Univerzita Karlova

Filozofická fakulta

Ústav Anglického jazyka a didaktiky

Bakalářská práce

Gleb Vorobev

Alternace syntetického a analytického komparativu v mluveném anglickém jazyce

Alternation of synthetic and analytic comparative in spoken English

Praha 2024

Vedoucí práce: Mgr. Kateřina Vašků, Ph.D.

Poděkování:

Chtěl bych hluboce poděkovat vedoucí své práce Mgr. Kateřině Vašků, Ph.D. za její pomoc ve všech etapách teto práce, za vysvětlení postupu práce se statistickými testy, a nejvíc za její trpělivost a pochopení.

Prohlášení:

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně, že jsem řádně citoval všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 5. srpna 2024

Gleb Vorobev

Abstrakt

Mnoho anglických adjektiv mohou vytvářet komparativ dvěma různými formy: synteticky se sufixem *-er*, nebo analyticky s *more*. Tato bakalářská práce se zaměřuje na determinanty, které ovlivňují (ovlivňující?) výběr jedné ze dvou forem v hovorové angličtině. V odborné literatuře bylo odhaleno mnoho proměnných, které mají vliv na tu volbu: teoretická část shrnuje ty, které lze zkoumat v korpusu, a ukazuje různé přístupy a zjištěni o jejich dopadu a významnosti. Všechny z nich měli vliv v písemném jazyce – proto cíli teto práce je porovnat distribuce komparativních forem a účinek proměnných, které jí ovlivňují, mezi dvěma korpusy.

Analytická část je provedena na dvou vzorcích – jeden o 612 tokenů pro analýzu interních, a jeden o 216 tokenů pro analýzu externích proměnných. Oba jsou sestavené ze stejných 53 typů. Všechna data jsou vzata ze mluvené časti korpusu BNC2014. Deset proměnných a jejich vliv na vytvoření komparativu jsou zkoumané v datech. Materiál ze mluveného korpusu je porovnáván se zjištěními ze písemného BNC aby zjistit rozdíl ve variabilitě.

Klíčová slova: alternace komparativů, fonologické proměnné, komparativní alternace, mluvený jazyk, morfologické proměnné, srovnání adjektiv, syntaktické proměnné

Abstract

Many adjectives in the English language can take either of the two comparative forms: a synthetic one with the suffix -er, or the analytic one with more. This thesis focuses on the determinants that influence the choice of one or the other in spoken English. Many variables that have an effect on the choice have been discovered in academic literature: the theoretical part takes stock of those that can be studied in a corpus, and shows different approaches and findings on their impact and significance. All of them were found to have an effect in the written language – making the aim of this thesis to contrast the comparative distribution and the impact of the variables influencing it between the two corpora.

The analytical part is conducted on two samples – one of 612 tokens for the analysis of internal, and one of 216 tokens for external variables. Both consist of the same 53 types. All data is taken from the spoken BNC2014. Ten variables and their influence on the comparative formation are examined in the data. The material from the spoken corpus is compared to the findings from the written BNC to also establish the difference in variability.

Keywords: adjective comparison, spoken language, comparative alternation, phonological variables, morphological variables, syntactic variables

Table of Contents

List of abbreviations

BNC	the British National Corpus
cf.	*confer* (Latin), compare
CGEL	a Comprehensive Grammar of the English Language
CQP	Corpus Query Processor
ibid	*ibidem*, in the same place
NZE	New Zealand English
OED	Oxford English Dictionary
ONZE	the Origins of New Zealand English corpus
Spoken BNC2014	British National Corpus 2014, spoken part
TEA	the Toronto English Archive corpus
WWE	the Wellington Corpus of Written New Zealand English

List of tables and figures

1. Introduction

This bachelor's thesis studies the variation of comparative forms of adjectives in contemporary spoken language. An adjective can form its comparative in two ways: either by adding the suffix *-er*, as in *prouder*, or the adverb *more*, as in *more proud*: the choice is based on the joint influence of a number of determinants. Comparative alternation, as well as the variables that govern it, is not a new topic of discussion: already in the CGEL, a grammar book from 1985, it is established that the choice of comparative is governed by the adjective's length in syllables. It also lists endings that identify adjectives that can "most readily take inflected forms" (Quirk et al., 1985, p. 461-462). Since then, many synchronic studies have been conducted to identify the determinants governing the alternation, as well as discover the exact influence each one has on the comparative forms' distribution (Mondorf 2002, 2003, 2009; Hilpert 2008; Matsui 2010; Cheung and Zhang 2016; Watanabe and Iyeiri 2020). However, the majority of the research was conducted in the domain of written language: there seem to be few works that focus on the comparative distribution and its determinants in spoken language (D'Arcy 2012, 2014). Even then, in her more recent study, D'Arcy finds the two sociolinguistic corpora she examines to be almost completely invariable: only four adjectives are variable in their comparative in one (2014, p. 226), and eight in the other (ibid., p. 234). It seems that this register is ripe for more research that would either confirm or falsify her findings, as well as examine the potential influence of the determinants of comparative alternation established in the written language on the patterning of the two forms in the spoken one.

The theoretical part of this thesis firstly establishes the definition of the adjective and distinguishes its features from those of an adverb. It briefly describes all the determinants of comparative alternation that have been heretofore attested in the academic literature, and then focuses on the effect – or lack thereof – of every variable that can be analyzed in corpus-based research. It provides multiple, oftentimes contested, findings and theories from scholarly works, both for written and spoken language, and finally features a section describing the proposed reasons for the comparative variation in language.

The practical part analyses two samples: a wider one, consisting of 612 adjectives of 53 types, which is used to gauge the influence of internal variables on the comparative patterning of adjectives, and a smaller, manually assembled one, consisting of 216 tokens of the same types. The smaller sample is used for the analysis of the influence of the external variables on the alternation. It also compares the overall variability in the spoken corpus with that in the written

BNC. The empirical part aims to replicate the findings in literature on the effect of 10 determinants of comparative distribution in the written register, but in the spoken language.

2. Theoretical part

2.1. Defining the adjective

Adjectives are used to "describe qualities of people, things and states of affairs" (Biber et al., 1999, p. 64): their "most characteristic function is to modify nouns" (Huddleston & Pullum, 2002, p. 527). It is not usually possible to state that the word is an adjective simply based on its form, since while certain suffixes do only occur with adjectives, like *-able* in *comfortable*, many of them, such as *young*, do not have an identifying form (Quirk, 1985, p. 402), and many suffixes are not exclusive to adjectives: an example is the suffix *-al*, which can occur both in adjectives and nouns derived from a verb: *abysmal – survival* (Dušková 2.1.). On a morphological level, adjectives can be inflected for comparison, however, many adjectives, like *disastrous* do not allow grading by attaching the suffix. Some adverbs also have similar inflected forms, making it impossible to judge if a word is an adjective by its gradability alone, for instance:

(1) (He worked) *hard - harder – hardest* (Quirk, 1985, p. 402)

On a syntactic level, an adjective can be the head of the adjective phrase, or the noun phrase in special cases like (2):

(2) *Of course he was rich, but **the rich** were usually mean* (Biber et al.,1999, p. 64).

Quirk et al. (1985, pp. 402-403) posit four criteria that define a word as an adjective:

a. The possibility for it to occur in an attributive function, premodifying the noun and positioned between it and a determiner (including zero article), as in *an **ugly** painting*, *the **round** table*, ***dirty** linen.*
b. The possibility for it to occur in a predicative function, as either subject complement, as in (3), or object complement, as in (4):
   (3) *The painting is **ugly**.*
   (4) *He thought the painting **ugly**.*
c. An adjective should be able to be premodified by the intensifier *very*:
   (5) *The children are very **happy**.*
d. Being able to take comparative and superlative forms, either synthetic ones by adding the endings *-er* and *-est*, or periphrastic ones using the premodifiers *more* and *most*:
   (6) *The children are **happier** now.*
   (7) *They are the **happiest** people I know.*
   (8) *These students are **more intelligent**.*
   (9) *They are the **most beautiful** paintings I have ever seen.*

Biber et al. additionally consider the adjective's semantic descriptive role in their discussion of its defining features (1999, p. 506), and Huddleston and Pullum establish another syntactic function pertaining to the adjective, namely its ability to take the postpositive – or postnominal, the term this paper uses – position, as in *someone* ***happy***. (2012, p. 528).

Nevertheless, all three grammars agree that not every adjective fits all the criteria. Two groups – central and peripheral adjectives – are distinguished by both Quirk et al. (1985, p. 404) and Biber et al. (1999, p. 527), but the latter consider an adjective peripheral when it lacks at least one of the characteristics, while the former only regard attributive and predicative function as determinants of central adjectives. Those words that only possess one, both or neither of the last two features are considered adverbs (Quirk et al., 1985, p. 404).

2.2. Choice of comparative

In English, three forms exist for comparison of higher degree: absolute (or positive), comparative and superlative. The positive form is the same as the base form of the adjective. The comparative and superlative forms can be realized either morphologically by adding the *-er* or *-est* suffix respectively, as in (10), or synthetically by using the adverbs *more* and *most*, as in (11):

(10) *high-high**er**-high**est***

(11) *complex-**more** complex-**most** complex* (Quirk, 1985, p. 458)

Some adjectives form their comparative and superlative forms irregularly: some adjectives with high frequency like *good* or *bad* have stems in these forms which are different from the base: *better-best* and *worse-worst* respectively (ibid., pp. 458-459). Such suppletive forms preclude the formation of periphrastic comparative in the corresponding adjectives: a form such as *more good* is impossible because of the existence of *better* (Hilpert, 2008, p. 398). They are therefore irrelevant to this study.

It is part of the general scholarly consensus (Quirk, 1985, pp. 461-462) that monosyllabic adjectives mostly take inflectional comparative and superlative, as in *low-lower-lowest,* while the adjectives with three and more syllables, like *beautiful*, mostly form comparison periphrastically: *more beautiful – most beautiful*. Adjectives with four or more syllables, except for some particular cases (*cf.* Section 2.2.2), are invariable and always form their comparison analytically. The most variation happens in the field of disyllabic adjectives, like with *polite*, which can either have *politer* or *more polite* as its comparative form. Throughout the years, scholarly research has

discovered many determinants of comparative alternation, both in the disyllabic adjectives and, in special cases, mono- and trisyllabic ones as well – in the fields of phonology, morphology, syntax, language use, lexicon, semantics and pragmatics (*cf.* Hilpert, 2008, Mondorf, 2009), as well as sociolinguistic factors such as formality (*cf.* Watanabe and Iyeiri, 2020). For the sake of this study, the most attention will be paid to the phonological, morphological and syntactic determinants, as well as those of positive frequency and the positive-comparative ratio in the domain of language use – the areas that have received extensive enough treatment in the academic literature and that are the most suitable for corpus-based research.

2.2.1. Phonological determinants

The phonological determinants of the comparative distribution in adjectives (aside from the syllable quantity) that are discussed in literature are their final segment, the possibility of stress clash, as well as avoidance of identical adjacent elements. As the most variation happens in disyllables, the effect of the variables is both the most pronounced and the most studied in this group. Nevertheless, some studies also find determinants of variation for mono- and trisyllabic adjectives, as well as a number of particular exceptions. Some scholars also seem to dispute a number of the findings in the previous studies.

Disyllabic adjectives take the synthetic form more readily if ending in [l̩], [ə^{r}], or an unstressed vowel. Quirk et al. group such adjectives according to their endings:

-*y*: early, easy, funny, etc.

-*ow*: mellow, narrow, shallow, etc.

-*le*: able, feeble, gentle, etc.

-*er*, -*ure*: clever, mature, obscure, etc. (Quirk et al., 1985, p. 462)

However, it is possible to further divide the -*y* group into adjectives ending in -*ly* (*/li/*) and those ending in -*y* (*/i/*). While the adjectives ending in */i/* form synthetic comparison more often than analytic, the */li/* adjectives usually occur with periphrastic comparison, as is the case with *lively*:

lively { *livelier-liveliest*
more lively-most lively (ibid., 462)

Analyses by Hilpert (2008, p. 409) and Matsui (2010, pp. 191,195) confirm this, with the findings indicating an increase in periphrastic comparison in the adjectives with the */li/* ending, while showing an affinity of the adjectives ending in */i/* for forming synthetic comparison.

In her 2003 paper, Mondorf performs a quantitative analysis of select disyllabic adjectives ending in *-er* and *-re* (*/r/*), the results of which show that in fact almost none – *slender* being the exception with 40% – of the 13 adjectives examined form synthetic comparison more than 20% of the time. In fact, at least ten of the adjectives do not form the inflectional comparative even in 10% of the instances (Mondorf, 2003, pp. 280-281). Such behavior is explained in terms of "avoidance of identity effects", specifically the phenomenon of haplology: the disinclination to use "identical and adjacent [morphological] elements or structures" (ibid., 2003, p. 278). The unpleasant *-rer* ending in the synthetic comparative of these adjectives is therefore avoided with the use of analytic comparison (ibid., 2003, p. 279). In Mondorf's explanation of another case of such avoidance she also posits that inflectional comparison can serve as a buffer to avoid a stress clash. Since, as she explains, most nouns in English are stressed on the first syllable according to the Germanic pattern, in a situation where a finally-stressed adjective is followed by a noun stressed on its first syllable, as in *a móre próud cándidate*, synthetic comparison may be chosen instead of the periphrastic option to avoid three subsequent stresses, like in *a próuder cándidate* (Mondorf, 2009, pp. 17-18). By ascertaining the amount of times the use of the analytic comparative would have led to a stress clash for a range of finally-stressed adjectives, Mondorf finds the use of synthetic comparative for such a purpose more common in the attributive position (Mondorf, 2009, p. 20). Nevertheless, both groups of adjectives she examines overwhelmingly prefer the analytic comparison in any position (Mondorf, 2009, p. 22). To this finding, Hilpert (2008, p. 400) adds that a stress clash is possible in other syntactic contexts as well, as in (12) or (13):

(12) *It's cóoler nów than it was last week.*

(13) *The drúnker Róbert gets, the more he likes Jessica.*

Another phonological factor influencing the choice of comparative Mondorf discusses is minimally distinct consonant clusters – namely the choice of analytic comparison in adjectives with endings such as *-pt*, *-kt* and *-ct*. In her study of 26 mono- and disyllabic adjectives ending in a consonant cluster, she finds that all 21 of the disyllabic ones exclusively form the periphrastic comparative (Mondorf, 2009, pp. 31-32). Hilpert's examination of this determinant, which he however extends to include all consonant clusters, confirms it to be a factor in adjectives forming analytic comparison (Hilpert, 2008, p. 407). Dušková (2009, 6.82) also partially confirms the influence of this factor, stating that adjectives ending in *-ct* and *-nt*, such as *distinct* and *recent*,

most often form the analytic comparative. She adds that adjectives with an uncommon phonological or morphological structure, such as *antique*, also usually form their comparative periphrastically.

Despite mostly forming the comparative analytically, trisyllabic adjectives can take *-er* in special cases: Biber et al. (1999, p. 522) claim that trisyllabic adjectives ending in *-y* can occasionally form synthetic comparison, providing an example in the forms *almightiest* and *unhappiest*. Nevertheless, *unhappiest* could instead fall under the rule formulated by Quirk et al. (1985, p. 462), according to which trisyllabic adjectives with a negative *un-* prefix, such as *unhappy* or *untidy*, are an exception and can form inflected comparatives and superlatives. Interestingly though, both examples provided by Quirk et al. also have the *-y* ending. Additionally, Mondorf (2003, p. 257) stresses a caveat in the generally accepted rule: the criterion influencing the formation of either comparative seems to not be the amount of syllables in the positive form of the adjective, but instead in its anticipated form in the synthetic comparative. As an example, she provides the trisyllabic *sensible*, which takes an inflected form as well as the periphrastic one due to the elision of a syllable in the resulting synthetic form *sensibler*. For monosyllabic adjectives, Hilpert (2008, p. 408-409) finds an exception to the general consensus: adjectives like *dull*, *pale*, *real*, *vile* – ending in a singular */l/* that is not part of a consonant cluster – tend to form analytic comparative. In her study of avoidance of identity effects, Mondorf (2009, pp. 31-32) also finds that, just like the disyllabic ones, monosyllabic adjectives ending in minimally distinct consonant clusters like */-pt/* or */-kt/* readily form periphrastic comparison. Out of the five adjectives examined, *strict* is the only one that unequivocally prefers to form comparison synthetically – once again, seemingly due to its much higher frequency – while three out of five almost completely preclude the formation of the synthetic comparative.

