:17) prefer the periphrastic form, whereas *lively* (49:76), *lovely* (7:42), and *ugly* (4:41) do not (Hilpert 2008: 414–415). It is possible that the significant effect of the /li/ ending toward the periphrastic form observed in Hilpert (2008) was driven by highly frequent adjectives like *likely*. The fixed-effects-only model used in that study treats the final segments of adjectives as unique to each observation, allowing highly frequent adjectives to disproportionately influence the results. This potential confound is addressed in the present study by incorporating individual adjectives as random effects and treating adjective-dependent variables as second-level predictors (see Section 3.2 for details).

With regard to the degree of bias in comparative forms, D'Arcy (2014) observes that individual adjectives do not exhibit robust variation in casual speech. In the Origin of New Zealand English (ONZE) corpus, which comprises 1.5 million words of casual speech, only a few disyllabic adjectives alternate between the comparative variants. While the author attributes this lack of variation to the style and "simplified grammar" characteristic of casual conversation (D'Arcy 2014: 238), partitioning toward either variant within the English comparative alternation is also observed, albeit to a lesser degree, in other registers, as noted in the previous section with reference to the British National Corpus. Although it is standard practice in variationist research to omit lexical elements that exhibit near-categorical distributions (Pijpops 2020), the existence of such partitioning itself raises theoretically interesting questions for the alternation. This is particularly relevant given the proposed view of the present study, which posits that competition occurs at multiple levels of schematicity.

What remains as a challenge is a systematic examination of variation and idiosyncrasies within the comparative alternation. This requires a method that can quantify the influence of conditioning factors and the idiosyncrasies of individual adjectives simultaneously, while accounting for one another. A mixed-effects regression model, incorporating the nested structure of the dataset, offers a robust solution. The next section provides an overview of the dataset and explains the basics of this regression model and its structure for the present study.

3 Methodology

3.1 Data and annotation

The dataset from Hilpert (2008) that is re-analyzed in the present study is based on comparative usage in the British National Corpus, which is a corpus of 100 million words that contains both written and spoken data.[2] Morphological comparatives are marked with a dedicated part-of-speech tag, which facilitates their exhaustive retrieval. For the periphrastic comparatives, a regular expression was used to retrieve instances of *more* to the left of an adjective. False positives such as *We publish more fine books than ever before* were manually identified and discarded. The dataset only includes alternating adjectives, which yields a total of 247 adjective types. The representation of morphological and periphrastic comparatives in the data set is highly uneven. Morphological comparatives (71,622 tokens) vastly outnumber periphrastic comparatives (8,256 tokens). Each token is annotated for eleven variables (cf. Table 1 below). Besides the number of syllables and morphemes, the dataset holds information on the final phonological segment of the adjective (e.g., final /i/). Tokens are further annotated for the presence or absence of adjective-final stress (e.g., *robúst*) and word-initial stress of its right-side collocate (e.g., *stronger númbers*). With regard to syntactic contexts, the dataset offers annotation on attributive and predicative uses, tokens with *to*-infinitive complementation, a following *than*, and the presence of premodification. The dataset also includes frequency information. Besides the frequency of each adjective type in its positive form, the ratio of positive to comparative forms, i.e., the relative frequency of comparatives, is determined for each adjective type. For the present study, minor

2 A full description of the data retrieval process and the annotation of the data can be found in Hilpert (2008: 400–405).

inconsistencies in the data were corrected.[3] A table with all adjective types and their frequencies in the morphological and periphrastic comparative is offered in Appendix 1.

3.2 Mixed-effects logistic regression

The present study employs logistic mixed-effects models as the primary method of investigation. Logistic regression is especially well-suited for research on linguistic alternations, as it predicts a binary outcome on the basis of multiple continuous and categorical variables. The mixed-effects model incorporates random effects alongside fixed predictors, accounting for dependencies within random groups when estimating fixed effects. For a detailed and practical introduction to applying mixed-effects logistic regression models in corpus-based alternation research, Schäfer (2020) offers a highly informative resource.

Mixed-effects logistic regression models have gained significant traction in corpus-based studies on alternations in recent years. Most studies, however, include random effects primarily to control for group-dependent variability, often without delving deeply into them (a notable exception being Röthlisberger et al. 2017). As Winter (2019, Section 15.9) points out, random effects can provide valuable insights in linguistic research, such as uncovering individual differences or lexical specificity. In this study, individual adjectives receive their own intercepts in the estimation of the fixed effects, offering insights into the extent to which specific adjectives deviate from the overall probabilistic tendencies.

Another advantage of mixed models, or more specifically multilevel models, is their ability to account for the nested structure within a dataset. Table 1 presents a list of fixed effects included in Hilpert's (2008: 407) model. The predictors are grouped into two categories. First-level predictors, such as whether the comparative form is used attributively or predicatively, vary at the level of individual observations. Second-level predictors are group-level variables; they vary across random groups (in this case, individual adjectives) but remain constant within them.[4] For instance, each adjective has a fixed number of syllables, which does not change across its

3 Specifically, a few adjectives in the original dataset received multiple values for variables that should not vary within adjectives (e.g., final elements of adjectives). With this error, such a variable would not be treated as a second-level or group-level predictor in a mixed-effects model (see later in this section) using the lme4 package.

4 A distinction can be made within second-level predictors: those associated with groups of adjectives (e.g., number of syllables, final elements) and those specific to individual adjectives (e.g., frequency counts). The former category can be regarded as meso-schematic constructions (e.g., monosyllabic ADJ – *er*).

Table 1: Fixed effects in Hilpert's (2008) model and their levels in the mixed model in the present study.

Levels	Predictors	Tendency reported in Hilpert (2008)
First-level (observation-level) predictors	*to*-infinitive complementation	Periphrastic
	Attributive and predicative usage	
	Attributive	Morphological
	Predicative	Periphrastic
	A following *than*	Morphological
	Premodification	Periphrastic
	Initial stress of right collocate	Not significant
Second-level (group-level) predictors	Number of syllables	Periphrastic
	The final elements	
	Final /i/	Morphological
	Final /l/	Periphrastic
	Final /r/	Periphrastic
	Final /li/	Periphrastic
	Final consonant cluster	Periphrastic
	Number of morphemes	Periphrastic
	Final stress	Periphrastic
	Frequency of positive form	Morphological
	Comparative/positive ratio	Morphological

observations. Without treating these variables as second-level predictors, highly frequent adjectives overrepresent the data points for these variables. To illustrate, in Hilpert's (2008) model, *simple* contributes 1,175 data points for the syllable count, while *hazy* contributes only 3. By designating these variables as second-level predictors, each adjective contributes only one data point for these variables, ensuring fair representation among adjectives.

4 Results

4.1 Mixed-effects logistic regression

All analyses presented in this section were conducted in R, version 4.4.2 (R Core Team 2024). Mixed-effects logistic regression models were implemented using the lme4 package (Bates et al. 2015).[5] To ensure model convergence, the bobyqa optimizer was

5 For the analysis of random effects, it may be preferable to use a Bayesian model since random effects (group-level effects) are treated as parameters on par with other predictors (Bürkner 2017). In

used. The call for our final model is presented in (3). This model incorporates all the fixed effects listed in Table 1, with their respective levels specified. The frequency of positive form was log-transformed to base *e*, since frequency effects in language generally accrue logarithmically. Additionally, the comparative/positive ratio was logit-transformed to create an unbounded continuous scale, enabling a more interpretable coefficient estimate.[6] The model did not encounter multicollinearity issues (Zuur et al. 2010). Generalized variance inflation factors (GVIF) for the fixed effects were calculated, with the highest corrected $GVIF^{1/2df}$ being 2.26 for the final stress (FINSTR). The random effects include 247 adjectives, each assigned its own intercept to account for group-level variability.