Some sources seem to disprove some of the information in either Quirk et al., Mondorf or Hilpert: Mondorf's findings, for instance, partially refute the claim CGEL makes on the patterning of adjectives ending in */r/*: both *mature* and *obscure* form their comparative periphrastically more than 90% of the time in her data, directly contradicting the example CGEL provides of adjectives which "most readily take synthetic forms" (Quirk et al., 1985, p. 462). The last remaining adjective from CGEL, *clever*, is the only adjective from the group to prefer synthetic comparison in all syntactic positions – but it is owing to, as Mondorf hypothesizes, its high frequency and the resulting entrenchment of the synthetic comparative (2009, p. 21). This is confirmed by the data from Hilpert (2008, p. 407), who finds the */r/* ending in adjectives to influence periphrastic comparative formation. The */li/* ending, which the CGEL, Mondorf and

Hilpert all find to be biased towards periphrastic comparison, is found to lack significant influence in the formation of comparatives in the statistical tests performed by Cheung and Zhang (2016, p. 576), despite their research being largely based on Hilpert's model. Meanwhile, the */i/* ending is discovered by them to be second in importance only to the number of syllables. They theorize that the reason for the different outcome in Hilpert's study can be due to the inclusion of the highly frequent adjective *likely* in his dataset, which Cheung and Zhang exclude from the analysis "due to its strong analytic bias" (ibid., p. 574). Hilpert (2008, p. 400) also disproves Mondorf's claim about the influence of a possible stress clash on the likelihood of synthetic comparison by ascertaining the initial stress of the following word to have no important contribution in the alternation. He does, however, find the final stress in the adjective itself important: in adjectives like *mature*, *intense* and *remote* it is found by him to be relevant to the formation of periphrastic comparative (Hilpert, 2008, pp. 406-409).

Mondorf also argues that the final segment of the adjective would be better regarded as "a formal categorisation parameter rather than elevating it to the status of an explanation"(2009, p. 34). She analyzes 15 adjectives of the same syllable amount and final segment (*/i/*) in non-attributive uses and finds much variability despite them being identical in two parameters: while *lucky*, *crazy* and *silly* form the analytic comparative 10 or less percent of the time, adjectives like *guilty* and *worthy* do so in around 80% of the instances in her data. Mondorf claims that the final segment of an adjective is only a corollary of avoidance of identity effects and morphological complexity (2009, p. 34).

2.2.2. Morphological determinants

Although Quirk et al. (1985, p. 462) state that disyllabic adjectives ending in *-le* mostly take synthetic comparison, Mondorf claims that "[l]umping all instances of disyllabic adjectives with final /-l/ together […] misses an important morphological generalization" (Mondorf, 2009, p. 35). Her examination of 24 adjectives ending in *-l* and *-le* (*/l/*) shows different formation of comparative based on the morphological complexity of the adjective. Adjectives in her data consisting of two morphemes, like *awful* and *brutal*, invariably form the analytic comparative, while monomorphemic adjectives vary significantly, from *stable* using the analytic form almost 100% of the time, to *gentle* comparing periphrastically in less than 20% of the cases. Mondorf finds that "[m]orphological complexity strongly affects the choice of comparative variant" (ibid, 2009, pp. 35-36).

Research conducted by Matsui (2010, pp. 193-194) explores the effect of morphological complexity on comparative alternation further: she analyzes all affixed adjectives in her data and reaches a conclusion that all of them – both with prefixes such as *a-* or *ad-* and suffixes like *-al*, *-en*, *-ful,-less* and *-ic*, among many others, as well as participial adjective suffixes *-ed* and *-ing* – nearly always form the analytic comparative irrespective of their length. Hence, she considers the approach taken by Mondorf incorrect: Matsui deems it necessary to distinguish the suffixes *-ful*, *-al* and *-ile/il* from the *-l/-le* endings. It is important to note that her paper does not acknowledge either *-y*, in adjectives like *sleepy* or *easy*, or *-ly*, in adjectives like in *costly* or *likely*, as suffixes, only talking about them as endings. Adjectives ending in *-y*/*-ly* in her data, which seemingly include the aforementioned suffixes, are more variable. More than 77% of adjectives ending in *-y* and 23% of adjectives in *-ly* prefer the synthetic comparative in 60 up to 100 percent of the cases in her study. Her data also does confirm that adjectives ending in *-le* form periphrastic comparison more readily. Her findings are partially supported by Biber et al. (1999, p. 523), who claim that disyllabic adjectives with suffixes *-ful*, *-less*, *-al*, *-ive* and *-ous* – ones Matsui also mentions – as well as participial adjectives like *bored* and *tiring* all usually form analytic comparatives. Matsui also groups compound adjectives, like her example *carefree*, with the affixed adjectives in taking the synthetic comparative in more than 90% of the instances.

Here her research is once again at odds with Mondorf, who finds that although the particular type of compound Matsui provides as example – one spelled as a single word – does form the analytic comparative in 97% of the cases in her corpus (2009, p. 46), the hyphenated compounds like *broad-based* exhibit variation, and the comparison of those formed by two words mostly patterns synthetically (2009, p. 51). Mondorf theorizes that one reason for such distribution of different compound types is the degree to which they are fused, which is reflected in their form: the low degree of lexicalization is signaled by their realisation as two words, it is higher in those connected by the hyphen, and the most fused compounds are realized as one word (Mondorf, 2009, p. 51). This factor, however, cannot be analyzed in this study, since it is reflected in the spelling of a compound, a type of data which is unreliable in a spoken corpus. Studying the hyphenated compound variant in her corpus, Mondorf does nevertheless find a significant increase in the analytic comparative in adjective parts of these forms, even with dimension adjectives like *long* or *broad*, which otherwise never form it in her data. A possible reason Mondorf posits for the increase of analytic comparative formation is the avoidance of "discontinuity effects produced by the insertion of an infix-like *-er*" in compounds with the adjective as the left-most element. The intervening *-er* ending is more acceptable in less entrenched compounds, however (Mondorf, 2009, p. 54). Nevertheless, disyllabic compounds

like *street-wise* with the adjective as the right-most element also overwhelmingly prefer periphrastic comparative, despite their structure and syllabic length (Mondorf, 2009, p. 48). Interestingly, some of the adjectives in her data also pattern very dissimilarly despite having the same left constituent and the same number of syllables: an example of this is *full-blooded*, which takes the periphrastic comparative in 93% of the cases, and *full-flavored*, which only does so in 6% of the instances in her data (Mondorf, 2009, p. 49). Such uncommon distribution prompts Mondorf to argue that word length in syllables cannot be the only or even the most prominent determinant of comparative alternation in compounds and the increased preference for the periphrastic comparative is "triggered by the morphological complexity and their lexical status as compounds" (Mondorf, 2009, pp. 46-47).

2.2.3. Syntactic determinants

2.2.3.1. Distinguishing complements and adjuncts

To be able to ascertain the effect of the *to*-infinitive and prepositional complements on the formation of the comparative, it is important to distinguish complements from the similarly structured adjuncts. There is, however, no clear delineation between complements and adjuncts found in literature. Two groups of complements exist: obligatory and optional. Compared to obligatory ones, optional complements can be left out – yet they differ from adjuncts in that they are still implied by the clause: in (14) the complement *for departure* can be excluded, however it is still implied in the sense that the boat is ready *for something* (Quirk, 1985, pp. 65-66), while in (15) the element *to tax* is necessary in both supplying the semantic meaning of the adjective and completing the grammatical structure (Hawkins, 1999, p. 242).

(14) *The boat was ready **for departure***.

(15) *All sales are subject **to tax.*** (Quirk, 1985, p. 66)

In some cases, an optional complement of an adjective can only be omitted when its interpretation and the information it normally expresses can be found in the context, like in (16):

(16) *Kim was very keen **to take part*** (Huddleston & Pullum, 2002, p. 542).

Here, the clause *Kim was very keen* can only stand on its own with the material from the context compensating for the missing complement (ibid, p. 542). Nevertheless, Quirk et al. (1985, p. 66) state that "this criterion is not always clear-cut, since the need for semantic satisfaction is a matter of degree". Finally, adjuncts are not bound at all to the adjective and can be excluded without any loss of meaning or grammaticality, like the infinitival construction *to volunteer* in (17):

(17) *I was mad* ***to volunteer*** (Huddleston & Pullum, 2002, p. 1256).

Although in all of the cases mentioned above it is relatively simple to distinguish the three domains, not every instance is so clear: Mondorf describes the construction in (18) as a borderline case, since although the construction in bold should be considered a complement, she claims it is not as closely bound to the adjective head as more prototypical complements:

(18) *I am even more lucky* ***to have had the overwhelming love and support of my family*****.** (Mondorf, 2009, p. 61)

In fact, Mondorf states that the three domains form "a continuum ranging from tightly bound obligatory complements via loosely bound optional complements to adjuncts" (Mondorf, 2009, pp. 61-62).

The movement, substitution and question tests Mondorf lists as those commonly introduced in grammars for differentiating between complements and adjuncts are also not always applicable: an example of this can be seen in (19):

(19) *To discuss melodrama, then, is to raise questions about 'culture' itself and the categories and oppositions by which we conceptualise it.* (Huddleston & Pullum, 2002, p. 1256)

Here, it is impossible to prepose the complement without changing the meaning of the sentence: *To raise questions about 'culture' itself and the categories and oppositions by which we conceptualise it is to discuss melodrama* is not the same statement as in (19) (Huddleston & Pullum, 2002, p. 1256).

Taking into consideration these factors, complements are distinguished from adjuncts in the data analyzed in this study by utilizing the movement, substitution and question tests, as well as the entailment test described in Hawkins (1999, pp. 241-242), which in this case checks for both grammatical and semantic independence of the adjective from the *to*-inifinitival or prepositional construction to determine whether that construction is a complement or an adjunct. Nevertheless, as Mondorf notes, all such operations are necessarily subjective to some degree (Mondorf, 2009, p. 63).

2.2.3.2. Effect of complements

Although it seems to be part of the scholarly consensus that complements of adjectives influence comparative alternation, the effect of the individual ones is still debated – and in the

case of structures with *than*, their very nature as complements is unclear. The effects of three elements – *to*-infinitive complements, prepositional complements and *than*-constructions are usually studied under this umbrella.

Mondorf (2009, pp. 63-64), controlling for the effects of prepositional complements and end-weight, examines the influence of the infinitival complement for 28 mono- and disyllabic adjectives. She excludes those in the attributive position from her analysis, citing the statement by Quirk et al. (1985, p. 420) that "[a]djectives with complementation (…) normally cannot have attributive position but require postposition". Although 12 of the adjectives examined do not appear once with an infinitival complement in her data, and eight of them have less than 5 instances of being modified by an infinitival complement, none of the remaining adjectives contradict the hypothesis that the presence of the infinitival complement positively influences the formation of the analytic comparative. Monosyllabic adjectives especially exhibit an increase in forming the analytic comparative with a *to*-infinitive complement. One example of this is the adjective *apt,* which jumps from only forming comparison with *more* in 40% of the cases to 97% when complemented by an infinitival construction (Mondorf, 2009, pp. 65-67). Next, controlling for the effects of end-weight and *to*-infinitive complements, Mondorf examines the relation between the prepositional complement and the choice of comparative. Excluding those that do not take a prepositional complement once in Mondorf's data, the same adjectives almost universally – 7 out of 9 monosyllabic and 9 disyllabic – exhibit an increase in the formation of analytical comparative when occurring with a prepositional complement (Mondorf, 2009, pp. 72-74).

Finally, for adjectives followed by a construction with *than*, Mondorf refutes the statement by Quirk et al. that adjectives can be compared periphrastically "more easily when they are predicative and are followed by a *than*-clause", as in (20):

(20) *John is **more mad** than Bob is* (Quirk, 1985, p. 462).

She cites Hilpert, whose findings indicate that a following *than*, in fact, increases the likelihood of the inflectional comparative (Hilpert, 2008, p. 407). Mondorf (2009, pp. 78-80) explains the difference in the influence between *than*-constructions and the complements she examines by excluding *than*-phrases from complement status on the grounds of them being less dependent on the adjective than either the infinitival or prepositional complement, claiming that structures with *than* are in fact more dependent on the degree marker. According to Mondorf, this degree of independence is exhibited in *than*-constructions not being contained in the adjective phrase and more readily permitting material to enter between it and the adjective, as in (21):

(21) *I have never been more proud of Neil than I have during the (...)* (ibid, 2009, p. 79).

She attributes the findings of Quirk et al. to positional factors rather than the effect of following *than*. Accordingly, both Watanabe and Iyeiri (2020, p. 90) and Cheung and Zhang (2016, p. 573) find *than*-constructions to be non-significant to the alternation of comparative, despite Hilpert's findings to the contrary.

Some of the findings made by Mondorf have been questioned, however: Gonzalez-Diaz, referring to Mondorf's study of the effect of infinitival complements and how so many of the adjectives examined present no or scarcely any data, states that her findings "should perhaps be taken as suggestive until more tokens/types are examined, for the adjectives on which she draws her conclusions […] cannot be considered representative of either monosyllabic, or […] the disyllabic, class" (2004, p. 369). Hilpert, too, notices that Mondorf's data produces no particular result and considers that the effect as discovered by her analysis "does not hold uniformly across different classes of adjectives and different syntactic environments" (Hilpert, 2008, p. 401). Additionally, despite Mondorf citing the claim made by Quirk et al. that adjectives with complementation are unable to take the attributive position, such cases are possible if the complementation is postposed, as in (22):

(22) They have a *larger* house *than yours*. (Quirk et al., 1985, p. 420)

Nevertheless, such a sentence is considered more informal than its regular version with the adjective in the predicative position (ibid, 1985, p. 420).

2.2.3.3. Premodification

Lindquist (2000, cited in Cheung & Zhang, 2016, p. 561) finds that adjectives premodified by an adverb such as *considerably*, *even*, *ever*, *far*, *much*, etc. (Cheung & Zhang, 2016, p. 565) more readily form the comparative analytically. Hilpert (2008, p. 407) confirms this claim in his study. Watanabe and Iyeiri (2020, p. 90), however, do not detect clear evidence of the influence of premodification on the formation of comparative, and Cheung and Zhang (2016, p. 573) find premodification to be non-significant in determining the type of comparison in their data.

2.2.3.4. Adjective position

It is considered general knowledge in linguistics that while attributive position favors inflectional comparison in adjectives, the trend for both predicative and postnominal positions is the reverse. Examining 38 disyllabic adjectives divided into groups based on their endings and stress placement, as well as 20 monosyllabic ones, Mondorf (2009, pp. 80-84) finds that although monosyllables and disyllables ending in */i/* tend to form synthetic rather than analytic comparison in all three positions, the disyllabic adjectives ending in */l/* and */r/* show different distribution. While the adjectives in */l/* in attributive position still form synthetic comparative in a little more than 50% of the cases in her data, periphrastic comparison is much more prevalent in predicative and postnominal position. Adjectives ending in */r/*, both iambic and trochaic, almost never allow for inflectional comparative in any of the positions. Mondorf attributes the ratio in the last group to identity effects, namely the undesirability of the potential *-rer* ending in synthetic comparison leading to the choice of the analytical realization instead. Hilpert (2008, p. 407) confirms her findings: in his data, an adjective in attributive position is biased towards forming the comparative synthetically, while one in a predicative position is more likely to compare periphrastically. The study by Cheung and Zhang (2016, p. 573) confirms the effects of these adjective positions on the comparative form, and analyses the effect of the postnominal placement – an adjective used postnominally in their data influences the formation of analytic comparative.

2.2.3.5. Parallel constructions

Quirk et al. notice that "[a]djectives also seem to be more free to form periphrastic comparative forms with the correlative construction *the…the"* (Quirk, 1985, pp. 462-463). It is, however, only one of the constructions which allow the adjectives in correlation to form comparison in the same way. In her study, Matsui finds "a tendency for the same type of comparison to be favored" (2010, p. 198) in coordinated adjectives. She finds cases of adjectives with prevalently analytic comparison creating forms with *-er* when following a synthetic comparative in coordination, as in (23), as well as adjectives usually taking the *-er* comparative forming comparison with *more* when coordinated with an adjective in analytic comparative, like in (24).

(23) *(…) the Prime Minister (…) certainly seemed* ***more courageous*** *and* ***more firm*** *in his conviction that a General Election was necessary.*

(24) *It's someone who envies creativity, and wishes to be* ***younger***, ***handsomer*** (…). (Matsui, 2010, p. 197)

However, adjectives that nearly never form the synthetic comparative, like those with the suffixes *-ful*, *-ing*, *-ous* and *-(i)al* in her data, don't take inflectional comparison even in this case:

(25) *The nearest birds began to pitch on a **higher**, **longer**, **more anxious** note.* (Matsui, 2010, p. 198)

Mondorf additionally finds cases of parallelism in proportional clauses, such as in (26):

(26) *Quite often, the **more risky** the investment, the **more promising** the potential.* (Mondorf, 2009, p. 12)

Interestingly, a similar parallelism effect in adjective comparison was found by her even with the comparatives of adverbs (27) and *more* as a quantifier (28):

(27) *The **harder** they work, the **righter** they are about that.*

(28) *The **more fierce** the sun, the **more** power there is to cool the building.* (Mondorf, 2009, p. 12)

Mondorf excludes the cases like (27) and (28) from her data on the grounds of them potentially influencing the findings to misrepresent the real distribution of comparatives (Mondorf, 2009, p. 12). She does not analyze the effect of parallelism on the choice of comparison, however, on the grounds of it being too unwieldy to conduct (Mondorf, 2009, p. 13).