(3) Response = {periphrastic, morphological}
Response ~ TO + ATT_PRED + THAN + PREMOD + RIGHTSTR + SYL + FINAL + MOR + FINSTR + logPOSFREQ + logitCOMPPOSR + (1|ADJ)

4.1.1 Fixed effects

Table 2 presents the results of the fixed effects in the mixed-effects model. A brief explanation of the columns in the table is called for. First, the coefficient estimates indicate both the direction and strength of each variable's effect.[7] Since the model predicts the probability of choosing the periphrastic variant over the morphological variant, positive values indicate a bias toward the periphrastic variant, while negative values indicate a bias toward the morphological variant. The 95 % confidence intervals (CI) reflect the precision of the estimates. Odds ratios provide a more intuitive sense of effect size compared to coefficients on the log-odds scale. For instance, the odds of disyllabic adjectives taking the periphrastic variant are 82.02 times higher than the odds of monosyllabic adjectives doing so. An odds ratio of 1 corresponds to a 50 % probability, meaning values less than 1 indicate a bias toward the morphological variant.

---

this study, however, we prioritized comparability with Hilpert (2008), which relies on frequentist estimation. Furthermore, we note that the fundamental mechanism of the estimation of random effects remains consistent between the two approaches (see Nicenboim et al. 2025 Ch. 5 for a discussion).

6 This is because the ratio ranged from 0 to 0.6, rendering a +1 unit increase for coefficient interpretation not particularly meaningful. Furthermore, logit-transformed proportions are more amenable to logistic regression, since it models the log odds (logits) of the dependent variable.

7 The coefficients for categorical and continuous variables (i.e., the two frequency values) are not directly comparable because they are estimated on different scales. Coefficients for categorical variables reflect contrasts relative to a reference level, whereas coefficients for continuous variables represent the effect of a one-unit (+1) increase in those variables.

Table 2: Coefficient table of model (3).

Levels	Variables	Coefficient estimates	CI low	CI high	Odds ratio	*p* value
	(Intercept)	−7.16	−8.97	−5.35	0	<0.001
First-level	*to*-infinitive complementation	0.65	0.38	0.92	1.91	<0.001
	Attributive or predicative					
	Attributive	−0.6	−0.74	−0.46	0.55	<0.001
	Predicative	0.43	0.28	0.58	1.54	<0.001
	A following *than*	−0.21	−0.37	−0.05	0.81	0.011
	Premodificaiton	0.38	0.24	0.52	1.46	<0.001
	Initial stress of right collocate	−0.07	−0.2	0.07	0.94	0.323
Second-level	Number of syllables	4.41	3.71	5.11	82.02	<0.001
	The final elements					
	Final /i/	−1.24	-2.12	-0.36	0.29	0.006
	Final /l/	1.37	0.4	2.34	3.92	0.006
	Final /r/	1.5	0.57	2.43	4.49	0.002
	Final /li/	0.36	−0.71	1.43	1.43	0.513
	Final consonant cluster	1.1	0.53	1.67	3.01	<0.001
	Number of morphemes	−0.13	−0.74	0.48	0.88	0.685
	Final stress	1.89	1.04	2.75	6.64	<0.001
	Frequency of positive form (log)	−0.67	−0.81	−0.53	0.51	<0.001
	Comparative/positive ratio (logit)	−0.84	−1	−0.67	0.43	<0.001

As expected, the overall pattern of the fixed effects aligns with Hilpert (2008). However, it is noteworthy that many second-level predictors have either lost statistical significance or show markedly reduced coefficient estimates or odds ratios.

The number of morphemes and final /li/ sound are no longer significant predictors in the current model. In Hilpert's (2008) model, the number of morphemes had a relatively weak effect, despite reaching the 0.05 significance threshold ($\beta = 0.2$, $p = 0.032$). The current model completely nullifies this effect ($\beta = -0.13$, $p = 0.68$), consistent with Hilpert's (2008: 408) interpretation that the number of morphemes may have little to no effect on the alternation that is independent of syllable counts. There is no effect of the final /li/ sound in the current model ($\beta = 0.36$, $p = 0.51$). As discussed in the next section, *-ly* adjectives (phonemically /li/) exhibit considerable variation, a point also noted by Watanabe and Iyeiri (2020: 79). As hypothesized in Section 2.3, the significant bias of final /li/ toward the periphrastic variant observed in Hilpert (2008) is most likely driven by highly frequent adjectives like *likely*, which exhibits significant idiosyncrasies (see Figure 1 in the following section).

It is also worth noting that the effects of other second-level predictors are substantially reduced. As with the two variables above, this reduction is attributable to a key feature of the mixed-effects model, namely its ability to account for nested structure and lexically specific variability. First, the multilevel structure properly handles the nested nature of the dataset, preventing highly frequent adjectives, which may exhibit considerable skewness (see Section 4.2), from disproportionately influencing the results. Second, adjective-specific variability is now accounted for through the random effects, which removes variance that would otherwise have been attributed to the fixed effects. As a result, for example, the odds ratio for final /l/ is significantly reduced from 36.71 in Hilpert (2008) to 3.92 in the present study. This means that although the current model also finds effects for these variables, their effect sizes had been overestimated in Hilpert's (2008) model.

4.1.2 Random effects

In the mixed-model presented in (3), each adjective has its own intercept for the regression model. Importantly, this approach differs from fitting separate regression lines for each adjective (i.e., the "no-pooling approach"; Gelman and Hill 2007: Ch. 12). In a mixed-effects model, the random groups are assumed to originate from a shared population and are therefore expected to share certain characteristics (Bates 2010: Section 3.4). As a result, the intercepts are "shrunk" or drawn toward the global intercept (Gelman and Hill 2007; Winter 2019: Section 15.9). This shrinkage effect is more pronounced for groups with fewer observations, as there is less evidence to suggest that these groups deviate significantly from the overall tendency (Gelman and Hill 2007; Schäfer 2020).

Random effects can be analyzed by focusing on those that exhibit significant deviations from the overall tendency. Figure 1 displays the top 10 adjectives with the largest positive intercepts, while Figure 2 displays those with the largest negative intercepts. In other words, these figures illustrate adjectives that show either a preference for the periphrastic variant (Figure 1) or the morphological variant (Figure 2) to a degree that is unexpected based on the fixed effects alone. It is important to interpret the rankings with caution due to the shrinkage effect. The figures include 95 % prediction intervals for the intercepts, which are generally inversely related to the number of observations.[8]

We can now examine the adjectives shown in the figures and explore how their idiosyncrasies manifest themselves. Starting with Figure 1, the adjective that stands

8 Following Schäfer (2020), 95 % prediction intervals were constructed using the conditional variance-covariance matrices obtained through the ranef() function from the lme4 package (Bates et al. 2015).

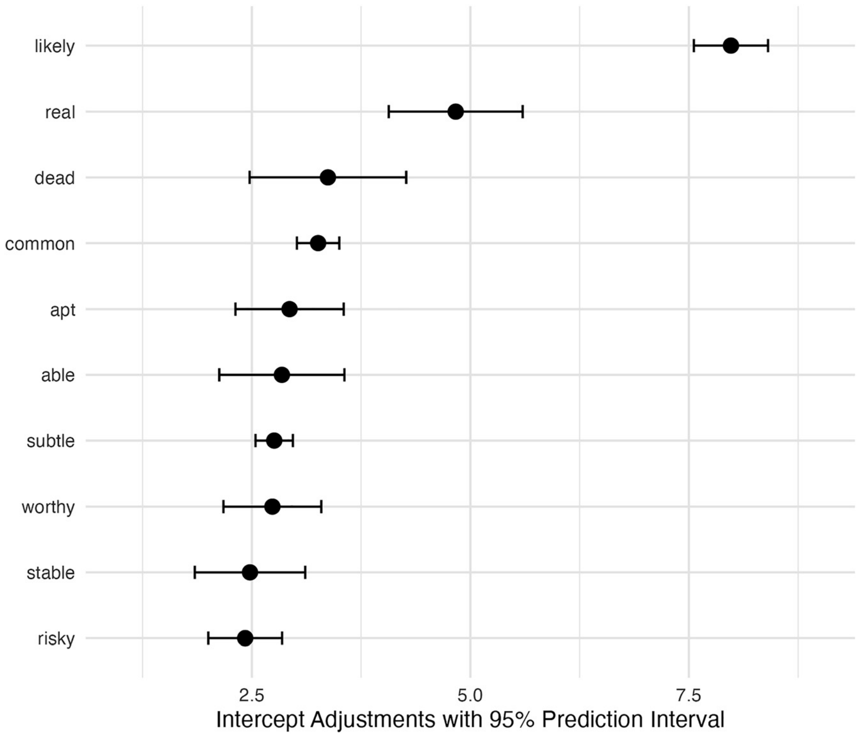


Figure 1: Top 10 adjectives with largest positive intercepts (idiosyncratic biases toward the periphrastic variant).