2.2.4. Frequency

Citing the Survey of English Usage, Quirk et al. state that although inflected comparatives are more common than those formed with *more*, many of their occurrences are represented by "a small number of relatively frequent adjectives" (Quirk, 1985, p. 463). Multiple studies (Cheung & Zhang, 2016, Hilpert, 2008, Mondorf 2009) confirm this relation in their data. The rate of analytic comparative formation in the data of Mondorf (2009, p. 41) seems to generally correlate with the overall frequency of the adjective: less frequent adjectives, like *right*, *proud* and *handy* with less than 500 instances of comparative forms in her corpus form analytic comparison more readily than adjectives like *happy, easy* and *hard* with more than 2500 instances each, which almost exclusively prefer the morphological comparative. Nevertheless, the findings are not uniform – the adjectives *true* and *proud* have approximately the same rate of comparatives formed with *more*, despite the former adjective being more than three times more frequent than the latter, and for the disyllabic adjectives, *likely* is once again an outlier, forming the periphrastic comparative almost exclusively, despite being the most frequent in the data. This

prompts Mondorf to conclude that frequency of a given adjective, although an important variable determining the choice of comparative, can be overwritten by other, more impactful factors.

Another variable related to language use which is explored in the academic literature is the ratio of positive and comparative forms of a given adjective (Cheung & Zhang, 2016, Hilpert, 2008). As Hilpert (2008, p. 397) explains, a low ratio of positive and comparative forms is a sign of an adjective's low gradability, which corresponds to the likelihood of periphrastic comparative formation, since novel and less entrenched information tends to be expressed through the analytic form, owing to "its additional phonological material". For instance, in Hilpert's data, the non-gradable adjective *able* has a comparative-to-positive ratio of only 0,0058, while the highly gradable *humble* has a ratio of 0,1141. Accordingly, while *able* only forms the comparative morphologically in 2,8% of the instances, *humble* does so in 72,8%. Therefore, as confirmed in the studies of both Hilpert and Cheung and Zhang, the ratio of positive to comparative form of an adjective is a determinant increasing the likelihood of the formation of synthetic comparative.

2.3. Double marking

An interesting case for this study is presented by the phenomenon of double marking – the production of adjectives marked for comparison both analytically and synthetically:

(29) *This way, it's **more easier** to see.*

In a small number of cases an adjective that usually forms its comparative or superlative irregularly can also be realized with a regular *-er* or *-est* ending, as in (30):

(30) *This is the **bestest** one you can read.*

What is especially important in this study is that such instances are stigmatized in Standard English and usually occur in the domain of conversation (Biber et al., 1999, p. 525). However, a study of comparative variability in two spoken corpora (D'Arcy, 2014) found the amount of doubly marked forms in them to be very small – only 1% and 2% of all instances of comparison in both corpora. Interestingly, it is most prominent in the disyllabic adjectives: 7,5% and 4% of instances of comparison for the disyllables in each corpus are doubly marked.

2.4. Spoken alternation

The amount of academic research on comparative alternation in the spoken language appears to be fairly small. Most notable is a study by D'Arcy (2014), who explores the two

modes of comparison in two sociolinguistic corpora: one of Canadian English and one of New Zealand English, which she also compares to a traditional written corpus of NZE. Despite the similarity of the vernacular corpora to the established patterns in more traditional, written ones in terms of both the high ratio of synthetic to analytic forms and the categorical comparative realization in both mono- and trisyllables (D'Arcy, 2014, pp. 223-225), the comparative distribution in the vernacular corpora appears to be fixed. It shows extremely little variability for the individual forms – for instance, the only forms that actually alternate in the Origins of New Zealand English corpus (ONZE) are *clever*, *common*, *silly* and *cool* with three, two, three and eight occurrences in the comparative respectively, every other adjective in the data having exclusively settled on one or the other mode of comparative formation. Even then, only *clever* out of all instances varies within the comparative form: all the other types exhibit variation either within the superlative or across two forms (D'Arcy, 2014, p. 226). Examining the written NZE corpus (the Wellington Corpus of Written New Zealand English, or WWE), however, D'Arcy finds that its adjectives clearly show individual variation on a scale very similar to that of the BNC: based on these findings, she dismisses the possibility of language variety influencing the distribution. The result of her study leads her to suggest the register of the corpus to be a decisive factor in the lack of variability of the comparative. Although acknowledging the possible issue of smaller corpora not featuring the most interesting rare cases of variation, she nevertheless shows the size of the corpus to only be a minor factor influencing the displayed comparative distribution, since WWE, despite being smaller than ONZE, demonstrates much more variation (D'Arcy, 2014, pp. 229-233). The results of the study are further consolidated by the comparison with TEA, a corpus of Canadian vernacular English, an analysis of which yields results strikingly similar to those gathered from ONZE (ibid, pp. 233-235).

Contemplating on the reason why structurally similar items in her data, like *yellower* and *more mellow*, do not form comparative in the same way, D'Arcy hypothesizes that in the spoken register the comparative variant is either specific to particular lexemes, or is dependent on their frequency. By testing for the frequency of the individual disyllabic items in her data, she finds no correlation between it and the choice of comparative, however. D'Arcy theorizes that the distribution could instead be dependent on the frequency of the various grammatical contexts in which the lexeme is commonly found in (D'Arcy, 2014, pp. 236-237). D'Arcy concludes that "it seems that variation is suppressed in the vernacular, not intentionally but by general cognitive and rhetorical factors that are backgrounded in vernacular usage" (D'Arcy, 2014, p. 238). She claims that since speech lacks the structures requiring more processing effort due to its less complex nature, and is not influenced by factors like euphony – another tendency to avoid

repeating adjacent structures, but in the domain of phonology – and symmetry, among other determinants which evince infrequent, but variable structures, the scope of variation in the spoken register is severely limited.

2.5. Reason for variation

Although the scholarly literature has established many variables that certainly influence the comparative alternation in adjectives, there appears to be no consensus among the academics on the reason for such variation. Mondorf (2009) seems to be the only author to offer a possible explanation, coining the term *more-support*. She invokes Rohdenburg's Complexity Principle, which posits that "[i]n the case of more or less explicit grammatical options, the more explicit one(s) will tend to be favored in cognitively more complex environments" (Rohdenburg, 1196, cited in Mondorf, 2009), and explains the mechanism of *more-support* as the preference of the language users for the more overt analytical comparative in the environments requiring more processing complexity (Mondorf, 2009, p. 6). She goes on to relate each determinant of the analytic comparative she examines to this principle: for instance, the complexity of producing the potential *-rer* ending in morphological comparatives of */r/*-final adjectives or the syntactic complexity of an adjective complement can be alleviated through the use of *more* – just as it can reduce the processing cost of producing a comparative of an infrequently used adjective.

Hilpert, however, warns against such an "attractive solution": loooking at the effects of */i/* and */l/* final segments in the frame of *more-support*, he comments on the improbability of processing complexity as explanation for their influence on the adjective comparative (2008, p. 412). Since the */l/* ending influences the formation of analytic comparison, according to the theory of *more-support*, it should mean that the production of an *-ier* ending is less complex than that of *-ler* – something that does not seem to be the case. Hilpert notices that adjectives ending in */i/* can form an intrusive */j/* glide, as in his provided example */kriːmijər/*, which would make the realization of its comparative form more complex than that of any adjective ending in */l/*. He also mentions the production of similar constructions, namely the *-er* nominalizations of nouns ending in */l/*, such as *caller* or *filler*, which are easily formed and do not seem to bring about a bigger processing load. He concludes that, although many of the factors of comparative variation do have a relation to processing complexity, it should not serve as a grounds for "a monocausal explanation […], which carries some risk of circularity" (Hilpert, 2008, pp. 412-413).

3. Material and Methods

3.1. Material

The data for this study comes from the spoken part of BNC2014, a synchronic corpus of contemporary British English. The spoken part contains 11,5 million words recorded between the years 2012 and 2016 in informal settings. The corcondancer used for the analysis was the Corpus Query Processor (or CQP), created and maintained by the Lancaster University[1]. The corpus uses the UCREL CLAWS6 tagset for part-of-speech tagging: adjectives in the synthetic comparative form are tagged with the JJR tag, allowing for their retrieval using the following CQP query: [pos="JJR"]. Another tag, RGR, corresponds to an adverb of comparative degree, allowing for the search for all analytic comparatives within the corpus with the following query: [lemma="more"&pos="RGR"] [pos="UH"]{0,} [pos="JJ"]. The UH tag is used to search for interjections, which, due to the features of the spoken register, can appear between any two words in the corpus: [pos="UH"]{0,} allows for any amount of interjections between *more* and the adjective, although realistically there would only be one or two. The JJ tag equals an adjective in its absolute form. The results of the two queries were sorted by frequency by choosing the "Frequency breakdown" option and then searched for types present in both lists. Doing this yielded 120 total types alternating between morphological and periphrastic comparative in the corpus. Three adjectives appeared in two different forms for one query for various reasons. *Drunk* appeared in two forms in its analytic comparative – *more drunk* and *more drunken*, *far* had two irregular forms *farther* and *further* in synthetic comparative, and the morphological comparative of *tired* was divided into two types in the corpus: *tireder* and *tired-er*. The second comprised two tokens from the same speaker engaged in a meta-linguistic conversation:

S0578: that 's how tired I am S0576: ***tired-er*** *? S0578: what ? S0576: what ? S0578: that 's how tired I am S0576: little bit* ***tired-er ?*** *S0578: no*

For the sake of this thesis, in each of these cases the two divergent types were counted as one.

The resulting list was then compared to the one in Hilpert (2008, pp. 414-415), which features 250 alternating adjectives from the BNC, to analyze the general differences and trends between the two registers. For ease of comparison with a mainly written corpus, the adjectives not in his list were then removed from the original list of 120 types, leaving 72 in total for further analysis. The adjectives with nearly categorical comparative distribution (those with more than

[1] The BNC2014 corpus, more information about it, as well as the CQP, are available at https://cqpweb.lancs.ac.uk/.

95% of instances of comparative in either synthetic or analytic form) were also removed from the data, leaving 61 adjectives. The obtained sample (referred to as “wider sample” here and elsewhere in the thesis) is shown in *Table 2*. Every adjective was then manually searched for in both comparative forms with the [lemma="*adjective*"&pos="JJR"] and [lemma="more"&pos="RGR"] [pos="UH"]{0,} [lemma="adjective"] queries, the concordance line order randomized, and the first 3 instances chosen for analysis. If there were less than 3 instances, the remaining ones were chosen. Each token was controlled to be from a different speaker (if possible) to exclude idiolect influence.

During the initial steps of the analysis, many instances of incorrectly tagged data were encountered. For eight adjectives, namely *fresh*, *poor*, *sweet*, *sad*, *straight*, *wealthy*, *hungry* and *black*, it led to their exclusion from the data altogether. In many cases, instances that were tagged as analytic comparatives were excluded when *more* as a comparative determiner was erroneously tagged as a comparative degree adverb, as in (31):

(31) S0428: have you got some *more* sweet photos ?

In some cases, what was returned by the corcondancer as an instance of a particular adjective was instead a part of a compound adjective or a fixed expression, as with *heavy duty* and *wet lettuce* in (32) or (33):

(32) S0192: so these are more *heavy duty* than these they 'd make a smoothie and something else in this very cool […]

(33) S0041: I think --ANONnameM is more *wet lettuce* than –ANONnameM

In some cases, an adjective in a comparative form was used by a speaker in the role of an adverb, yet still tagged as JJR, like *simpler* in (34):

(34) S0566: >>the the the government I think told them get some system going to allow it to happen *simpler* or we 'll kick you and the banks have done it

For *fresh*, a unique picture emerged: 16 out of 23 total instances of the form *fresher* tagged as JJR were actually an incorrectly tagged slang word for the term *freshman*:

(35) S0221: I walked into one in fresher 's week saying that

Twice, actual instances of an adjective’s comparative form had to be removed from the analysis: both times the adjectives were used metalinguistically, without being a part of the sentence, as in (36), and so these instances were not able to be analysed:

(36) so it 's all quite em quiet em em eh what 's the opposite of common ? Rare scarcer ?

S0041: rare

S0086: rarer

S0041: more rare

S0086: more rare

All such occurrences were excluded from the data.

After eliminating all the mistagged instances, the final number of alternating adjectives in the sample dropped to 53, with 216 total comparatives. From these, 120 are synthetic and 96 are analytic. Of the total number, 26 types are monosyllabic and 27 are disyllabic. There are no trisyllabic adjectives in the sample. Table 1 shows the numbers of synthetic and analytic comparatives that were taken for analysis for every type out of the 53 (referred to here and elsewhere in the thesis as "smaller sample"):

adjective	synthetic	analytic	adjective	synthetic	analytic	adjective	synthetic	analytic
angry	1	3	funny	3	3	rough	3	3
brown	3	2	grey	1	3	rude	2	1
busy	3	3	hairy	3	1	scary	3	3
chunky	3	1	handsome	1	1	sexy	2	2
classy	3	2	harsh	3	3	sharp	3	1
clear	3	3	healthy	3	3	simple	3	3
cloudy	1	1	heavy	3	2	soft	3	2
comfy	3	3	hot	3	1	spicy	3	2
crazy	1	1	keen	2	3	sticky	1	1
curly	2	1	lean	1	1	strict	3	3
dense	1	1	lucky	2	1	subtle	1	3
empty	2	1	mad	1	1	tidy	2	1
fair	3	1	moist	1	1	tricky	2	3
fine	3	1	narrow	3	1	weird	3	2
fit	3	1	naughty	1	1	wet	3	1
free	2	3	rare	3	1	white	3	1
friendly	3	3	red	1	2	yellow	1	1
full	3	1	roomy	1	1			

Table 1. The smaller sample of manually selected 2016 adjective tokens.

Since it was not feasible to go through all the data in the wider sample manually to eliminate mistaggings, they were only eliminated in the smaller sample in *Table 1*. However, the adjectives that did not have any instances in one or the other comparative form after the incorrectly tagged ones were removed were excluded from the wider sample in *Table 2* as well.

adjective	synthetic forms	analytic forms	adjective	synthetic forms	analytic forms
angry	1	4	lean	1	1
brown	3	2	lucky	2	1
busy	27	3	mad	1	1
chunky	3	1	moist	1	1
classy	5	2	narrow	11	1
clear	43	8	naughty	1	1
cloudy	1	1	rare	7	3
comfy	3	3	red	2	3
crazy	1	1	roomy	1	1
curly	2	1	rough	5	3
dense	1	1	rude	2	1
empty	2	2	scary	7	5
fair	8	1	sexy	2	2
finer	9	1	sharp	6	1
fit	11	1	simple	27	8
free	3	5	soft	30	2
friendly	9	7	spicy	5	2
full	7	1	sticky	1	1
funny	26	3	strict	10	7
grey	1	3	subtle	1	3
hairy	7	1	tidy	2	1
handsome	1	2	tricky	2	5
harsh	3	3	weird	22	2
healthy	30	9	wet	6	2
heavy	44	3	white	14	5
hot	45	4	yellow	2	3
keen	3	3			

Table 2. The wider sample of all the non-categorical variable adjectives in BNC2014 that overlap with data from the BNC (token n=612).

The double comparatives were also searched for with the [lemma="more"&pos="RGR"] [pos="UH"]{0,} [pos="JJR"] query, yielding 21 instances:

double comparative	occurrence N.	double comparative	occurrence N.
more better	3	more later	1
more further	2	more longer	1
more older	2	more oatier	1
more bigger	1	more posher	1
more broader	1	more riper	1
more calmer	1	more thicker	1
more closer	1	more worse	1
more harder	1	more younger	1
more keener	1		

However, since none of the double comparative forms corresponded to any of the 53 adjective types in my samples, they were not analyzed.

3.2. Method

Firstly, the overall amount of variability in the corpus was analyzed by comparing the overall frequency of the alternating forms in the BNC2014 data with that in Hilpert (2008), as well as the data provided on the variability in the two spoken corpora in D'Arcy (2014).