out is *likely*. The idiosyncrasies of this adjective have also been noted in previous research (e.g., Mondorf 2009: 41; Watanabe and Iyeiri 2020: 79). The average estimate for all comparative tokens of this adjective, without considering the random intercept (+7.98), is −2.68, indicating that the fixed effects alone predict a tendency toward the morphological variant. This is partly because a majority of disyllabic adjectives favor the morphological variant, and because final /li/ has no effect on the alternation in the current model. Thus, *likely*'s substantial bias toward the periphrastic form (3724:17) is remarkably unexpected.

Figure 1 also includes several monosyllabic adjectives, such as *real* (109:4), *dead* (19:4), and *apt* (36:13), whose preference for the periphrastic variant is unexpected based on the fixed effects. For *dead* and *real*, their infrequent use in comparisons, reflected in their very low comparative/positive ratios (0.002 and 0.005, respectively), partially explains their preference for the periphrastic variant. In the case of *apt*, the final consonant cluster contributes to this preference (Sweet 1968 [1891]: 326).

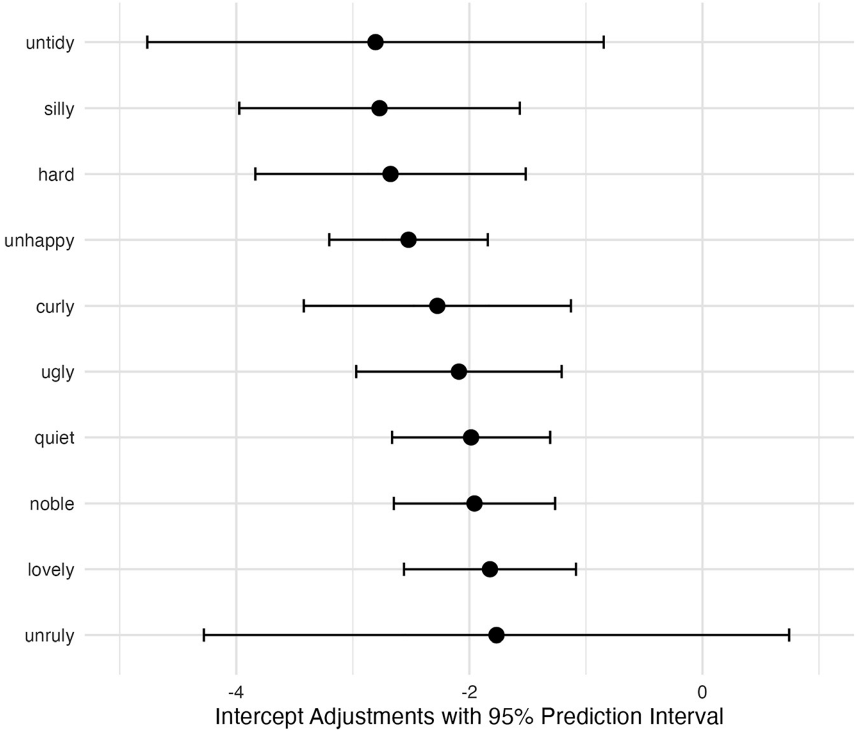


Figure 2: Top 10 adjectives with largest negative intercepts (idiosyncratic biases toward the morphological variant).

Nonetheless, since monosyllabic adjectives overwhelmingly favor the morphological variant, these monosyllabic adjectives require large positive adjustments to their intercepts to account for their unexpected behavior.

Turning to Figure 2, the majority of these adjectives are disyllabic, as the preference of monosyllabic adjectives for the morphological variant can largely be explained by the fixed effect (i.e., syllable count). However, the almost categorical distribution of the monosyllabic adjective *hard* (1:1745) requires a substantial adjustment to its intercept, particularly given that it ends with a consonant cluster, which favors the periphrastic form.[9] The disyllabic adjectives in Figure 2 – *silly*, *curly*,

9 One anonymous reviewer pointed out that in non-rhotic British English (the BNC variety), *hard* does not end in a consonant cluster. While we agree with this point, for the purpose of this paper, we decided to use the same adjective categories as Hilpert (2008). Otherwise, it would produce differences in the results that might obscure the aspects this study focuses on. However, it it advised that future research take this issue into account in their analysis.

ugly, *quiet*, *noble*, and *lovely* – substantially favor the morphological variant (see Appendix 1). Notably, half of the adjectives in the figure (*silly*, *curly*, *ugly*, *lovely*, and *unruly*) end with the final /li/ sound, despite the other *-ly* adjectives such as *costly*, *deadly*, or *likely*, which exhibit the opposite tendency. Additionally, trisyllabic adjectives containing the prefix *un-* such as *untidy*, *unhappy*, and *unruly* are described as "exceptions" by Quirk et al. (1985: 462). Their occurrence in the morphological form is particularly unexpected given their syllable counts.

Figures 1 and 2 indicate that the current model effectively accounts for lexically specific variability by incorporating random effects. The degree to which random effects explain the variance can be assessed using pseudo R^2 for logistic regression. Nakagawa and Schielzeth (2013) introduced a method to distinguish between marginal R^2 (variance explained solely by fixed effects) and conditional R^2 (variance explained by both fixed and random effects). In this case, the marginal R^2 was 0.49, while the conditional R^2 was 0.68, highlighting that a considerable proportion of variance is attributed to random effects.

4.1.3 Model performance

We are now examining whether the mixed-effects model offers a better fit for the dataset than the fixed-effects model. For this purpose, we independently constructed a fixed-effects model using the same predictors in (3), since the present study involved minor corrections to the annotations and a non-linear transformation of the frequency predictors. The prediction accuracy of the fixed-effects model was 96.1 %.[10] Prediction accuracy was calculated by dividing the number of correct predictions by the total number of predictions, using a probability threshold of 0.5 to determine the predicted outcome. Note that 89.7 % of the tokens in the dataset are in the morphological form, meaning the model must exceed this baseline to demonstrate meaningful accuracy. By comparison, the mixed-effects model achieves an improved prediction accuracy of 97.6 %.