The sample in *Table 1* was quantitatively analyzed for 4 determinants of comparative alternation:

1. number of syllables

2. the final element (*/i/*, */l/*, */r/*, */li/*, *consonant cluster* and *other*)

3. final stress

4. morpheme number

The 216 instances of synthetic and analytic forms from *Table 1* were then analyzed for 6 other determinants of comparative alternation:

1. initial stress of the word following the comparative form

2. complementation by a to-infinitive complement

3. complementation by a prepositional phrase

4. the adjective's function in a sentence (attributive, predicative or postniominal)

5. the presence or absence of following *than*

6. the presence or absence of premodification

The determinants of positive frequency, as well as of the comparative to positive ratio, were not tested since their values are specific for each of the 53 types, and so could not be analyzed by the method employed for all the other variables.

Other possible influencing factors, such as parallel constructions, were also noted. Positive frequency was gleaned by using the [word="*adjective*"&pos="JJ"] query. As the resulting concordance list also included comparative forms with the comparative degree adverb *more*, these forms were not counted as positives. Comparative frequency was found by using the [lemma="*adjective*"&pos="JJR"] | [lemma="more"&pos="RGR"] [pos="UH"]{0,} [lemma="*adjective*"] query. The morpheme and syllable number, stress and the final segment for

all adjectives, as well as the words following them, were consulted in the OED [2] and *Online Etymology Dictionary*[3]. Monosyllabic weak forms, which are never stressed or only stressed in particular instances, were determined by consulting Roach (2010, pp. 89-95). Adjectives that were followed by an unstressed weak form received negative coding for initial stress of right collocate. Those at the end of an utterance or followed by a short pause, transcribed in the corpus as *(.)*, also received this coding. Monosyllabic adjectives were coded positively for final stress. Adjectives ending in /l/ as part of a consonant cluster, as, for instance, *subtle*, only received coding for the consonant cluster. The resulting influence (or lack thereof) of each determinant was considered one by one and compared to its expected effect on comparative formation based in theory. The statistical significance of each variable was controlled by performing the chi-square test[4].

3.3. Hypothesis

Based on the results of the analysis of two spoken English corpora by D'Arcy (2014), the comparative variability is expected to be lower and more restricted in the spoken corpus when compared to a written one. This is due to, as D'Arcy (2014, p. 236) theorizes, decreased complexity in terms of grammar and processing, as well as the lack of features that are inherent to written language, such as euphony and symmetry, which influence variation in forms that are generally not prone to variation (D'Arcy, 2014, p. 232).

The 12 determinants that affect the alternation in the written language are expected to influence the comparative alternation in the spoken corpus in the same way. As per Hilpert (2008), the determinant of most importance for variability is anticipated to be the number of syllables: monosyllabic are expected to be biased towards the synthetic comparative, with more variability in disyllabic ones. This determinant is expected to show the strongest effect on adjective comparison. In terms of phonological determinants, the final endings, specifically /l/, /r/ and /li/ for the periphrastic forms and the final /i/ for the morphological ones are expected to effect the biggest difference between the two types of comparatives. For syntactic determinants, complementation by both prepositional and to-infinitive constructions is predicted to influence the choice of analytic comparison in the biggest way, while adjectives in attributive position are expected to most readily take the synthetic comparative. Finally, for both determinants in the realm of frequency, the bias towards morphological comparison is expected. Meanwhile,

[2] Available at https://www.oed.com/.
[3] Available at https://www.etymonline.com/.
[4] For this, software available at https://www.korpus.cz/calc/ was used. The mode chosen was "2 words in 1 corpus".

determinants such as following *than* and the initial stress of the following word are predicted to have comparatively weak influence or show no comparative bias at all.

4. Analysis

4.1. Variability of the comparative forms

Analyzing comparative alternation in the BNC, Hilpert (2008, pp. 414-415) finds 252 variable types in his data. That is more than double the amount of 120 alternating types in my BNC2014 data. Not all of them overlap, however: only 72 adjectives (including the categorical ones, as well as mistagged instances) are attested in both corpora. All of the forms that are uniquely variable in the BNC2014 can be seen in *Table 3*.

adjective	synthetic	analytic	adjective	synthetic	analytic	adjective	synthetic	analytic
good	3220	12	pretty	12	3	runny	2	1
old	1112	1	loose	9	2	stiff	2	1
bad	911	1	green	6	5	truer	2	3
young	724	2	little	6	3	bolshier	1	1
long	445	3	chubby	5	1	catchier	1	1
high	370	7	flat	5	1	juicier	1	1
low	258	1	lazy	5	1	meaner	1	1
short	134	1	cute	3	2	meatier	1	1
far	127	1	drunk	3	1	sicker	1	1
fast	101	2	fancy	3	1	sleeker	1	1
deep	34	3	interesting	3	82	squarer	1	1
cool	31	3	mild	3	1	tarter	1	1
dear	25	1	adulty	2	1	tinier	1	1
posh	19	7	chilly	2	1	tireder	1	9
calm	16	2	crispy	2	1	trashier	1	1
wise	13	2	plain	2	1	windier	1	2

Table 3. Adjectives that are exclusively variable in the spoken BNC2014.

Nevertheless, the frequencies of only 10 adjectives: *good*, *long*, *high*, *deep*, *cool*, *posh*, *pretty*, *green*, *little* and *interesting* actually proved to be statistically significant after performing the chi-square test. It showed that a comparative form with absolute frequency of less than three instances had a chance to not appear at all for any given corpus of around 10 million words. Although the BNC also features 180 (114 after the chi-square test) alternating adjectives that are invariable in my data, the size of the corpus also plays a big role: as D'Arcy (2014, p. 231) states, the size of the corpus is implicated in the appearance of rare analytic forms which "are the interesting ones regarding analytic comparative formation", such as *apt*, *fit*, *free* and *proud*. The

BNC is 8,5 times larger than the spoken BNC2014, and does in fact feature all four types, whereas my data only includes *fit* and *free*.

More research of comparative variability between a spoken and a written corpus needs to be conducted to learn more concretely if a spoken corpus is less variable then a written one. As of now, it seems that although a written corpus like the BNC is more variable, at least some of the variability, like in rarer forms such as *apt* and *proud*, is effected by its larger size. Despite this, the spoken register seems to feature some alternating forms that are invariable even in the larger amount of written data. This seems to suggest that at least for some adjectives, the variability is, in fact, more restricted in written language.

4.2. Phonological determinants

4.2.1. Number of syllables

Based on the findings in the academic literature, this determinant was expected to influence the formation of periphrastic comparative. It was tested on the larger population of all the alternating adjectives that overlapped with the list found in Hilpert (*Table 2*). In total, my data contained 314 monosyllabic and 298 disyllabic forms. For monosyllabic adjectives, 246 instances formed the comparative synthetically, while 68 did so analytically. For disyllabic ones, 224 comparatives were morphological, while 74 were periphrastic. The results can be seen in *Figure 1* and *Table 4*.

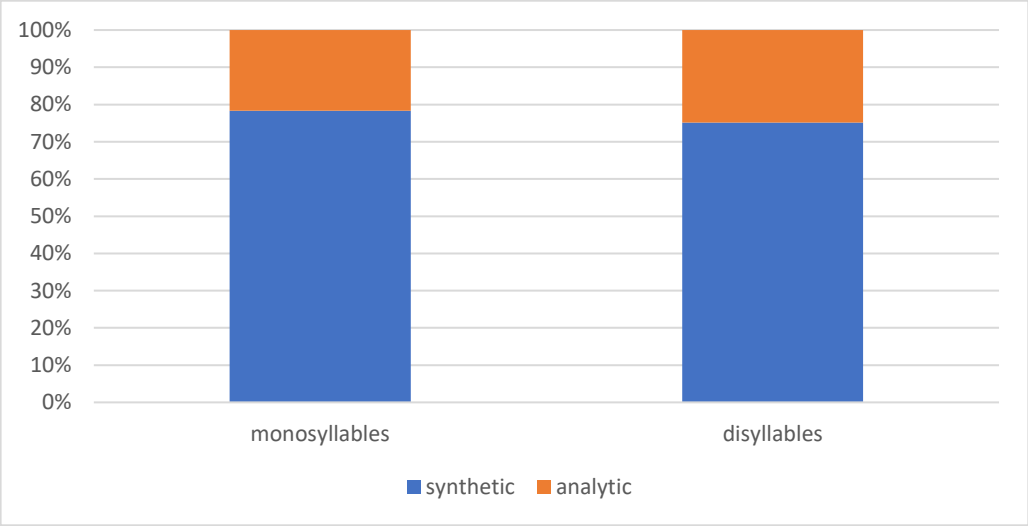

Figure 1. The distribution of both comparative forms in mono- and disyllabic adjectives in the wider sample.

	synthetic	analytic	**total**
monosyllables	246	68	314
disyllables	224	74	298

Table 4. Absolute frequencies of the adjectives in Figure 1.

After performing a chi-square test, these results were found to be statistically significant on the level of 5%. For monosyllables, the result of the analysis seems to support the picture presented in theory – it is established that monosyllables massively prefer the synthetic comparison, and almost 80% of them compare morphologically. The disyllables in my data, however, also showed preference for the synthetic comparative, with 75% of them forming the comparative morphologically, which does not correspond with the expected bias towards the periphrastic comparative found by Hilpert (2008) and Cheung and Zhang (2014). This is given by the forms like *busy*, *funny*, *healthy*, *happy*, *narrow* and *simple* in *Table 2* which still massively prefer the synthetic form despite being non-categorical. There is no such patterning in the sample for the analytic forms. One reason for such distribution may be the general frequency of synthetic and analytic comparatives. As mentioned previously, the morphological comparatives represent a small amount of very frequent types. A direct consequence established in literature is that synthetic comparison is consistently more frequent than analytic (Quirk, 1985, p. 463). The size of the corpus is also once again implicated: the BNC2014 is larger than any of the two corpora in D'Arcy (2014), however it is still much smaller than the BNC: this means that it does not capture the rare comparative forms that open the adjective to the analysis of its variability. The rare synthetic forms of disyllabic adjectives like *able* and especially *likely* (6 and 17 instances of synthetic comparative in Hilpert's BNC data), which are overwhelmingly biased towards the analytic comparative, are not captured by the BNC2014, meaning they cannot be analyzed as variable adjectives. Other analytically biased adjectives like *subtle*, which represents a significant amount of analytically compared disyllables in the BNC (Hilpert, 2008, p. 415), is captured by BNC2014, but only has the positive frequency of 36 in the whole corpus and just 4 total instances of comparative (*cf. Table 2*). As such, the possible reason for such unexpected distribution of disyllables in *Figure 1* may be the general larger frequency of the synthetic comparatives, as well as the smaller size of the corpus which leads to rarer, analytically biased variable adjectives not being captured in my data.

4.2.2. Final element

Based on the available theoretical material, four out of five final segments covered in this thesis influence the formation of adjective comparative periphrastically: */li/*, */r/* and */l/*, as well as

the consonant cluster. The fourth, */i/*, is connected with the bias towards the morphological comparative. The wider sample included 233 total instances of adjectives with the */i/* ending, 176 of them forming the synthetic comparative and 57 analytic. For */r/*, the data contained 58 instances of synthetic comparative and 12 instances of the comparative being formed analytically, and 11 synthetic and 8 analytic comparatives for */li/*. Only one adjective in the data had an */l/* ending: *full* with 7 occurrences in the synthetic and one in analytic comparative. Because of such a restricted amount of data, it was excluded from the analysis. The data includes 69 occurrences of synthetic and 21 of analytic comparatives of adjectives ending in a consonant cluster. The synthetic/analytic ratios for every final element can be seen in *Figure 2* and *Table 5*. The result for every ending except */li/* proved to be statistically significant on the level of 5%.

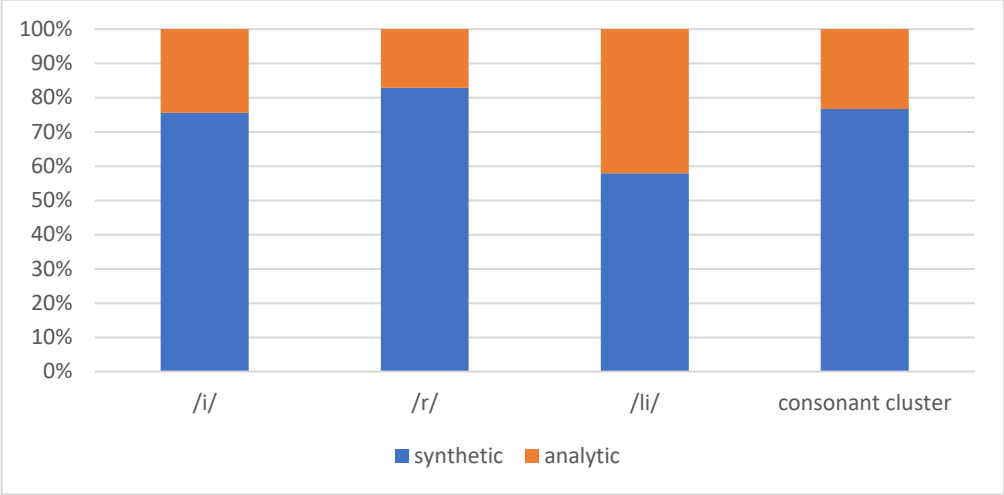

Figure 2. The distribution of both comparative forms for every final element in the wider sample, excluding /l/.

	synthetic	analytic	**total**
/i/	176	57	233
/r/	58	12	70
/li/	11	8	19
cons. cluster	69	21	90

Table 5. Absolute frequencies of the adjectives in Figure 2.

The preference of the final */i/* element for morphological comparison corresponds to the tendency described in described in the CGEL, Hilpert (2008), Mondorf (2009) and Matsui (2010). For both */r/* and the consonant cluster, however, a preference for the synthetic comparative is unusual. The reason for such distribution for */r/* is most probably due to the nature of */r/*-final adjectives in my sample. All of them – *clear*, *fair* and *rare* – are monosyllabic, which biases them to form the synthetic comparative. For the consonant cluster, two out of the three adjectives that were most frequent in my data – *soft* with 30 instances and *strict* with 10 were once again monosyllabic The reason for the lack of significance of */li/* is most likely the size of

my sample and of the corpus itself: I only analyzed two */li/* adjectives: *friendly* and *curly,* however there is only one other variable adjective in the data with the */li/* ending – *curly* – which only has 2 instances in the synthetic and one instance in the analytic comparative. Interestingly, Cheung and Zhang (2016, p. 574) find this ending to have no significant impact on the comparative alternation in their study. This is, as they explain, due to the exclusion of the adjective *likely* from their data, an adjective which was highly biased towards forming the analytic comparative in Hilpert (2008, p. 414). This adjective is invariable in the spoken part of the BNC2014: it is therefore possible that the smaller size of the corpus as compared to the BNC prevents the infrequent form *likelier* from appearing in the data and skewing the results for */li/* towards the analytic comparative.

4.2.3. Final stress

Every disyllabic adjective in the wider sample is stressed on the first syllable. Since the monosyllabic adjectives are coded positively for final stress, the outcome is a distribution that is identical to that of syllable number. The amounts of finally-stressed adjectives in both comparative forms correspond to those of monosyllabic adjectives, and the numbers of disyllabic adjectives in the two forms are the same as those of initially-stressed ones. Therefore, it is impossible to analyze this determinant in my sample, since its potential effect on the comparative alternation cannot be separated from that of syllable length. Although the same could theoretically be said for syllable length itself, both Hilpert (2008, p. 407) and Cheung and Zhang (2016, p. 573) find this determinant to be the most significant for comparative alternation in their respective studies. As such, it was chosen for analysis over the determiner of final stress.

4.2.4. Initial stress of right collocate

The presence of an effect on comparative alternation by this determinant is debated in literature. Although Mondorf (2009) argues that morphological comparative plays a role in preventing a stress clash between the finally-stressed adjective and a following initially-stressed word, Hilpert (2008, p. 406) seems to disprove this claim by ascertaining the statistical insignificance of this variable in his research. This variable was analyzed using the smaller sample of 216 alternating adjectives. Since the basis of this variable is the prevention of stress clash, only the 104 finally-stressed adjectives were chosen for the analysis. Filtering the sample by these parameters yielded 83 forms that were not followed by an initially-stressed word: 44 of them synthetic and 39 analytic. Twenty-one finally-stressed tokens were followed by a word stressed on the first syllable: 16 compared synthetically and 5 periphrastically, as in (37) and (38).

(37) *CLEARER1*: S0238: so are you free is **clearer** is n't it ?

(38) MORECLEAR1: S0520: oh god see they should have made that **more clear** should n't they ?

The result for the comparatives with an initially-stressed right collocate proved to be statistically significant on the level of 5%. *Figure 3* and *Table 6* shows the synthetic/analytic ratios for both groups.