Additionally, we are interested in comparing the performances between these models at the aggregated level for individual adjectives. Specifically, we aim to determine how each model predicts the probability of selecting the periphrastic variant for each adjective. Figure 3 presents a density plot showing the predicted proportions of choosing the periphrastic form for the alternating 247 adjectives, based on the fixed-effects model and the mixed-effects model. The actual proportion is represented by the (largely overplotted) red density curve. A density plot illustrates the concentration of data points (in this case, adjectives) across the dataset. For instance, we observe that in both the fixed-effects and mixed-effects models,

10 In Hilpert's (2008) model, the prediction accuracy was 95.8 %.

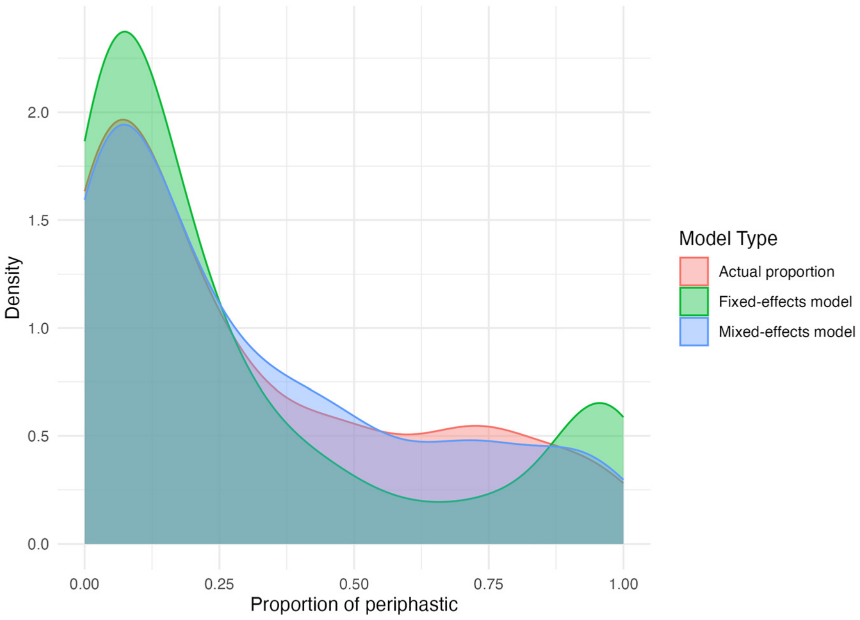


Figure 3: Density plot of predicted proportions of the periphrastic variant per adjective (n = 247).

adjectives are concentrated toward low proportions of the periphrastic form, which corresponds to high proportions of the morphological form. Additionally, we can observe that the curves slightly rise toward the high proportions of the periphrastic variant, creating a highly asymmetrical U-shape. This indicates that adjectives are generally skewed toward either variant.

The close alignment between the mixed-effects model's estimates and the actual proportions is expected, as this model incorporates the proportions of each adjective to calculate the random effects. What can be seen in the graph is that the fixed-effects model tends to overestimate the skew in proportions for both variants. In other words, the fixed-effects model predicts that adjectives are strongly biased toward one variant over the other, whereas in reality, the bias, particularly for the periphrastic form, is more nuanced.

Why does the fixed-effects model behave this way? It is important to reemphasize that the fixed-effects model does not 'know' that the data originates from 247 adjectives with varying frequencies (see Section 3.2). As a result, the model tends to generalize data from highly frequent adjectives to less frequent ones, since the former contribute the majority of data points. Therefore, the extreme skew predicted by the fixed-effects model reflects the tendency of highly frequent adjectives to exhibit a stronger bias toward one variant over the other.

Note that the fixed-effects model includes frequency variables, namely the logarithm of positive frequency and the logit transformation of comparative/positive ratio. However, these variables are not treated as adjective-dependent in the model's structure, as all variables are handled as observation-level predictors. If these frequency variables are excluded from the fixed-effects model, the skews become much more extreme. A corresponding figure is included in Appendix 2. This suggests that frequency plays a pivotal role in determining the strength of bias (skewness) in the present dataset. The following section discusses this issue in more detail.

4.2 Token frequency and strength of bias

In light of the results from Section 4.1.3, we systematically investigate the effects of frequency on the strength of bias for individual adjectives in the comparative alternation. To this end, the following mixed-effects model (4) was constructed. The purpose of this model is to examine the effect of frequency while controlling for other adjective-dependent factors that may also impact skewness.

(4) Response = {preferred, dispreferred}
Response ~ logFreq * Preferred Variant Type + SYL + FINAL + MOR
+ FINSTR + (1|ADJ)

This model predicts the probability of selecting the preferred variant for an adjective, whether it is the periphrastic or morphological form. Consequently, the predicted probability will always exceed 0.5. The frequency variable reflects the comparative forms (periphrastic + morphological) rather than the positive form, aligning with prior studies like Mondorf (2009). The frequency values were log-transformed to base *e* and subsequently centered. The interaction term with "Preferred Variant Type" (a factor variable with two levels: morphological and periphrastic) was included because frequency effects may influence skewness for one variant but not for the other. All second-level predictors from model (3) were included, except for the positive frequency and the comparative-to-positive ratio. The model did not encounter multicollinearity issues, with the highest corrected $GVIF^{1/2df}$ being 2.71 for the final stress (FINSTR).

Adjectives with no preference (i.e., a completely balanced distribution of variants) were excluded ($N = 9$). Additionally, adjectives with token frequencies less than six were removed ($N = 11$). This exclusion was necessary because of the unbalanced distribution of variants in the data set. The mean proportion of the preferred variant among adjectives was 0.84. Adjectives with fewer than or equal to six comparative tokens could not achieve this proportion (5/6 = 0.833) and were therefore excluded. As a result, only 227 adjectives were included for this analysis.

Table 3 presents the coefficient table for model (4). The effects of frequency are divided into two rows: one for when the preferred variant is morphological (the top row after the intercept) and the other for when it is periphrastic (the interaction term). In both cases, frequency is positively associated with the degree of bias (morphological: $\beta = 0.88$, $p < 0.001$; periphrastic: $\beta = -0.40$, $p < 0.001$). The coefficient for the latter case should be interpreted as a "slope adjustment term" (Winter 2019: 138), indicating that while the effect is weaker when the periphrastic variant is preferred, it remains positive (β = 0.88–0.40 = 0.48). Note that the preceding model (3) and previous research (e.g., Cheung and Zhang 2016; Hilpert 2008; Mondorf 2009), with the exception of D'Arcy (2014: 237), suggest positive effects of frequency (in positive or comparative forms) on choosing the morphological variant. Mondorf (2009) attributes this effect to the ease of cognitive processing resulting from the higher entrenchment of frequent forms, which she argues biases speakers toward using the morphological variant. While this explanation is certainly plausible and aligns with the results of model (3), it appears that once a preference is established for an adjective, repeated usage reinforces that bias regardless of the direction.

Table 3: Coefficient table for model (4).

Variables	**Coefficient estimates**	**CI low**	**CI high**	**Odds ratio**	***p* value**
(Intercept)	1.54	0.2	2.87	4.64	0.024
Log frequency of comparatives	0.88	0.78	0.97	2.41	< 0.001
Preferred variant (periphrastic)	−0.51	−0.91	−0.12	0.6	0.011
Frequency × variant	−0.4	−0.62	−0.17	0.67	< 0.001
Number of syllables	0.23	−0.41	0.87	1.26	0.476
The final elements					
Final /i/	0.48	−0.1	1.06	1.61	0.106
Final /l/	−0.14	−0.82	0.54	0.87	0.683
Final /r/	−0.37	−0.95	0.22	0.69	0.220
Final /li/	0.52	−0.18	1.23	1.69	0.144
Final consonant cluster	−0.15	−0.51	0.22	0.86	0.440
Number of morphemes	−0.14	−0.57	0.29	0.87	0.534
Final stress	1.12	0.44	1.79	3.06	0.001

There is also a significant effect for the preferred variant ($\beta = -0.51$, $p = 0.011$). Since the morphological variant serves as the reference level, the negative coefficient indicates that the degree of bias is stronger for the morphological variant.

For all second-level predictors except for final stress, no significant effects were observed. This is because their effects, if present, are unidirectional. For instance, final consonant clusters predict a bias toward the periphrastic variant. However, the current model predicts biases toward both variants, which counteracts and diminishes the impact of unidirectional effects. The significant effect of final stress ($\beta = 1.12$, $p = 0.001$) may appear puzzling, as it biases adjectives toward the periphrastic form (see Table 2). However, this can be attributed to an interaction between final stress and syllable count. Since monosyllabic adjectives automatically received positive values for final stress (Hilpert 2008: 400), the effects of this variable operate in two ways: one favoring the periphrastic variant (from the stress effect) and the other favoring the morphological variant (from monosyllabic adjectives). In the previous mixed-effects model (3), this dual effect was controlled by the significant effect of syllable count. In the current model, however, with no significant effect of syllable count, the dual effect persists. When the interaction term between final stress and syllable count is included, the main effect of final stress is no longer present ($\beta = 1.05$, $p = 0.206$), demonstrating that it is the interaction that explains the observed puzzling effect.