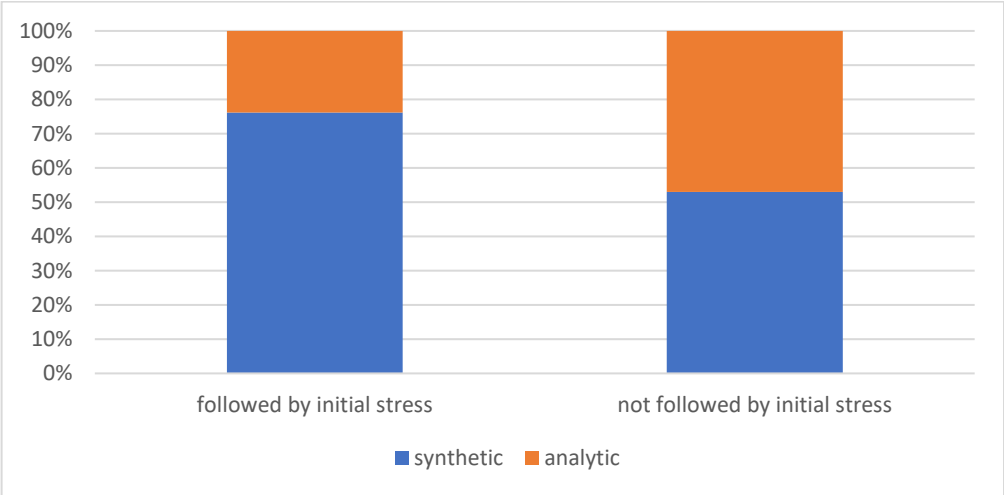

Figure 3. The distribution of adjectives in both comparative forms in the smaller sample followed and not followed by a word with initial stress.

	synthetic	analytic	**total**
initial stress following	16	5	21
initial stress not following	44	39	83

Table 6. Absolute frequencies of the adjectives in Figure 3.

The result corresponds to the findings of Mondorf (2009) – adjectives followed by an initially-stressed word form are much more likely to form the synthetic comparative than those that are not. Nevertheless, it is important to note that since every monosyllabic adjective in my data is marked as finally-stressed, and no adjective in my sample has stress on the second syllable, it was impossible to control for the effect of syllabic length on this distribution.

4.3. Morpheme number

Available theory states that adjectives that are morphologically complex are predisposed to taking the periphrastic comparative: in fact, in the data of both Mondorf (2009) and Matsui (2010), adjectives consisting of two morphemes almost exclusively prefer the analytic comparative. There are 446 monomorphemic adjectives in the wider sample in my data: 358 of those form their comparative synthetically, and 88 do so analytically. One example is the adjective *clear* that has 43 occurrences in the synthetic comparative and 8 in the analytic. From

the 166 dimorphemic adjectives, 112 are found in synthetic comparative and 54 in analytic. The adjective *friendly*, for instance, consisting of morphemes *friend* and the suffix *-ly*, has 9 instances in the synthetic and 7 in the analytic comparative. The ratios in both groups are statistically significant on the level of 5% and are shown in *Figure 4* and *Table 7*.

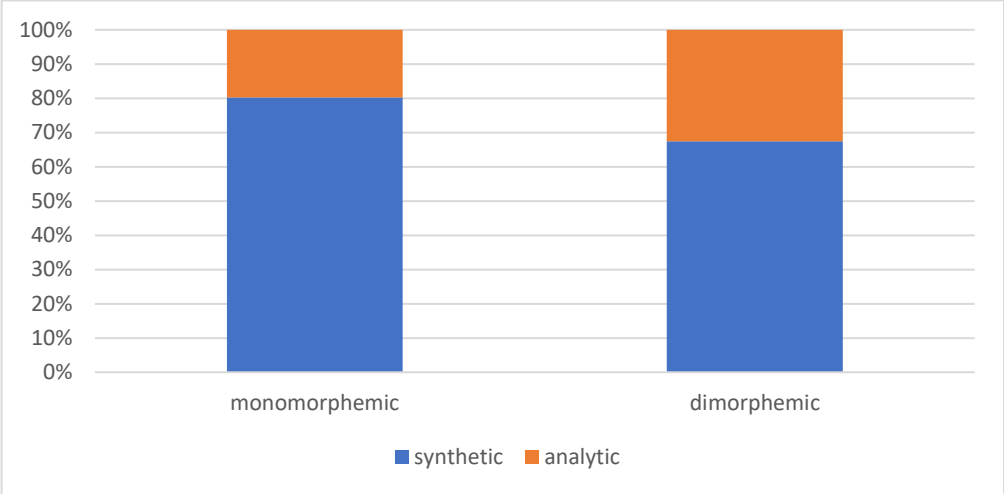

Figure 4. The distribution of mono- and dimorphemic adjectives in both comparative forms in the wider sample.

	synthetic	analytic	**total**
monomorphemic	358	88	446
dimorphemic	112	54	166

Table 7. Absolute frequencies of the adjectives in Figure 4.

Although an increase in morphological complexity corresponded to an increase in analytic comparison as was expected, the increase is quite slight: for monomorphemic adjectives the frequency of synthetic comparative formation is 13% higher than that for disyllabic ones. Nevertheless, it seems to generally correspond to the findings in Hilpert (2008). Although in both Mondorf (2009) and Matsui (2010) the dimorphemic adjectives seem to exclusively prefer the periphrastic comparative, in the case of Mondorf, her data only consisted of adjectives ending in */l/*, which may have contributed to such a major preference. In Matsui, the *-y* and *-ly* suffixes are not considered as such and are analyzed as endings: this seems to be the reason why her findings indicate all suffixes except *-ed* form the analytic comparative in more than 90% of all occurrences (2010, p.192).

4.4. Syntactic determinants

4.4.1. Complementation by prepositional phrase

Based on the information available in the academic literature, a prepositional complement following an adjective in the comparative form should influence analytical comparison. Out of

216 adjectives in the smaller sample, only 6 are followed by a prepositional complement – 3 are synthetic and 3 are periphrastic. Examples of both are provided in (39) and (40):

(39) *ROUGHER2*: S0358: then that started squealing and jumping around and (.) and that just made her more excited so she got even **rougher** *with it*

(40) *MOREKEEN2*: S0435: they 're much **more keen** *about teaching about household wiring*

In (39), the adjective *rough* in the synthetic comparative form is followed by a prepositional complement *with it*: both the question test, described in Mondorf (2002, p.69), and the entailment test, described in Hawkins (1999, pp. 241-242), identify it as a complement. It can be asked for by *what* and followed by the verb *to do*, creating a sentence in (41):

(41) What did she get even rougher with? – With it.

As for the entailment test, although the sentence *She got even rougher* functions correctly in terms of grammar, the semantic property added by *with it* is not entailed in it: this result also establishes *with it* as a prepositional complement. For (40), the situation is the same: a question like in (42) can be constructed, and although *They're much more keen* makes sense both grammatically and semantically, the additional meaning of *about teaching […]* modifies its semantic properties to the point where it cannot be omitted.

(42) What are they much more keen on? – Teaching about household wiring.

The remaining 209 adjectives are not followed by a prepositional complement: 117 of them are synthetic, and 92 are analytic.

The small amount of instances of this variable in the sample is due to its size. Since only 6 total prepositional complements following the comparative were available for analysis, no information on the influence of this variable on the comparative alternation could be inferred from the available data. Moreover, due to the number of occurrences after both forms being the same, the data is not statistically significant.

4.4.2. Complementation by a to-infinitive

Similarly to the prepositional complement, complementation by a to-infinitive is expected to bias the comparative to form analytically. My sample of 216 adjectives only contains two instances of such construction following the comparative: one – in (43) – for synthetic and one – in (44) – for analytic.

(43) *HARSHER2*: S0198: no (.) you should n't they 're much they 're much **harsher** on your liver *to digest*

(44) *MORETRICKY1*: S0417: ex-wife and kind of erm er (.) in the place that 's **more tricky** for me *to find work* and

Similarly to the variable of prepositional complement, the entailment test was used to distinguish a complement from an adjunct. For (43), a sentence like *They're much harsher on your liver* is grammatically correct, yet semantically different without the *to*-infinitive complement *to digest*. Similarly, for (44), the clause does not entail *The place that's more tricky for me*, since although it can function grammatically, the semantic content supplied by tricky is different without the complement *to find work*.

Once again, due to the size of my sample, the amount of analyzable instances for this variable is too small to offer any information that would either support or falsify the hypothesis.

4.4.3. Adjective position

Both the predicative and postnominal position were found in literature to influence the formation of periphrastic adjective comparative. The attributive position, on the other hand, predisposes adjectives towards the synthetic comparative. The smaller sample features 21 instances of an adjective comparative form in the attributive position. Seventeen attributives have the *-er* form, and 4 are formed with *more*:

(45) *HEAVIER1*: S0227: no but you want the the **heavier** one like ?

(46) *MORECOMFY2*: S0141: do people want to go and grab a seat out a **more comfy** seat and stuff and em ?

The sample contains 188 instances of comparatives in predicative position: 98 of these are synthetic and 90 are analytic. An example of each is provided in (47) and (48):

(47) *HOTTER1*: S0198: mm and the chillies and garlic is certainly **hotter** here we think

(48) *MOREHEAVY2*: S0511: yeah the fork 's more **more heavy** than the comb right ?

Finally, for postnominal position, six instances were recorded – 4 morphological, as in (49), and 2 periphrastic, as in (50):

(49) *SHARPER3*: S0434: I need something **sharper**

(50) *MORESPICY1*: S0115: like d' you want that or do you want uh um something **more spicy** ?

One other instance, in (51), is indeterminate:

(51) *CURLIER2*: S0415: --ANONnameF had the total curly hair

S0417: really ?

S0415: yeah t- ten times **curlier** than him

Here, the noun in the third utterance is omitted, creating an adjective phrase: for the sake of the analysis, it is possible to put the noun *hair* in such a way as to render the adjective *curlier* in the postnominal, or in the attributive position. Since one correct position here cannot be determined, this instance wasn't included in the data for this part of the analysis. *Figure 5* and *Table 8* show the overall distributions of comparative forms in all three positions in the sentence:

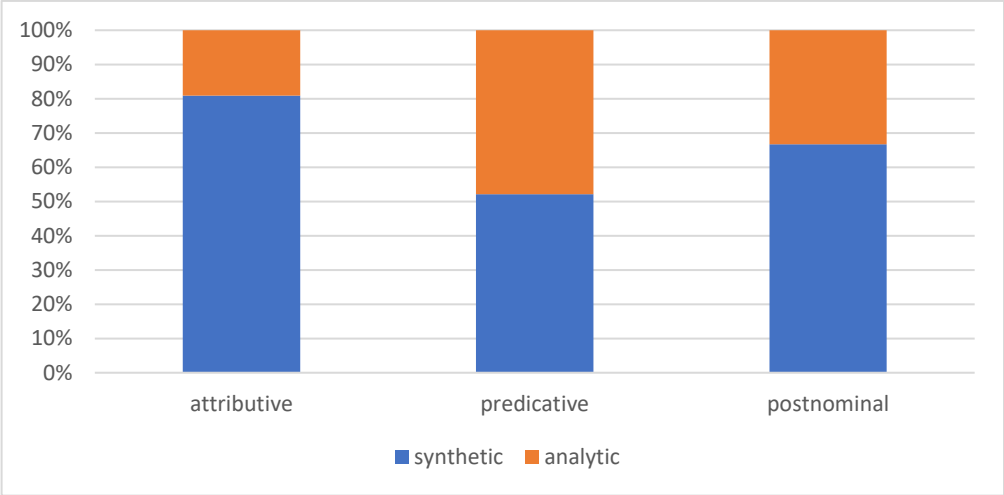

Figure 5. The distribution of adjectives in both comparative forms in different sentence positions in the smaller sample.

	synthetic	analytic	**total**
attributive	17	4	21
predicative	98	90	188
postnominal	4	2	6

Table 8. Absolute frequencies of the adjectives in Figure 5.

Only the comparative frequencies for the attributive adjective position yielded results that were significant on the level of 5%. The comparative variation in this position corresponds to the picture established in the literature: adjectives in the attributive position in my sample form the synthetic comparative in more than 80% of the cases, which corresponds to the findings in Hilpert (2008), Mondorf (2009), and Cheung and Zhang (2016). In the case of the predicative position, the reason for such even distribution and the resulting statistical insignificance appears to be the way in which the sample was assembled. Adjectives in predicative position are the vast majority in the smaller sample – 187 out of 216, or ~87% of the occurrences. Since I took from 1

to 3 (depending on their availability) instances of each comparative form for every adjective in *Table 2*, the numbers of the comparative forms in the smaller sample are quite close (the comparative/periphrastic ratio is 120 to 95: the difference is just 6%). All this causes comparatives in the predicative position to pattern quite evenly. *Figure 6* and *Table 9* compare the comparative distribution in predicative position to the ratio of synthetic and analytic comparatives in the whole smaller sample:

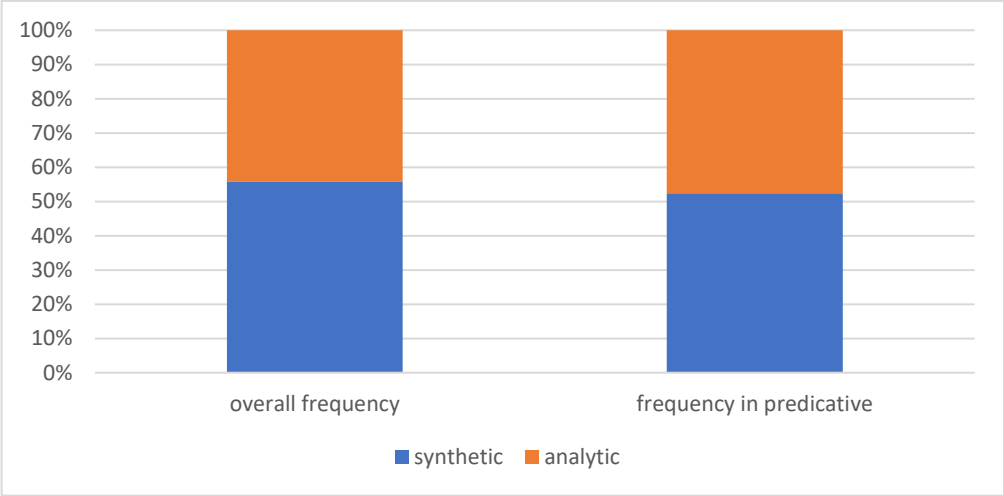

Figure 6. Comparison of the overall frequency of comparative forms in the smaller sample and the frequency of those forms in predicative position.

	synthetic	analytic	**total**
overall frequency	120	95	215
frequency in predicative	98	90	188

Table 9. Absolute frequencies of the adjectives in Figure 6.

The difference between ratios is less than 3%. Since the number of synthetic tokens in the sample is larger than that of analytic ones, predicative adjectives, being almost 90% of the whole sample, have to pattern accordingly. Finally, for the adjectives in the postnominal position, the statistical insignificance can be explained by the small number of occurrences in the sample, which itself is a consequence of its small size. For these reasons, my data on the variables of predicative and postnominal adjective position cannot either support or falsify the hypothesis.

4.4.3. A following than

The findings in the academic literature disagree on the influence of *than* following an adjective on the form of its comparative. It influences the formation of the synthetic comparative according to the CGEL, yet Hilpert (2008) finds it to instead cause bias for the analytic comparatives in the adjectives it follows. Finally, the studies by Cheung and Zhang (2016) and Watanabe and Iyeiri (2020) find this determinant to be insignificant in the formation of adjective

comparatives. There are 40 instances of *than* following a comparative in my data: 21 of those are synthetic, and 19 are analytic.

(52) *WHITER3*: S0618: they are are n't they ? (.) it has n't been built on for a while cos it 's **whiter** than white (.) so --ANONplace to the r-

(53) *MOREBUSY3*: S0555: >>I think it 'll be **more busy** than Turkey Turkey a lot of it was just walking around

For the 176 adjectives not followed by such a construction, 99 form the comparative with *-er*, and 77 do so with *more*. *Figure 7* and *Table 10* show the ratios of both groups.