Figure 4 presents a scatterplot illustrating the relationship between the log frequency of comparative forms and the proportion of the periphrastic form. Several representative data points are labeled with the names of the corresponding adjectives, including *gentle* and *remote*, whose positions somewhat diverge from other adjectives. The plot highlights the bidirectional nature of frequency effects, showing a positive relationship between frequency and the strength of bias (skewness) for both directions of the bias. However, it should be kept in mind the dataset contains fewer adjectives that favor the periphrastic form ($N = 49$) compared to those favoring the morphological form ($N = 178$).

Additionally, it is noteworthy that the relationship between frequency and skewness exhibits a heteroskedastic pattern: there is substantial variability in the proportion among low-frequency adjectives, while high-frequency adjectives exhibit much less variability. For instance, on the low-frequency end, adjectives such as *sticky* and *risky* are relatively evenly distributed, whereas *gross* and *obscure* display considerable skewness. On the high-frequency end, adjectives are generally much less variable and skewed toward their preferred variant. This is evident from the near absence of data points in the middle-right section of the plot.

The robust variation reflects the probabilistic nature of the variation; adjectives may be skewed toward one variant or remain more balanced depending on the conditioning factors associated with the adjectives or the contexts in which they are

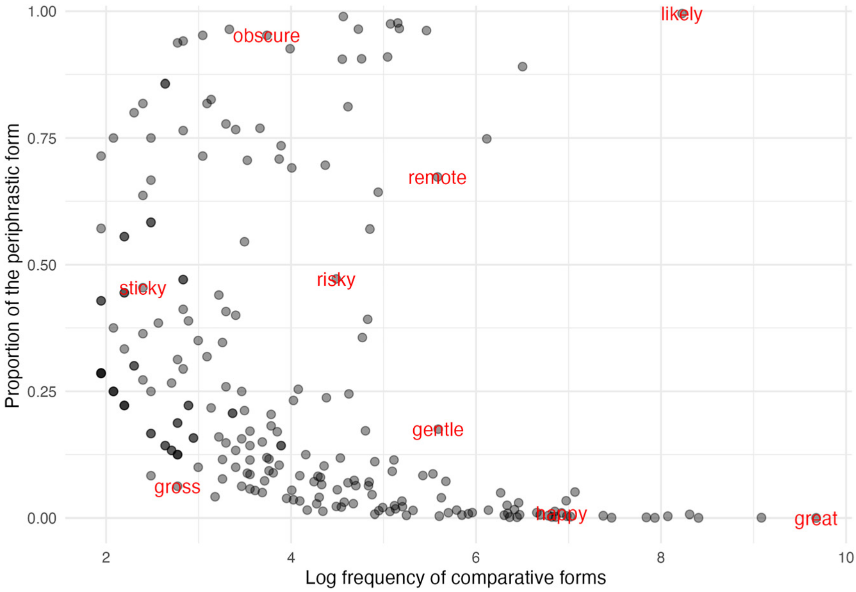


Figure 4: The relationship between log frequency and proportion of the periphrastic variant of 227 alternating adjectives in the BNC.

used. Conversely, on the high-frequency end, most adjectives show near-categorical distributions, including *happy*, *likely*, and *great*. The low variability within this group suggests that these adjectives escape probabilistic constraints and become almost exclusively associated with their preferred variant.[11]

It is worth comparing this result to that of Bermel and Knittl (2012), who report a substantial number of alternating nominal endings with a relatively balanced distribution in Czech (see Section 2.1). They find that noun groups exhibiting variation in the genitive singular and locative singular are, on average, highly frequent compared to groups that show no variation within the same grammatical paradigm (Bermel and Knittl 2012, Table 2). However, they do not provide frequency distributions within these variable groups, leaving it uncertain whether token frequency influences skewness among nouns capable of forming alternating endings.

11 There are also statistical aspects that contribute to the observed heteroskedasticity. Since lower frequencies mean a smaller number of observations, there is larger sampling error for less frequent adjectives. Note that this issue is taken into account in regression model (4), where adjectives with a smaller number of observations contribute less to the estimation of frequency effects. For technical details on how a mixed-effects (multi-level) model handels this issue, Raudenbush and Bryk (2002) Chapter 3 directly addresses it.

To summarize Section 4, the results demonstrate that individual adjectives in the BNC exhibit robust idiosyncrasies within the comparative alternation. While most fixed effects remain significant factors in the alternation, the strength of their effects is substantially qualified in the mixed-effects model. A closer examination reveals that individual adjectives show preferences that diverge from the general tendencies of the conditioning factors. The mixed-effects model can further account for the strength of bias in individual adjectives, which positively correlates with the token frequency of adjectives in comparative forms.

5 General discussion

In this section, we discuss our findings with regard to the broader question that we outlined in the introduction, which focuses on lexically specific constructions and their implications for the English comparative alternation.

The robust idiosyncrasies observed among adjectives align with the usage-based approach to language, where both individual comparative forms and their parent schema can be cognitively represented in a speaker's mind. The deviating patterns found across adjectives may indicate that speakers directly access individual comparative forms without relying on the probabilistic constraints of their parent schema. The degree to which speakers depend on concrete representations depends on the level of entrenchment, which is commonly operationalized through token frequency in usage-based linguistics (Bybee 2010).

However, it is important to note that the adjectives exhibiting the most divergent patterns (Figures 1 and 2) are not necessarily highly frequent. Nonetheless, one aspect not considered in this study is the clustering of adjectives. For instance, disyllabic adjectives ending in /li/ that favor the morphological variant (e.g., *silly, curly, ugly, lovely,* and *unruly*) in Figure 2 may form clusters separate from other /li/ adjectives that exhibit the opposite tendency. The strength of their deviating patterns may be attributed to the entrenchment of such clusters. This interpretation aligns with the view of Construction Grammar presented in Section 2, where linguistic generalizations based on clustering can emerge at any level of schematicity.

Furthermore, the observation that highly frequent comparative forms tend to be skewed toward one of the two variants also supports the idea of redundant representations in language. At the lexically specific level, variation is more idiosyncratic and less governed by factors that have predictable systematic effects. As discussed in Section 4.2, the extent to which this occurs is largely proportional to the token frequency of adjectives in comparative forms. We believe that this idiosyncrasy characterizes the English comparative alternation because retaining alternative variants for the same adjective offers little functional advantage in communication.

For example, *riskier* and *more risky* serve to express nearly identical meanings and functions. To illustrate, below are some examples of the adjective *risky* in comparative forms from the BNC, which exhibits a roughly equal distribution (42:47, Appendix 1).

(5) *Yes, but that would have become progressively **more risky** as people began to stir.* (fictional prose)

(6) *When the Antwerp market began to contract, he diversified into **more risky** ventures.* (biography)

(7) *Altering the diet is also far **more risky** for a child than it is for an adult, so there are more difficult decisions to be made before embarking on an elimination diet* (non-academic prose)

(8) *The second is that the institutions' loan portfolios are likely to grow **riskier**, should political imperatives substantially increase their exposure to former Soviet republics.* (periodicals)

(9) *Giving £10,000 credit for one month is always **riskier** than giving £5000 credit for one month.* (non-academic prose)

(10) *They are forced to work in isolation, even though it is **riskier** to do so.* (non-academic prose)

The choice of variants in these examples are influenced by the first-level predictors, such as attributive or predicative usage, and pragmatic factors that were not considered in the present study, such as emphasis (Mondorf 2009:112). For example, one could argue that the presence of premodification *far* in Example (7) biases the speaker to use the periphrastic variant. However, it is clear that, at least in most examples, the use of the alternative variant would have expressed the same intended message by the speaker. As a consequence, these forces can be overridden by the conventionality of a given variant, as evidenced by adjectives that exhibit extreme skews beyond what is explained by the probabilistic constraints at the schematic level. It is more probable that speakers choose *more risky* over *riskier* for additional emphasis than *more easy* over *easier* for the same purpose.