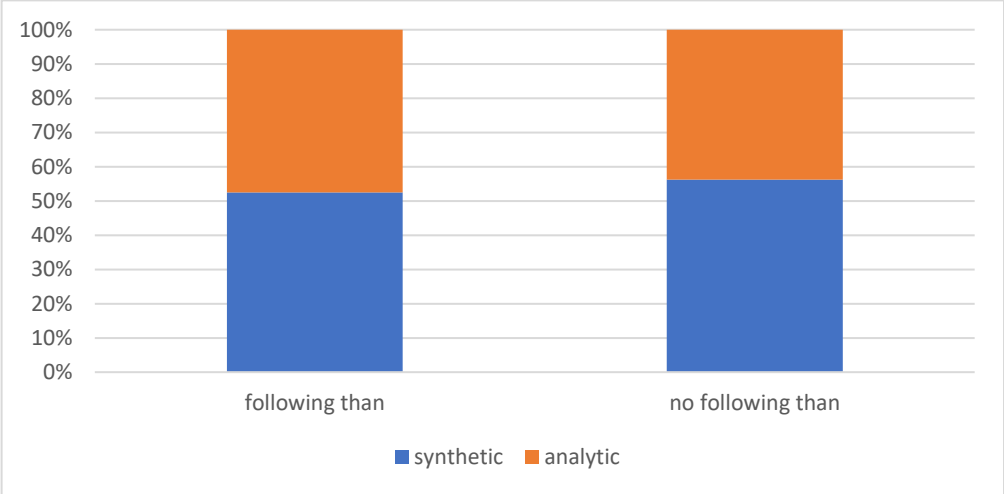

Figure 7. The distribution of adjectives in both comparative forms with and without following than in the smaller sample.

	synthetic	analytic	**total**
following *than*	21	19	40
no following *than*	99	77	176

Table 10. Absolute frequencies of the adjectives in Figure 7.

The result of a chi-square test for the distribution of comparatives followed by *than* showed that it is not statistically significant on the level of 5%. This means that my data on this variable cannot either support or falsify the hypothesis. Interestingly, this correlates with the results in the research of Cheung and Zhang (2016) and Watanabe and Iyeiri (2020), who do not find evidence of this determinant influencing the distribution. Nevertheless, more research on this variable needs to be conducted – especially in the domain of the spoken register – to be able to definitively mark following *than* as irrelevant in the comparative formation.

4.4.3. Premodification

The influence of this determinant on the comparative alternation is also debated in literature. Hilpert (2008), Cheung and Zhang (2016) and Watanabe and Iyeiri (2020) all analyze the significance of this variable in their research. Hilpert finds an adjective premodified by an adverb such as *even*, *much* or *a bit* to be more likely to form its comparative synthetically. The other two studies, however, find premodification to have no influence on the variation. There are 68 instances of comparatives with premodification in my sample: 32 of those are synthetic, and 36 are analytic.

(54) *CLEARER2*: S0144: >>oh they 're *a bit* **clearer** yeah

(55) *MOREHEALTHY3*: S0008: and it 's *much* **more healthy** as well

For 148 remaining comparatives without premodification, 88 are morphological and 59 are periphrastic. The comparative frequencies for both groups can be seen in *Figure 8*. The difference in frequencies of premodified comparatives was not statistically significant on the level of significance of 5%. As such, my data on this variable cannot be used to uphold or contradict the hypothesis. Still, this finding also corresponds to those of Cheung and Zhang (2016) and Watanabe and Iyeiri (2020) on the irrelevance of premodification for comparative alternation. Once again, more research needs to be done to confirm the lack of this determinant's influence, even more so in spoken language. *Figure 8* and *Table 11* show the distribution of comparative forms for both premodified and non-premodified adjectives in the smaller sample.

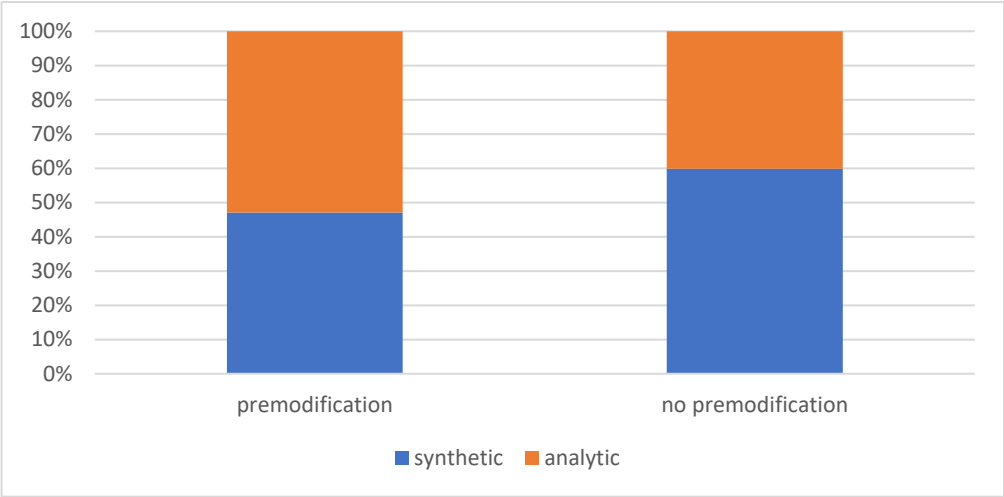

Figure 8. The distribution of adjectives in both comparative forms with and without premodification in the smaller sample.

	synthetic	analytic	**total**
premodification	32	36	68
no premodification	88	60	148

Table 11. Absolute frequencies of the adjectives in Figure 8.

4.5. A note on the process of analysis

Many of the unexpected results of the analysis seem to be connected in some way or other with the lower frequency of the analytic comparative forms. For the unusual morphological preference of the disyllables, it is the smaller size of the corpus, which does not capture the rare forms that would allow for the treatment of, for instance, *apt*, *narrow* and *likely*, as variable in my data. These are all analytically biased adjectives in the data in Hilpert (2008), but also feature synthetic forms. For the predicative position, it is the larger frequency of the synthetic comparative in the smaller sample (stemming from lower availability of analytic forms) which shows predicative adjectives in my sample as preferring the synthetic comparative as opposed to the expected analytic bias. In both situations, the disyllables and predicative adjectives would always be more frequent in the synthetic rather than analytic form. For the larger sample, on which the effect of syllable number was analyzed, the amount of more forms is 142, which is more than lower than 298, the number of disyllables in the data. Even if every analytic form in the sample was disyllabic, adjectives consisting of two syllables would still show a synthetic bias. It is the same for the smaller sample, on which the effect of adjective position was analyzed. The number of predicative adjectives in the sample is 187, which is 92 adjectives more than there are analytic forms in the whole sample, meaning that predicative adjectives would always show a synthetic "bias". Although the smaller sample was compiled by adding 3 instances of each comparative form for every one of the 53 types in Table 1 to add more variance to the sample without capturing too many frequent synthetic tokens, it still has 56% to 44% ratio of synthetic to analytic forms – there was not enough analytic forms available for every type. To account for this, it would have probably been better to present the findings for both forms as relative frequencies per 100, to be able to ascertain how the adjectives influenced by any given determinant would pattern if the number of both forms of the sample was the same. An example of such presentation can be seen in *Figure 9*. All of the blue bars correspond to all of the synthetic comparatives in the smaller sample (120 tokens), and the orange bars represent the total number of analytic ones (95 tokens). The absolute frequencies from *Table 12* can be seen converted into relative frequency per 100 in *Table 13*. As can be seen from the bar chart, predicative adjectives would show a bias towards the analytic comparatives (94 out of 100 analytic comparatives would be predicative), while the attributives would pattern synthetically (14 periphrastic to 4 analytic). The distribution in the postnominal position could still be

explained by the small size of the sample. This is not to say that the expected distribution would be the "correct" one, but to showcase the deceptiveness of conducting the analysis on absolute frequencies. Nevertheless, it could also more deceiving than dealing with the absolute frequency, since the relative frequency number is always abstract.

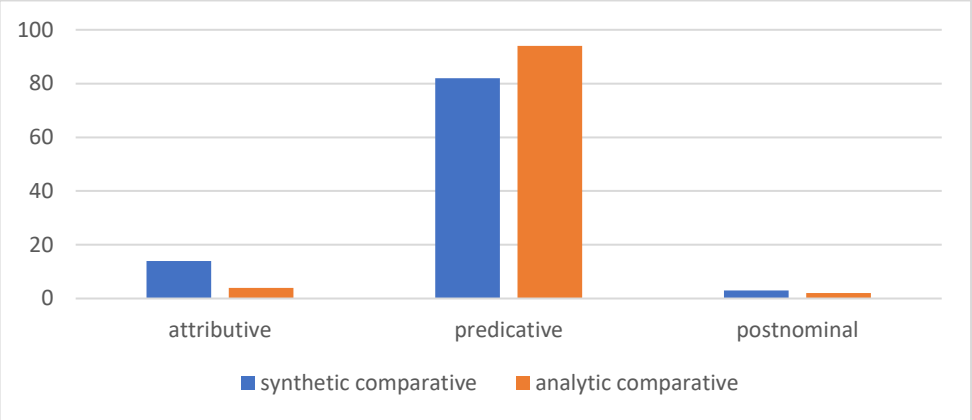

Figure 9. An example of the distribution of adjectives in both comparative forms in different sentence positions in the smaller sample, frequency relative to a 100.

	synthetic	analytic	**total**
attributive	17	4	21
predicative	98	90	188
postnominal	4	2	6
unclear	1	0	0
total	120	96	216

Table 12. Absolute frequencies of the adjectives in Figure 9.

	synthetic	analytic
attributive	14	4
predicative	82	94
postnominal	3	2
total	100	100

Table 13. Instances of adjectives from Table 12, relative to a 100, shown in Figure 9.

5. Conclusion

This study's main focus has been the two comparative form of the adjective, namely the suffix *-er* and the adverb *more*, and the variables that influence the speaker's choice of one or the other in spoken language. All data analyzed in this work was taken from the Spoken BNC2014, a synchronic corpus of spoken English. The analysis was conducted on two samples. For the qualities internal to the adjective that have been examined in literature as variables of comparative alternation, a wider sample of 53 non-categorical types that overlapped with Hilpert's (2008) data from the BNC was used. In total, it counted 612 tokens of both morphological and periphrastic comparative forms. The internal variables examined were the adjective's number of syllables, its final element, the placement of stress and its morpheme count. For the analysis of external variables, such as placement of stress in the right collocate, complementation by either a prepositional or a *to*-infinitive complement, its position in the sentence structure, presence or absence of premodification, and a following *than*, a manually assembled sample of 216 tokens was used. It featured up to three instances (if available) of each comparative form for each of the 53 adjective types.

The first hypothesis of this paper was that comparative variability in adjectives would be restricted in spoken English as compared to the written register. More than twice as many types alternated between the two comparative forms in the written as opposed to the spoken corpus: 252 variable types in the BNC to only 120 in the spoken BNC2014. However, at least ten adjectives – *good*, *long*, *high*, *deep*, *cool*, *posh*, *pretty*, *green*, *little* and *interesting* – were found to be uniquely variable in the much smaller spoken corpus. One hundred and fourteen types only alternated in the written corpus, however it is also 8,5 times larger than the spoken BNC2014, and as such includes forms that are not present in the smaller corpus, like *apt* and *proud*. Overall, while the BNC showed bigger variability in comparative forms, it is generally given by its bigger size, and the presence of adjectives that are variable in the spoken corpus but not in a much bigger written one shows that, at least for some adjectives, comparative variability is in fact less restricted in the spoken register.

The main hypothesis of the thesis was that the comparative alternation is governed in the spoken language by the same determinants as in the written register. For phonological determinants, only some actually corresponded to their expected influence: monosyllabic adjectives formed their comparative form synthetically with 78% frequency, adjectives with */i/* as their final segment compared morphologically in 75% of the instances and initially-stressed

words followed the synthetic comparative of a finally-stressed adjective in 80% of the cases. The only morphological determinant, number of morphemes, also proved to pattern according to the findings in the literature: monomorphemic adjectives formed the synthetic comparative in 80% of the cases, while the dimorphemic ones showed more variation: they only compared synthetically in 68% of the cases. Regrettably, not many conclusions could be drawn from the data on the syntactic variables. The data for both the prepositional and *to*-infinitive complement was too small to make any meaningful observations, and the differences in frequencies between the two comparative forms proved to be insignificant for both the predicative and postnominal adjective positions, as well as the variable of following *than* and premodification. Adjectives in the attributive position however did pattern as expected, with more than 80% of cases of morphological comparative.

Although the distribution for many of the variables proved to be statistically insignificant, for most of them this result actually correlated with similar findings in some of the previous studies. The determinants of final */li/*, following *than* and premodification were also found to be insignificant in the formation of comparative in Cheung and Zhang (2016) and Watanabe and Iyeiri (2020). For the remaining non-significant determinants, the size of my data for postnominal adjectives proved to be too small to provide any statistically significant results. The statistically insignificant patterning of the predicative adjectives in my sample was due to the way in which the sample was compiled: taking up to 3 instances of each comparative form for every type in the sample led the adjectives in this position – which comprised almost 90% of the sample – no choice but to pattern similarly to the overall distribution of comparative forms in the data.

Finally, the distribution of some determinants patterned differently to what was expected. The disyllabic adjectives which are found by Hilpert (2008) and Cheung and Zhang (2016) to be very biased towards forming the comparative analytically, instead formed the synthetic comparative in 75% of the cases. This distribution is thought to be influenced by both the size of the corpus and the small frequency of analytically biased adjectives. The very unusual patterning of */r/* with 82% of forms being morphological can be explained by the fact that all three /r/-final adjectives in the sample were monosyllabic, and so very biased towards the synthetic comparative. For the consonant cluster, which also had 78% of instances in the synthetic comparative, two out of three most prominent adjectives, *soft* and *strict*, with 30 and 10 instances respectively, were monosyllabic and so skewed the result for this ending to the side of the morphological comparative.

Overall, much of what this study set out to do in the beginning was, unfortunately, not accomplished. Although the variability in the written corpus was found to be greater, with much more uniquely variable types, much of it is given by the BNC's size, and spoken BNC2014 features uniquely variable types of its own. More research into the variability in the two corpora and its relationship to their sizes needs to be conducted to achieve more concrete discoveries. For the determinants of comparative alternation, only a couple, and mostly only partially, exhibited the behavior expected in the hypothesis. The distribution of some proved to be non-significant, but – which is more important – the behavior of many of the variables in the analysis was unexpected, and in all cases seemed to be linked to either the way in which the sample was assembled, or the size of the corpus. In fact, looking at the overall results of the analysis, it seems quite likely that the distribution was skewed towards the synthetic comparative due to the sheer number of synthetic adjectives in comparison to the analytic group, in both samples. A way to control for the uneven size should have been introduced into the analysis so its results can be more concrete and have more weight in the study of determinants of comparative alternation in spoken language. More research – which would account for the wildly uneven general frequencies of synthetic and analytic forms – needs to be performed in this area of language to be able to definitively state whether or not the comparative alternation in the spoken register is governed by the same determinants as it is in the written word.

References and sources

References:

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.

Cheung, L., & Zhang, L. (2016). Determinants of the synthetic–analytic variation across English comparatives and superlatives. *English Language & Linguistics*, 20(3), 559–583. https://doi.org/10.1017/S1360674316000368

D'Arcy, A. (2014). Functional Partitioning and Possible Limits on Variability: A View of Adjective Comparison from the Vernacular. *Journal of English Linguistics*, 42(3), 218–244. https://doi.org/10.1177/0075424214539702

Dušková, L. et al. (1994) Mluvnice současné angličtiny na pozadí češtiny. Praha: Academia.

González-Díaz, V. (2004). [Review of the book *Determinants of grammatical variation in English*, by Günter Rohdenburg and Britta Mondorf.] *English Language & Linguistics*, 8(2), 360–373. https://doi.org/10.1017/S1360674304221475

Hawkins, J. A. (1999). The relative order of prepositional phrases in English: Going beyond Manner–Place–Time. *Language Variation and Change*, 11(3), 231–266. https://doi.org/10.1017/S0954394599113012

Hilpert, M. (2008). The English comparative – language structure and language use. *English Language & Linguistics*, 12(3), 395–417. https://doi.org/10.1017/S1360674308002694

Huddleston, R. D., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.

Love, R., Dembry, C., Hardie, A., Brezina, V. and McEnery, T. (2017). The Spoken BNC2014: designing and building a spoken corpus of everyday conversations. In *International Journal of Corpus Linguistics*, 22(3), pp. 319-344.

Matsui, C. (2010). –Er Type or more Type Adjectives of Comparison in English. *Studia Neophilologica*, 82(2), 188–202. https://doi.org/10.1080/00393274.2010.521421

Mondorf, B. (2002). The Effect of Prepositional Complements on the Choice of Synthetic or Analytic Comparatives. In Hubert Cuyckens & Günter Radden (Eds.), *Perspectives on Prepositions*. Tübingen : De Gruyter. https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=e000xww&AN=674092&lang=cs&site=ehost-live&scope=site

Mondorf, B. (2003). Support for more-support. In Günter Rohdenburg & Britta Mondorf (Eds.), *Determinants of Grammatical Variation in English*. De Gruyter Mouton. https://doi.org/10.1515/9783110900019

Mondorf, B. (2009). *More Support for More-support: The Role of Processing Constraints on the Choice Between Synthetic and Analytic Comparative Forms (Issue v. 4)*. John Benjamins Publishing Co. https://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN=284026&lang=cs&site=ehost-live

Online Etymology Dictionary. Available at https://www.etymonline.com. (last accessed 27 July 2024).