One important notion related to the strength of bias among adjectives is *statistical preemption* (Boyd and Goldberg 2011; Goldberg 2011) or simply *preemption* (Clark 1987). Statistical preemption explains how speakers learn the conventional variant over a competing variant based on skewness in the input. For example, if speakers systematically hear *explain this to me* when *explain me this* might have been expected, they learn that the former is the conventional form for expressing the intended message in context (Goldberg 2019). In the case of English comparatives, the same mechanism can explain how speakers acquire the conventional form of a given adjective beyond what is predicted by schematic patterns. However, the primary concern of this study is the skewness itself rather than its acquisition. In other words, how does skewness in the input arise in the first place?

The present study has demonstrated that skewness positively correlates with the log token frequency of adjectives in comparative forms. Note that this differs from the log token frequency of the preferred variant, which equates with *confidence* in Goldberg (2011:140), where the correlational relationship with skewness is expected by chance (i.e., the more frequent the preferred variant is, the greater the likelihood that it outweighs the frequency of the recessive variant). Although it is crucial to distinguish correlation from causation, our results suggest that the repeated use of a given adjective in comparison leads to the emergence of a lexically specific construction – a symbolic unit – that ultimately blocks the competing variant from being retained in usage. As suggested by Levshina and Lorenz (2022: 250–251), this effect is not independent of statistical preemption. However, we believe the relationship involves reflexive interactions (Schmid 2020). More specifically, individual leaners aiming to speak conventionally is what drives the emergence and maintenance of a convention, which in turn facilitates its further acquisition. Although this claim cannot be directly substantiated based on our data, we find that the frequency effects on skewness align most closely with this proposed view.

At the schematic level of English comparative constructions, our findings show variation that can be predicted based on multiple variables, though the effect sizes of these predictors are substantially reduced when incorporating individual adjectives as a random effect. In other words, it is not the case that each adjective behaves so differently in forming a comparative that no systematic patterns emerge from them. At the same time, the substantial idiosyncratic patterns found among adjectives and the frequency effects on skewness suggest that speakers primarily rely on probabilistic patterns at the schematic level for adjectives with less entrenched comparative forms or less entrenched clusters encompassing these forms. For example, it is likely that speakers rely on schematic patterns when forming comparatives for newly emerging adjectives, such as *cringy* (based on Google Ngram, accessed in February 2025), assuming that they have had limited exposure to these adjectives being used in comparative constructions. In contrast, for highly frequent adjectives that exhibit pronounced skews or other diverging patterns, speakers are expected to rely less on schematic patterns. This interpretation aligns with previous studies discussed in Section 2, which have demonstrated the importance of lower-level constructions in alternations (e.g., Perek 2014; Pijpops et al. 2021).

The discussions presented thus far relate to a recent theoretical debate surrounding isomorphism (Haiman 1980) or the Principle of No Synonymy (Goldberg 1995). Leclercq and Morin (2023) argue that the functionalist principle rejecting true synonyms does not preclude the existance of variation in language. In the comparative alternation, individual adjectives can and do occur in both variants, as

Examples (5) to (10) above demonstrate. However, as Leclercq and Morin (2023) emphasize, when concrete comparative forms are *constructions* – that is, entrenched and conventionalized symbolic units – they are subject to isomorphism, whereby one lexically specific variant becomes dominant over the other. In other words, although schematic variants do license different ways of expressing the same meaning in individual usage, the Principle of No Synonymy ensures that there are no true synonyms in the structure or system of language. It is worth stressing that this effect is gradient, as we have demonstrated through frequency effects. There is a cline of variability depending on the varying levels of entrenchment of individual comparative forms, as illustrated in Figure 4.

To conclude the discussion, in English comparative constructions, we largely observe systematic variation at the schematic level and more idiosyncratic variation, including minimal variation, at the lexically specific level. We suggest that the extent to which idiosyncratic effects accumulate depends on the level of entrenchment and conventionality of lexically specific constructions, in this case individual comparative forms.

6 Conclusions

Using a mixed-effects logistic regression model, this study reanalyzed Hilpert's (2008) data on the English comparative alternation. Particular attention was given to the results of the random effects, which revealed diverging patterns among adjectives and skewness toward either variant, with skewness being positively correlated with the token frequency of adjectives in comparative forms.

From a methodological perspective, we emphasize that a closer examination of random effects is valuable for identifying idiosyncratic patterns among random groups in linguistic alternations. From a theoretical perspective, our results support the view that competition occurs at multiple levels of schematicity in English comparative constructions, with systematic variation happening at the schematic level and more idiosyncratic variation at the lexically specific level.

It remains to be seen to what extent the present findings can be extended to other alternations in English or other languages. More specifically, frequency effects on skewness at the lexically-specific level may manifest themseleves differently depending on the type of alternation concerned. It is expected that the manifestation varies depending on various features, including the degree of semantic overlap between alternating constructions, the length or number of open slots, and the presence of polysemy, among others. All these factors can influence the nature of competition between linguistic variants, potentially leading to varying outcomes.

Overall, the theoretical arguments we have presented require the accumulation of empirical evidence, with each piece considered in relation to the specific characteristics of the data at hand. We hope that future empirical research will continue to build this body of corroborating or conflicting evidence and contribute to addressing these broader theoretical questions.

Acknowledgement: This study has received funding from the Swiss Government Excellence Scholarship No. 2023.0078.
Data availability: The full dataset and analysis code are available in an OSF repository: https://osf.io/xv9br/.

Appendix 1

(1). The 247 alternating adjectives in the BNC (Hilpert 2008: 414): *able* to *mad*

ADJ	*more*	*-er*	ADJ	*more*	*-er*	ADJ	*more*	*-er*
Able	170	6	Cosy	6	23	Glossy	3	13
Absurd	16	1	Crafty	3	3	Grand	6	175
Ample	9	2	Crazy	3	23	Great	1	15,936
Angry	42	76	Creamy	2	13	Grey	6	67
Apt	36	13	Crude	5	71	Grim	5	38
Big	1	4,466	Curly	2	12	Gross	1	15
Black	5	85	Daft	2	14	Guilty	12	2
Bland	5	11	Damp	2	54	Hairy	2	14
Bleak	4	23	Dark	5	802	Handsome	24	10
Blind	3	38	Dead	19	4	Happy	10	1,007
Bloody	2	10	Deadly	30	9	Hard	1	1,745
Blue	6	29	Deaf	2	7	Hardy	3	32
Blunt	5	30	Dense	19	147	Harsh	3	166
Bold	4	178	Diffuse	34	14	Hazy	2	1
Brash	2	38	Dire	5	2	Healthy	26	500
Brave	2	70	Dirty	1	64	Heavy	4	1,072
Brief	2	50	Dreary	2	2	Hefty	3	7
Bright	4	642	Dry	4	163	Holy	5	18
Brisk	2	33	Dull	2	58	Homely	13	4
Broad	7	1,588	Earthy	7	4	Hot	1	270
Brown	4	39	Easy	28	4,031	Humble	25	77
Bulky	7	26	Empty	4	14	Hungry	5	43
Bumpy	2	7	Faint	2	139	Idle	6	2
Busy	8	118	Fair	5	322	Ill	7	5
Canny	3	3	Feeble	7	15	Intense	169	4
Cheap	1	2,786	Fierce	11	82	Just	8	9
Cheery	1	11	Fine	1	627	Keen	15	120
Choosy	3	3	Fit	3	201	Kindly	4	7
Chunky	2	5	Fond	3	31	Lax	7	10
Classy	3	13	Frail	11	16	Leafy	2	10
Clear	36	1,032	Free	21	270	Lean	2	87
Clever	21	101	Fresh	3	144	Leggy	2	2
Close	3	2,539	Friendly	73	55	Lengthy	6	23