Oxford English Dictionary. Available at http://www.oed.com. (last accessed 27 July 2024).

Quirk, R. (Ed.). (1985). *A Comprehensive grammar of the English language*. Longman.

Roach, P. (2010). *English phonetics and phonology: A practical course* (4. ed., reprinted). Cambridge Univ. Press.

Watanabe, A., & Iyeiri, Y. (2020). Explaining the variability of adjective comparatives and superlatives: Entering the twenty-first century. *WORD*, 66(2), 71–97. https://doi.org/10.1080/00437956.2020.1751961

Sources:

British National Corpus 2014, spoken part (Spoken BNC2014). Available via the Lancaster University's Corpus Query Processor website at https://cqpweb.lancs.ac.uk (last accessed August 4, 2024).

Resumé

Bakalářská práce se zabývá variabilitou adjektiv v komparativu a vlivem (nebo jeho absenci) determinantů alternace komparativu ve mluveném jazyce a srovnáním výsledků s očekávaným vlivem na zaklade podobného dopadu ve písemném jazyce. V angličtině existuji dvě možné podoby komparativu adjektiva: jedna se vzniká pomoci přípony *-er*, druha – pomoci adverbia *more*. V odborné literatuře bylo zjištěno, že vyber mezi dvěma formy je podmíněn sdíleným vlivem řady determinantů. Vědecký výzkum odhalil hodně proměnných, kterými se řídí alternace komparativu, ve mnoha různých oblastech lingvistiky: proměnné fonologické, morfologické, syntaktické, determinanty v oblasti sémantiky a pragmatiky (cf. Hilpert, 2008, Mondorf, 2009) a taký sociolingvistické determinanty jako formalita (cf. Watanabe and Iyeiri, 2020).

Práce se zabývá pouze determinanty, které jsou vhodné pro korpusový výzkum. Teoretická část práce popisuje čtyři skupiny proměnných: fonologické, morfologické, syntaktické a determinanty, spojené s frekvenci adjektiva. Fonologickým determinanty jsou počet slabik, koncový segment adjektiva, a postavení přízvuků jak v samotném adjektivu, tak v jeho pravém kolokatu. Počet slabik je už dávno ustáleným v oborové literatuře determinantem alternace komparativu: jednoslabičná adjektiva téměř vždy tvoří synteticky komparativ, zatímco tříslabičná skoro vždy jeho tvoři analyticky. Dvouslabičná adjektiva je doménou největší variace. Nicméně, Hilpert (2008) a Cheung a Zhang (2016) konstatují, že dvouslabičná adjektiva častěji tvoři analytický komparativ. Stejně jako determinanty alternace komparativu jsou v odborné literatuře nejčastěji zmíněny koncovky */i/*, */li/*, */r/*, */l/*, a souhláskový shluk na konci adjektiva. Jedinou koncovkou, která ovlivňuje tvoření syntetického komparativu je */i/*: každá jiná směruje adjektivum k tvoření komparativu analyticky. Mondorf (2009, pp. 17-18) tvrdí, že v případě, kdy po adjektivu s přízvukem na poslední slabice následuje substantivum s přízvukem na slabice první, komparativ adjektivu bude vytvořen synteticky, aby se zbavit potenciálnímu střetu přízvuků.

Ve morfologické doméně alternace komparativu Mondorf (2009) a Matsui (2010) tvrdí, že zatímco monomorfemní adjektiva tvoři v jejich datech obě formy komparativu, dimorfemní téměř vždy tvoří analytickou formu. Mondorf (2009) taky zjišťuje, že složeniny v její datech jsou tím více formují analytický komparativ, čím více jsou spojena. Ty, které měly formu jednoho slova, 97 % komparativů skládaly analyticky. Nejvíc variability bylo u těch s pomlčkou. Složeniny realizované jako dvě slova skoro vždy měly syntetický komparativ.

Syntaktickými determinanty, které se v literatuře vyskytuji, jsou komplementace buď to-infinitivem nebo předložkovým doplňkem, *than* následující za adjektivem, premodifikace, a taky buď predikativní, atributivní nebo postnominální pozice adjektiva ve větě. Hilpert (2008), Mondorf (2009) a Cheung a Zhang (2016) zjišťuji, že oba typy doplněni, stejně jako predikativní a postnominalní pozice ve větě, ovlivňuji tvořeni analytického komparativu. Adjektiva s premodifikaci anebo následujícím *than* naopak spolu s atributivním postavením adjektiva obvykle tvoři syntetický komparativ.

Nakonec determinanty, spojenými s frekvenci adjektiva, jsou jeho frekvence v positivní formě a podíl počtu jeho tokenu v obou komparativních formách ke počtu výskytu jeho pozitivní formy. Protože skupina adjektiv, obvykle tvořících syntetické komparativní formy, se skládá z malého počtu velice frekventovaných typů, malá frekvence se obvykle souvisí s analytickou realizaci komparativa. Potvrzuje to výsledky studii Hilperta (2008), Mondorf (2009) a taky Cheung a Zhang (2016). Malý poměr komparativních forem vůči pozitivním je taky determinantem pro výskyt analytické formy, protože znamená nižší stupňovatelnost adjektiva a vyšší úsilí pro jeho zpracování. Takové případy vyžaduji explicitnější analytickou formu.

Však o vlivu některých z těchto proměnných někteří vědci diskutují. Cheung a Zhang (2016) nepovažují koncovku */li/* jako relevantní pro alternaci komparativu. Watanabe a Iyeiri (2020) se připojují k Cheung a Zhang, protože spolu s nimi uznávají následující *than* a premodifikaci nevyznanými ve svých datech.

Nakonec v praktické části deset z těchto dvanácti determinantu komparativní alternace (mimo determinantu spojených s frekvenci) bylo analyzováno ve dvou vzorcích. První vzorec se skládal ze 612 alternujících se komparativních tokenů 53 typů a byl použit pro analýzu interních determinantů komparativní alternace adjektiv. Ty 53 typů jsou nekategorická alternující se adjektiva, která se překrývá s variabilními adjektivy ve datech z BNC u Hilperta (2008). Všechny ty omezeni byly provedené pro snazší porovnání dat v mých vzorcích s daty ze písemného korpusu. Druhý vzorek se skládal ze 216 manuálně analyzovaných tokenů stejných 53 typů, a byl sestaven z maximálně 3 tokenů každé komparativní formy pro každý z 53 typů. Ten vzorek byl použit pro analýzu vnějších determinantů komparativní alternace.

Byla také zkoumaná studie komparativní alternace ve dvou mluvených korpusech (D'Arcy, 2014). Výsledky teto studie jsou zajímavé: D'Arcy odhalila velmi omezenou a malou variabilitu v obou zdrojích. Z tohoto důvodu byla variabilita v BNC2014 porovnaná s variabilitou v BNC, aby bylo možné zjistit, jestli její pozorování jsou aplikovatelné ve větším mluveném korpusu v srovnaní s písemným. Po analýze celého korpusu v srovnání s BNC bylo

zjištěno, že i když písemný korpus stejně zahrnuje vetší variabilitu, je stálé vice než osmkrát větší než BNC2014 a proto muže zachytit méně časté formy. Nicméně, BNC2014 stejně ukazuje osobní variabilitu: aspoň 10 adjektiv jsou variabilní v BNC2014, ale ne v písemném BNC.

Konečné výsledky analýzy determinantů komparativní alternace byly velmi překvapivé: jen proměnné, které ovlivňovaly formaci syntetického komparativu, přinesly očekávané výsledky: jednoslabičná adjektiva, přízvuk na první slabice pravého kolokatu a adjektiva v atributivní pozici. Postnominalní pozice adjektiva, koncové /i/, následující *that* a premodifikace nebyly v mých datech statisticky signifikantní, a efekt finálního /l/, finálního přízvuku a obou typu komplementace nebylo možné analyzovat kvůli malému objemu dat nebo překrývaní efektů dvou determinantu. Zbývající determinanty – koncové */r/* a souhláskový shluk, prediktivní pozice, dislabičná a dimorfemní adjektiva mají překvapivou distribuci. I když vysokou frekvenci v syntetické formě pro dva koncové segmenty je možné vysvětlit – vetší část dát pro každý z nich je zastoupena monoslabičnými adjektivy – situace je zajímavejší pro ostatní tři determinanty. Každý z nich má větši počet analytických komparativů než ostatní pozice ve větě, monoslabičné nebo monomorfemní adjektiva, ale stejně mají větší poměr syntetických komparativů vůči analytickým. Je jich efekt na komparativní alternací viditelný, ale nezobrazuje se na celém grafu.

Závěrem bakalářské práce je, že metoda analýzy, kterou jsem původně použil, byla chybná. Když začal jsem pracovat nad praktickou části, nezohlednil jsem mnohem vetší počet syntetických komparativů v obou svých vzorcích. Pravděpodobně, metod analýzy, který by mohl zobrazit rozložení adjektiv v obou vzorcích jako relativní frekvence na 100, by mohl být správnou volbou. Na druhou stranu, protože se jedná o abstraktní hodnotu, možná by to bylo více zavádějící než jednání s absolutní frekvenci.

Appendix

The appendix contains all of the strings of text that feature the comparative forms that constitute the smaller sample in *Table 1*. Although only the sentences which contained the comparatives were analyzed, the entire utterances where the adjectives occurred are given here to supply context. When the comparative was divided into two or more utterances by pauses or intervening speakers, or needed context contained in the adjacent ones, those were provided as well. For utterances which, when viewed in the corcondancer, did not have a clear identifiable beginning or were too long, irrelevant context was omitted. Intervening utterances that were not relevant were also omitted. This omission is marked by the […] symbol in the table itself. For every instance, the comparative form is in bold.

HOTTER1	S0198: mm and the chillies and garlic is certainly **hotter** here we think
HOTTER2	S0321: >>yeah it it it does need a green chilli which is probably **hotter** so
HOTTER3	S0521: >>there was a thing in the paper was it the paper we saw it ? it said erm the forecast is that Europe gets ho- **hotter** but the northern countries get cooler because of the Gulf Stream that 's the forecast for the coming years
MOREHOT1	S0008: which the steam goes in the wood all and er so he improved that so he got **more hot** hotter and er that did n't work and then he chucked all his wood in his pond so it got you know
HEAVIER1	S0227: no but you want the the **heavier** one like ?
HEAVIER3	S0623: it would be slightly more than a pound S0611: be probably **heavier** is gold heavy or light ?
MOREHEAVY1	S0167: yeah but she 's quite nice though is n't she ? she was like well you know it does make your periods **more heavy** and it does make them quite painful and there is a there is a higher rate of failure and stuff but you know it still it still works and it 's not it does n't fail all the time
MOREHEAVY2	S0511: yeah the fork 's more **more heavy** than the comb right ?
CLEARER1	S0238: so are you free is **clearer** is n't it ?
CLEARER2	S0144: >>oh they 're a bit **clearer** yeah
CLEARER3	S0427: speaking to one or two of them it becomes **clearer** after a bit what they 're at
MORECLEAR1	S0520: oh god see they should have made that **more clear** should n't they ?
MORECLEAR2	S0084: sourcing paper and pa- prices and that sort of thing (.) tut (.) so split it m- so that it 's **more clear** what each person does
MORECLEAR3	S0558: but apparently my dad reckons the speakers are better on it he thinks he prefers the sound of that one to this one where I think this is **more clear**
HEALTHIER1	S0454: >>yeah I think it 's sometimes it 's **healthier** to have a […] proper break yes
HEALTHIER2	S0618: >>so you feel **healthier** for the first six months of a calendar year and less healthy (.) for the second six months
HEALTHIER3	S0120: that 's probably not much **healthier**

MOREHEALTHY1	S0441: yeah he looks a little bit **more healthy**
MOREHEALTHY2	S0284: erm (.) what Edward Norman called the (.) --UNCLEARWORD secrets and narrow privacy in family life *[...]* S0284: er are not nearly so s- (.) not nearly so so covered up as they as they once were *[...]* S0284: so much **more healthy**
MOREHEALTHY3	S0008: and it 's much **more healthy** as well
SOFTER1	S0455: oh good mm (.) when I 've listened to that I was surprised cos I I du n no if I my voice has got **softer** but I was more gently spoken than I thought
SOFTER2	S0445: bana- bananas are easier cos they 're **softer** and
SOFTER3	S0530: >>I still prefer my be- my home bed is a lot **softer**
MORESOFT1	S0369: they 'd let a bit out *[...]* S0369: and make it a bit so as it 's softer S0251: a bit spongier yes S0369: >>and er so you know it 's a bit fatter S0251: a bit **more soft**
MORESOFT2	S0587: my my dad was very er offensive to me he said that erm mine and my brother 's accents are like **more soft** now like like we say erm bath instead of bath I was like but I do n't
BUSIER1	S0342: >>when you 're younger as well you have more time for your friends do n't you ? and then you get **busier** when you 're older do n't you ?
BUSIER2	S0083: you so loved doing it and getting on with it so even even the hassles about being there were were nothing (.) I mean you you absolutely loved it and the **busier** you were (.) the more deadlines you had the more you loved it and I thought I do n't believe it (.) this is --ANONnameF (.) I bet she
BUSIER3	S0019: so that 's all I 've got to do today (.) been in to --ANONplace (.) been to the bank done some shopping (.) eh those are the total commitments for today (.) mm tomorrow is a bit **busier** (.) cos I got --ANONnameM eh --ANONnameM (.) chris coming at half past nine it is half past nine and then I 've got tennis at two (.) so that 's it (.) and then Wednesday I 've got --ANONnameN coming at half past nine again I 've got to get nanny to the hairdressers by half past eleven (.) I 've got the osteopath at half past three and then I 've got aqua aerobics at six o'clock
MOREBUSY1	S0549: >>I was gon na say yeah it was on the way back it was **more busy** was n't it ?
MOREBUSY2	S0475: I 'm actually **more busy** or busier
MOREBUSY3	S0555: >>I think it 'll be **more busy** than Turkey Turkey a lot of it was just walking around
SIMPLER1	S0192: he definitely seemed a little **simpler** when he came over practising through and
SIMPLER2	S0613: oh it 's so much **simpler** for men
SIMPLER3	S0251: well you 'd think they 'd just have one style for the whole university to make it **simpler** would n't you ?
MORESIMPLE1	S0680: well it yes I mean they have been choosing their own patterns but and they had patterns for jackets and things but er you know some of the **more simple** stuff I suppose the er what do you call it ? the revamping things
MORESIMPLE2	S0555: >>I think the re- the reason that we ca n't is mostly because of the fact that when we develop we (.) because of the fact that we 're so successful we developed such large brains and now we 're and (.) now we are too complicated to compare to primates who who *[...]* S0555: who are much **more simple**
MORESIMPLE3	S0586: it w- yeah it would be **more simple**
FUNNIER1	S0253: >>about people and and that if you read it in a big chunk it 's **funnier** because you 're kind of you 're seeing like how he
FUNNIER2	S0083: he looks even **funnier** from the other side