(continued)

ADJ	*more*	*-er*	ADJ	*more*	*-er*	ADJ	*more*	*-er*
Cloudy	8	4	Full	19	620	Light	2	953
Clumsy	5	12	Funky	3	3	Likely	3,724	17
Coarse	3	104	Funny	8	70	Lively	49	76
Cold	4	383	Fussy	4	2	Lofty	2	24
Comfy	5	4	Gentle	47	221	Lonely	7	13
Common	594	73	Ghostly	3	1	Lovely	7	42
Compact	18	4	Glad	5	8	Lowly	2	6
Corrupt	5	1	Glitzy	2	1	Lucky	5	65
Costly	82	19	Gloomy	7	42	Mad	3	71

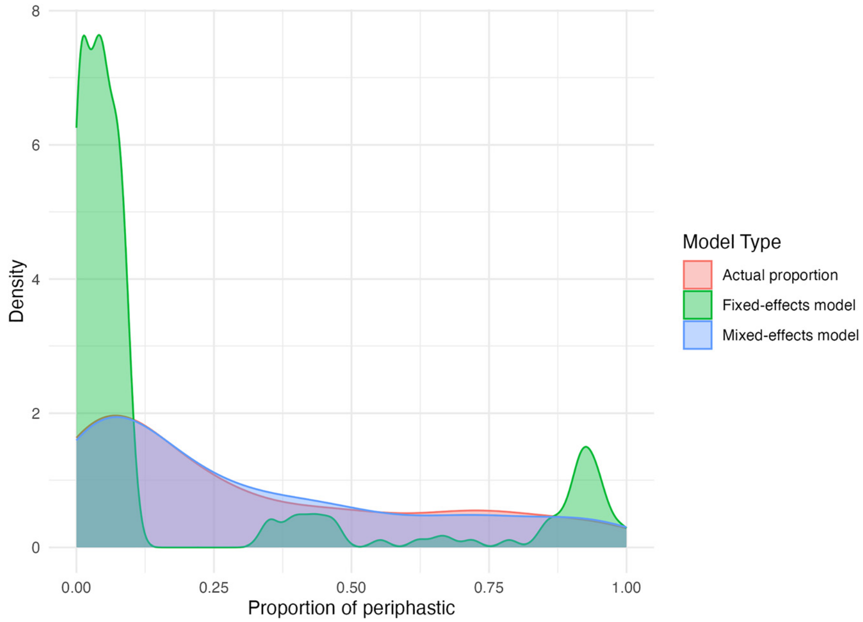
(2). The 247 alternating adjectives in the BNC (Hilpert 2008: 415): *manly* to *yellow*.

ADJ	*more*	*-er*	ADJ	*more*	*-er*	ADJ	*more*	*-er*
Manly	4	1	Rosy	4	11	Stormy	3	5
Mature	141	14	Rough	7	103	Straight	2	92
Mellow	11	14	Round	3	52	Strange	19	208
Messy	3	27	Rowdy	2	5	Strict	11	266
Mighty	2	30	Rude	4	26	Strong	11	3,194
Moist	3	6	Sad	7	94	Sturdy	2	18
Murky	2	5	Safe	12	933	Subtle	339	114
Narrow	14	550	Salty	4	5	Supple	15	1
Nasty	1	76	Scarce	5	55	Sure	19	61
Naughty	2	5	Scary	7	11	Sweet	2	157
New	10	601	Secure	156	4	Tasty	3	16
Noble	9	35	Severe	227	9	Tender	15	6
Noisy	6	34	Sexy	5	37	Tense	9	3
Obscure	40	2	Shadowy	8	2	Thin	3	544
Odd	8	56	Shaky	3	8	Tidy	8	36
Oily	3	7	Shallow	6	125	Tight	1	579
Pale	3	296	Sharp	3	368	Tricky	13	43
Patchy	3	2	Shrill	2	13	Ugly	4	41
Petty	2	4	Shy	2	12	Unhappy	18	15
Pleasant	90	50	Silly	1	23	Unlikely	27	1
Polite	23	7	Simple	60	1,115	Unruly	3	1
Poor	3	1,115	Sincere	12	2	Untidy	2	2
Profound	106	11	Sleepy	7	7	Vague	8	39
Proud	4	14	Slender	15	44	Vain	4	3
Pure	3	94	Slight	2	35	Vast	4	21
Quick	8	770	Small	5	8,816	Vile	3	3
Quiet	7	454	Sober	50	4	Warm	2	344
Quirky	3	4	Soft	5	563	Wary	21	6
Racy	2	6	Sorry	12	18	Weak	2	913
Rare	22	231	Sour	3	4	Wealthy	15	148
Raw	5	4	Sparse	7	20	Weary	3	9

(continued)

ADJ	*more*	*-er*	ADJ	*more*	*-er*	ADJ	*more*	*-er*
Ready	55	24	Speedy	6	69	Weighty	8	24
Real	109	4	Spicy	2	7	Weird	5	27
Red	8	100	Spooky	2	1	Wet	1	134
Remote	179	87	Sporty	2	14	White	9	118
Rich	4	877	Sprightly	2	1	Wild	1	189
Right	7	5	Stable	86	9	Worldly	20	1
Risky	42	47	Stale	2	6	Worthy	38	17
Robust	95	1	Stark	9	17	Yellow	8	9
Rocky	3	16	Steady	4	31			
Roomy	4	5	Sticky	5	6			

Appendix 2: Density plot of predicted proportions per adjective using the fixed-effects model without frequency variables (cf. Figure 3)



References

Arnon, Inbal & Neal Snider. 2010. More than words: Frequency effects for multi-word phrases. *Journal of Memory & Language* 62(1). 67–82.

Bates, Douglas. 2010. *Lme4: Mixed-effects modeling with R*. New York: Springer.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Bermel, Neil & Luděk Knittl. 2012. Morphosyntactic variation and syntactic constructions in Czech nominal declension: Corpus frequency and native-speaker judgments. *Russian Linguistics* 36(1). 91–119.

Biber, Douglas & Susan Conrad. 2019. *Register, genre, and style*. Cambridge: Cambridge University Press.

Booij, Geert E. 2013. Morphology in construction grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of construction grammar*, 255–273. Oxford: Oxford University Press.

Boyd, Jeremy K. & Adele E. Goldberg. 2011. Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language* 87(1). 55–83.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald R. Baayen. 2007. Predicting the dative alternation. In Gerlof Bouma, Irene Kraemer & Joost Zwarts (eds.), *Cognitive foundations of interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.

Broccias, Cristiano. 2001. Allative and ablative at-constructions. In Mary Adronis, Christopher Ball, Elston Heide & Sylvain Neuvel (eds.), *CLS 37: The main session. Papers from the 37th meeting of the Chicago linguistic society*, 67–82. Chicago: Chicago Linguistic Society.

Bürkner, Paul-Christian. 2017. Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). 1–28.

Bybee, Joan L. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Cheung, Lawrence & Longtu Zhang. 2016. Determinants of the synthetic–analytic variation across English comparatives and superlatives. *English Language and Linguistics* 20(3). 559–583.

Chua, Deborah. 2018. Understanding comparative alternation in *y*-adjectives: What else might we need? *Journal of Linguistics* 54(3). 459–491.

Clark, Eve. 1987. The principle of contrast: A constraint on language acquisition. In Brian MacWhinney (ed.), *Mechanisms of language acquisition*, 1–33. Hillsdale: Lawrence Erlbaum Associates.

Croft, William. 2022. *Morphosyntax: Constructions of the world's languages*. Cambridge: Cambridge University Press.

Dąbrowska, Ewa. 2009. Words as constructions. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New directions in cognitive linguistics*, 201–223. Amsterdam: John Benjamins.

D'Arcy, Alexandra. 2014. Functional partitioning and possible limits on variability: A view of adjective comparison from the vernacular. *Journal of English Linguistics* 42(3). 218–244.

De Smet, Hendrik, Frauke D'Hoedt, Lauren Fonteyn & Kristel Van Goethem. 2018. The changing functions of competing forms: Attraction and differentiation. *Cognitive Linguistics* 29(2). 197–234.

Diessel, Holger. 2016. Frequency and lexical specificity in grammar: A critical review. In Heike Behrens & Stefan Pfänder (eds.), *Experience counts: Frequency effects in language*, 209–238. Berlin, Boston: Mouton de Gruyter.

Diessel, Holger. 2023. *The constructicon: Taxonomies and networks*. Cambridge: Cambridge University Press.

Ferreira, Fernanda & Nikole Patson. 2007. The "good enough" approach to language comprehension. *Language & Linguistics Compass* 1. 71–83.

Gelman, Andrew & Jennifer Hill. 2007. *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

Goldberg, Adele. E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
Goldberg, Adele E. 2011. Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics* 22(1). 131–153.
Goldberg, Adele E. 2019. *Explain me this: Creativity, competition, and the partial productivity of constructions*. Princeton: Princeton University Press.
Gries, Stefan Th. 2003. *Multifactorial analysis in corpus linguistics: A study of particle placement*. London: Continuum.
Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collostructional analysis: A corpus-based perspective on "alternations". *International Journal of Corpus Linguistics* 9(1). 97–129.
Haiman, John. 1980. The iconicity of grammar: Isomorphism and motivation. *Language* 56(3). 515–540.
Hilpert, Martin. 2008. The English comparative - Language structure and use. *English Language & Linguistics* 12(3). 395–417.
Hilpert, Martin. 2021. *Ten lectures on diachronic construction grammar*. Leiden: Brill.
Hotta, Hikaru. 2025. Frequency does not predict the processing speed of multi-morpheme sequences in Japanese. *Linguistics Vanguard*. https://doi.org/10.1515/lingvan-2024-0083.
Jolsvai, Hajnal, Stewart M. McCauley & Morten H. Christiansen. 2020. Meaningfulness beats frequency in multiword chunk processing. *Cognitive Science* 44(10). https://doi.org/10.1111/cogs.12885.
Katz, Leonard, Karl Rexer & Georgije Lukatela. 1991. The processing of inflected words. *Psychological Research* 53(1). 25–32.
Kytö, Merja & Suzanne Romaine. 1997. Competing forms of adjective comparison in modern English: What could be more quicker and easier and more effective? In Terttu Nevalainen & Leena Kahlas-Tarkka (eds.), *To explain the present: Studies in the changing English language in honour of matti rissanen*, 329–352. Helsinki: Société Néophilologique.
Langacker, Ronald W. 1987. *Foundations of cognitive grammar: vol. I: Theoretical prerequisites*. Stanford: Stanford University Press.
Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow & Suzanne Kemmer (eds.), *Usage-based models of language*, 24–63. Stanford: Stanford University Press.
Leclercq, Benoît & Cameron Morin. 2023. No equivalence: A new principle of no synonymy. *Constructions* 15(1). https://doi.org/10.24338/cons-535.
Lehmann, Hans Martin & Gerold Schneider. 2012. Syntactic variation and lexical preference in the dative-shift alternation. In Joybrato Mukherjee & Magnus Huber (eds.), *Corpus linguistics and variation in English: Theory and description*, 65–75. Amsterdam: Rodopi.
Levshina, Natalia & David Lorenz. 2022. Communicative efficiency and the principle of No synonymy: Predictability effects and the variation of *want to* and *wanna*. *Language & Cognition* 14(2). 249–274.
Lindquist, Hans. 1998. The comparison of English disyllabic adjectives in *-y* and *-ly* in present-day British and American English. In Hans Lindquist (ed.), *The major varieties of English: Papers from MAVEN 97*, 205–212. Växjö: Växjö University.
Mondorf, Britta. 2003. Support for more-support. In Günter Rohdenburg & Britta Mondorf (eds.), *Determinants of grammatical variation in English (topics in English linguistics 43)*, 251–304. Berlin & New York: Mouton de Gruyter.
Mondorf, Britta. 2009. *More support for more-support: The role of processing constraints on the choice between synthetic and analytic comparative forms*. Amsterdam/Philadelphia: John Benjamins.
Nakagawa, Shinichi & Holger Schielzeth. 2013. A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology & Evolution* 4(2). 133–142.
Nicenboim, Bruno, J. Schad Daniel & Vasishth Shravan. 2025. Introduction to bayesian data analysis for cognitive science. https://bruno.nicenboim.me/bayescogsci/ (accessed 5 August 2025).

Pawley, Andrew & Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In Jack C. Richards & Richard W. Schmidt (eds.), *Language and communication*, 191–226. London: Longman.

Perek, Florent. 2014. Rethinking constructional polysemy: The case of the English conative construction. In Dylan Glynn & Jus Robinson (eds.), *Polysemy and synonymy: Corpus methods and applications in cognitive linguistics*, 61–85. Amsterdam/Philadelphia: John Benjamins.

Perek, Florent. 2015. *Argument structure in usage-based construction grammar: Experimental and corpus-based perspectives*. Amsterdam/Philadelphia: John Benjamins.

Pijpops, Dirk. 2020. What is an alternation? Six answers. *Belgian Journal of Linguistics* 34(1). 283–294.

Pijpops, Dirk & Freek Van de Velde. 2016. Constructional contamination: How does it work and how do we measure it? *Folia Linguistica* 50(2). 543–581.

Pijpops, Dirk, Dirk Speelman, Freek Van de Velde & Grondelaers Stefan. 2021. Incorporating the multi-level nature of the constructicon into hypothesis testing. *Cognitive Linguistics* 32(3). 487–528.

Quirk, Randolph, Sidney Greenbaum & Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. New York: Longman.

Raudenbush, Stephen W. & Anthony S. Bryk. 2002. *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

R Core Team. 2024. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rohdenburg, Günter. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics* 7(2). 149–182.

Röthlisberger, Melanie, Jason Grafmiller & Benedikt Szmrecsanyi. 2017. Cognitive indigenization effects in the English dative alternation. *Cognitive Linguistics* 28(4). 673–710.

Schäfer, Roland. 2020. Mixed-effects regression modeling. In Magali Paquot & Stefan Th. Gries (eds.), *A practical handbook of corpus linguistics*, 535–561. Cham: Springer.

Schmid, Hans-Jörg. 2020. *The dynamics of the linguistic system: Usage, conventionalization, and entrenchment*. Oxford: Oxford University Press.

Sereno, Joan A. & Allard Jongman. 1997. Processing of English inflectional morphology. *Memory & Cognition* 25(4). 425–437.

Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.

Sweet, Henry. 1968 [1891]. *A new English grammar: Logical and historical*. Oxford: Clarendon Press.

Szmrecsanyi, Benedikt. 2005. Language users as creatures of habit: A corpus-linguistic analysis of persistence in spoken English. *Corpus Linguistics & Linguistic Theory* 1(1). 113–150.

Szmrecsanyi, Benedikt. 2010. The English genitive alternation in a cognitive sociolinguistics perspective. In Dirk Geeraerts, Gitte Kristiansen & Yves Peirsman (eds.), *Advances in cognitive sociolinguistics*, 141–166. Berlin & New York: De Gruyter Mouton.

Thornton, Anna. 2019. Overabundance in morphology. *Oxford Research Encyclopedia of Linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.554.

Watanabe, Ayano & Yoko Iyeiri. 2020. Explaining the variability of adjective comparatives and superlatives: Entering the twenty-first century. *Word* 66(2). 71–97.

Winter, Bodo. 2019. *Statistics for linguists: An introduction using R*. New York: Routledge.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Zuur, Alain F., Elena N. Ieno & Chris S. Elphick. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1(1). 3–14.