FUNNIER3	S0173: so is he actually **funnier** when he 's just being himself ?
MOREFUNNY1	S0520: yeah it 's **more funny** when they come in occasionally and do like a silly thing
MOREFUNNY2	S0326: yeah --ANONnameF 's quite --ANONnameF 's lols le- see --ANONnameF 's **more funny**
MOREFUNNY3	S0330: he knows we 're not American it is and he also knows that we think it 's **more funny if** he speaks in Italian er rather than English
WEIRDER1	S0607: this one looks even **weirder**
WEIRDER2	S0187: yeah that 's weirder (.) that is definitely definitely **weirder**
WEIRDER3	S0152: so that was Friday nights (.) and last night dream was even **weirder** (.) it was like past and present all in one go
MOREWEIRD1	S0115: um in fact it 's I I 'd be surprised if people are n't even **more weird** than uh *[...]* S0115: than we than we think they are really
MOREWEIRD2	S0084: yeah (.) it 's a It 's a strange choice I 'll give you that (.) she 's just trying to be **more weird** than Lady Gaga and until you wear a meat dress
WHITER1	S0041: where is it ? **Whiter** teeth (.) bigger boobs (.) and an Audi
WHITER2	S0115: it 's done the uh well it 's done in like one second what like took us like fifteen minutes to do (.) it 's going a lot **whiter** in n it ?
WHITER3	S0618: they are are n't they ? (.) it has n't been built on for a while cos it 's **whiter** than white (.) so --ANONplace to the r-
MOREWHITE1	S0439: >>and just repaint the cupboards cos the cupboards just need like to be **more white** they look now looking like a bit of a weird yellowy mank colour
FITTER1	S0097: but Jan is probably ten times **fitter** than me which is the eh
FITTER2	S0618: >>--UNCLEARWORD was mostly younger **fitter** people as well
FITTER3	S0249: well I 'll move then *[...]* S0249: somewhere **fitter**
MOREFIT1	S0324: yeah but still you 're **more fit** than me
NARROWER1	S0251: >>yes I know sometimes it can be a bit tight that might be why it 's left because it 's a slightly **narrower** one
NARROWER2	I mean some have got dropped handle bars but the majority of them have got narrow straight handle bars S0086: like my bike S0083: like yours yeah but **narrower** than yours
NARROWER3	S0037: like so you know if you were knitting a jumper or something ? You might wan na make it **narrower** towards the top or (.) you know sleeves you make them smaller um at the bottom
MORENARROW1	S0183: they 're **more narrow** than we are
STRICTER1	S0044: they 're a lot **stricter** over there though in the
STRICTER2	S0123: that 's amazing (.) but I would love them to be um a little bit **stricter** on people who cycle about I mean without lights and that kind of stuff
STRICTER3	S0192: >>they 're getting **stricter**
MORESTRICT1	S0428: and that he 's like his company is **more strict** so
MORESTRICT2	S0012: it 's all relevant you got to be **more strict** you 're too soft you are you 're soft say no learn to say no you ca n't though can you ?
MORESTRICT3	S0556: so your parents get **more strict** about the times you go out because it 's
FINER1	S0586: >>again --ANONnameM liking the **finer** things in life
FINER2	S0328: one of --ANONplace 's **finer** clubs
FINER3	S0208: they 're slightly **finer** but you can still tell they 're brown
MOREFINE1	S0006: now if you think about the way that jewellery is made uh to make it **more fine** they have to file it and polish it

FRIENDLIER1	S0554: it was mostly Northampton fans and I was a Northampton fan so the atmosphere was great it 's so much **friendlier** than football
FRIENDLIER2	S0686: do you think it 's **friendlier** ?
FRIENDLIER3	S0024: I look **friendlier** in that one do n't I ?
MOREFRIENDLY1	S0144: no and it 's a lot **more friendly**
MOREFRIENDLY2	S0614: >>she 's a bit **more friendly** s- I went this morning and she she is very she 's very
MOREFRIENDLY3	S0008: I suppose to make it **more friendly** (.) they 've allowed one car parking space per house (.) that parking space is half on the road half on the footpath
FAIRER1	S0684: from that system where they just did n't get the --UNCLEARWORD to comprehensive comprehensive is seen as being more equitable **fairer** for people from different backgrounds you know that they would have to mix and you would n't have you would n't determine people 's children 's erm outcomes you know like careers at the age of eleven on the basis of one test they did
FAIRER2	S0259: it 's the tills in n it ? they should be **fairer**
FAIRER3	S0198: >>yeah I guess he 's **fairer** is n't he ?
MOREFAIR1	S0198: >>that 's kind of **more fair** because then they 're like
FULLER1	S0439: fuller bust (.) cos we all know that you need a bit of a S0441: **fuller** bust
FULLER2	S0058: we got **fuller** as it went on
FULLER3	S0024: but I want to get everything prepared and I need to get some milk prepared and stuff so that I I need to start practising a bit more of a **fuller** day out (.) cos I 'm gon na really miss her when I 'm teaching
MOREFULL1	S0497: it 's a lot **more** (.) **full**
HAIRIER1	S0192: and then you get hairier and **hairier** do n't you ?
HAIRIER2	S0619: cos I realised a long (.) time ago (.) you know if I 'm fatter or **hairier** or (.) whatever (.) someone somewhere will still have sex with me so it does n't really matter
HAIRIER3	S0192: >>I mean put it this way the entire cheeks are **hairier** than that
MOREHAIRY1	S0476: my father did n't erm er I mean yes I --ANONnameM does n't gr- er --ANONnameM has this --ANONnameN y- youngness but he 's er he 's **more hairy** than the rest of us
RARER1	S0115: yeah I mean well I would have said you like there is a slight risk with with uh with **rarer** names
RARER2	S0653: >>I 'm sure it 's probably **rarer** than emerald
RARER3	S0608: oh okay Conservative Scottish ? S0428: that 's S0608: >>even **rarer**
MORERARE1	S0671: and then he erm and then er there was an opportunity to take the flick and erm like er there 's er it 's like a penalty in football only *[...]* S0671: >> **more rare**
SCARIER1	S0192: >>and I was like that 's a bit lame so we actually made an effort but we decided that AIDS mosquitoes was **scarier**
SCARIER2	S0253: but that 's kind of **scarier** because
SCARIER3	S0100: watch it (.) honestly because I 'm terrified of them so it 's even **scarier** (.) it 's not scary cos it 's funny but because it clowns
MORESCARY1	S0655: and then it 's it 's called the attack of the invincible spiders and then they s- and then wait I 've just thought of something even **more scary**
MORESCARY2	S0197: it 's much **more scary**

MORESCARY3	S0216: >>it 's cos it was the middle of the night that 's why it was **more scary** and I was in the taxi by meself cos we 've just dropped me friend off like she lives at the back of the stadium and then he just like (.) at the lights and then I would n't I would n't have even noticed and then at the lights and he just went like that and I was like shaking her and that and then I punched him in the back and then erm I got out at the end of --ANONplace and just ran (.) it 's like Jesus I just had to run off (.) it 's just ridiculous
SHARPER1	S0192: they 're perfectly evolved they were around at the same time they grow up to like seven or eight metres in length in the sea which is like three or four times the length of your dad they 're pretty big S0198: just with a bit **sharper** teeth
SHARPER2	S0104: >>yeah (.) well there 's a better one here a **sharper** one (.) but erm what I could but it 's got two people on it that were --ANONnameF 's friends that did n't really know us (.) I could cut them off at the end cos that does n't cut off well no I ca n't quite without (.) without the without them without seeing them (.) well --UNCLEARWORD how about that ?
SHARPER3	S0434: I need something **sharper**
MORESHARP1	S0416: du n no makes it just sound **more sharp** and more (.) yeah
WETTER1	S0455: no and I also think it 's **wetter** in --ANONplace than here because
WETTER2	S0618: farming (.) but I do n't know and I 've got my half welly yard boots that I bought in Newmarket from the horse shop (.) so they 'll be good on colder or **wetter** days
WETTER3	S0192: >>it 's just getting the weather 's just gon na get colder and **wetter** and darker and
MOREWET1	S0008: one of --ANONnameF 's tricks even if she even if it 's raining she 'll put the washing out (.) I said why do you put your washing out when it 's raining ? It 'll just get **more wet** (.) oh well it airs it
CLASSIER1	S0329: yeah you would pick it up if it was lying around but you would n't buy it (.) whereas you might buy it if you could just have it zipped into your inbox (.) oh thanks mister so the shirt I bought --ANONnameM is a bit like that one that that guy 's wearing except S0329: except cl- **classier** short sleeved when I was going out with --ANONnameM one he used to laugh at men in short sleeved shirts
CLASSIER2	S0382: he 's like moved out of my range there 's a vanilla one or have you got something **classier** ?
CLASSIER3	S0198: it does n't get **classier** than that
MORECLASSY1	S0530: could we do Topshop instead ? because you know (.) it 's a lot nicer S0530: and a lot **more** (.) **classy** (.) erm (.) oh oh
MORECLASSY2	S0018: I do n't know (.) I just think they look a bit **more classy**
ROUGHER1	S0012: they 're a bit **rougher** on the outside than the cucumbers though
ROUGHER2	S0358: then that started squealing and jumping around and (.) and that just made her more excited so she got even **rougher** with it
ROUGHER3	S0253: why does everyone tell me Chester 's really rough at night *[...]* S0253: how I do n't understand how rough is it S0254: just fights no **rougher** than --UNCLEARWORD really
MOREROUGH1	S0233: so if I hang around with my dad for a bit I end up getting a little bit **more rough**
MOREROUGH2	S0416: okay (.) but being a forward is just a lot **more rough**
MOREROUGH3	S0192: that 's when I got my tattoo did n't like her do n't remember her name she was a bit rough a bit **more rough** around the edges
SPICIER1	S0192: yeah the chillies the chillies *[...]* S0198: are **spicier** yeah yeah that 's
SPICIER2	S0046: it 's just that but **spicier** (.) in a naan (.) it 's lovely

SPICIER3	S0584: >>they 're like spi- they 're like **spicier**
MORESPICY1	S0115: like d' you want that or do you want uh um something **more spicy** ?
MORESPICY2	S0198: yeah I think it 's got more flavour *[...]* S0229: **more spicy**
BROWNER1	S0416: I think if you made that tree a bit (.) like (.) **browner** (.) and you took away those bubbles (.) that are erm right in here (.) then (.) then you could say the next summer (.) you know what I mean ?
BROWNER2	S0498: does who looks **browner** ?
BROWNER3	S0502: >>but --ANONnameF I think it 's cos we 're used to seeing you **browner**
MOREBROWN1	S0084: are n't they supposed to be **more brown** than that ?
MOREBROWN2	S0502: yeah --ANONnameF is **more brown**
CHUNKIER1	S0255: it seems to be very much sort of **chunkier** heels now for the boots
CHUNKIER2	S0439: I always prefer like **chunkier** straps I think they
CHUNKIER3	S0041: mm (.) no your legs are smaller than mine (.) cos I do n't really mind my legs (.) but y- mine are **chunkier** than yours
MORECHUNKY1	S0267: I do but not y- you know sort of erm not huge ones and not stilettos they 're **more chunky** but they 're still very high –UNCLEARWORD
COMFIER1	S0689: well put your arm back on there it was **comfier**
COMFIER2	S0587: >>it 's good it 's like er like they 're **comfier** than jeans
COMFIER3	S0597: I 'll I 'll sit there it 'll be **comfier** (.) could you pass me a pillow though please ?
MORECOMFY1	S0530: some are but some are really like **more comfy** most of the time for me
MORECOMFY2	S0141: do people want to go and grab a seat out a **more comfy** seat and stuff and em ?
MORECOMFY3	S0597: --UNCLEARWORD (.) I 'm gon na sit over here it 's **more comfy**
FREER1	S0278: I 'd be **freer** in prison
FREER2	*[...]* although there 's an assumption in our society that younger people are **freer** and easier about sex and and er health et cetera this turns out from a recent piece of research to be not at all the case *[...]*
MOREFREE1	S0262: so you probably sit there and tap your feet like you 're bored cos you want to get on and do something and if you 're I expect if you 're kind of transitioned into a space then either you will be a lot **more free** or you will be totally constrained and go and do something else
MOREFREE2	S0253: but erm you know just trying to be a bit **more free** and just think okay well you know (.) it 's obvious that you know all this stuff 's obvious
MOREFREE3	S0405: >>conform- conformity (.) conventions probably yeah conventions of society are **more free** (.) and (.) that 's why there is such an argument that (.) erm the Joker was (.) erm a nihilist because he (.) because he did n't he did n't obey society 's laws but (.) then you you think that he was n't a nihilist because he did believe in things like anarchy and (.) creating chaos for a purpose not just for the sake of it but (.) you know (.) comic book
HARSHER1	S0525: that 's what I thought yeah (.) they 've got quite a lot of evidence that you 'd be stupid not to plead guilty because if she pleads not guilty and gets found guilty the sentence 'll be *[...]* S0525: **harsher** (.)
HARSHER2	S0198: no (.) you should n't they 're much they 're much **harsher** on your liver to digest
HARSHER3	S0417: >>and Ireland 's quite harsh on that in *[...]* S0417: >> **harsher** than the UK
MOREHARSH1	S0456: for the thur (.) but their thur sound is not a soft one it 's sort of **more harsh**

MOREHARSH2	S0344: >>because like if you went into the ins and outs of it like the reason people are so easily erm the reason people might go oh be be maybe slightly **more harsh** and more bitchy than they would in any other situation is because they do n't really feel like --ANONnameF would behave any differently if it was them in their their shoes because obviously she 's snaked a lot of people
MOREHARSH3	S0392: over lexicalisation yeah so you 're using over lexis more letters then you 've got to we use things like hedges off which I think is where people sort (.) oh what is that one or is it softeners ? where you have things like maybe or to be honest or something to to make it sound like if you just said could I borrow your thing ? then that 's obviously a bit **more harsh** than please could I that would it be okay if I borrow or would it be okay if I
KEENER1	S0058: actually I do n't think we 've got anyone out of that (.) which I 'm quite surprised (.) I would have thought the students would have been **keener**
KEENER2	S0607: and it 's not what I 'm trying to do but it 's like making them **keener** and then I sort of think why do you wan na meet up with me ? cos I 'm awful
MOREKEEN1	S0058: >>so anyway yeah (.) so I think she yeah erm I mean I did n't know how keen she was in this one at Basingstoke which presumably she 's **more keen** on the one in Hull
MOREKEEN2	S0435: they 're much **more keen** about teaching about household wiring
MOREKEEN3	S0427: no I- I- I think for the you know the **more keen** an- or the better students that they 'll get a lot more out of it my I may have got this wrong but my impression is there is a light touch to this because I know part of it is about dropping stuff and and --ANONnameF was saying she 'll be available
CURLIER1	S0041: so she used to straighten it before she went out and then it 'd just gradually during the night S0041: --UNCLEARWORD **curlier** until by the end of the night she was Diana Ross
CURLIER2	S0415: --ANONnameF had the total curly hair *[...]* S0415: yeah t- ten times **curlier** than him
MORECURLY1	S0058: thank you (.) I 'm sure this is different to last time (.) it 's more **more curly**
EMPTIER1	S0326: I have n't booked it but probably in September it 's cheaper **emptier** and nicer
EMPTIER2	S0517: >>even though it 's **emptier** and
MOREEMPTY1	S0326: my mum her house is actually **more empty**
LUCKIER1	S0262: and that 's only that 's only if er you know if it had got a bit earlier so --UNCLEARWORD bit **luckier**
LUCKIER2	S0230: yeah but he was always **luckier** than me before
MORELUCKY1	S0008: but you were quite you know in one way you were **more lucky** cos like as far as you 're concerned you needed thirty years ' stamps
REDDER1	S0220: >>my face was getting **redder** and redder –UNCLEARWORD
MORERED1	S0423: they do well they 're **more red**
MORERED2	S0423: they do well they 're more red *[...]* S0421: >>maybe a bit **more red** than green er
RUDER1	S0525: yeah and the baby --ANONnameF was the only one that took an interest so --ANONnameM sat with --ANONnameF and and held --ANONnameM --ANONnameM held him and let --ANONnameF hold him which --ANONnameF was fine with (.) but the boys --ANONnameF completely ignored her and did n't even ask about him nothing and that was the first time she 'd met him (.) she just gets **ruder** that woman
RUDER2	S0331: that 's why people would be **ruder** at like (.) say (.) somewhere that 's (.) booked out all the time they do n't care what you think

MORERUDE1	S0330: I know he is funny but like even when he 's not like totally fucked up and like paralytic in the corner and really embarrassing everybody even even when he 's not that he 's still really really really w- like way **more rude** than everybody else I 've ever met
SEXIER1	S0439: to be honest I did say to her at least you 're the guest could you imag- I would hate if I was not that I ever would cos it not really into that kind of thing but if I was to have a threesome you have to be the guest if not that couple was gon na have it 's opening up a can of worms are n't you ? if I had a threesome and I would be like did you find her **sexier** ? was she was she better than me ?
SEXIER2	S0208: first --UNCLEARWORD that 's even **sexier**
MORESEXY1	UNKMALE[??]: if you hide it it 's more I reckon it 's **more sexy** if you hide it
MORESEXY2	S0024: yeah I think it would be nice *[...]* S0024: be a bit **more sexy** than wearing those
TIDIER1	S0198: over here just need to get and this will see what your dad brings up see how that because I mean we 're gon na he 's gon na have a bit of stuff I think I think we 're probably gon na want as least another bookshelf or something erm to put some book to like to make it **tidier** basically
TIDIER2	S0018: these boys are quite these boys are probably cleaner and **tidier** than the girls in the house actually
MORETIDY1	S0394: >>we 'll try and be **more tidy** next time you come
TRICKIER1	S0198: so stuff could come back and it 's like but it is gon na get **trickier** because your family A are very talkative and B you were at home but when we 're just the two of us in --ANONplace I do n't know how we 're gon na
TRICKIER2	S0525: no but thanks for the tip there in that I 've got a split second (.) c D (.) it 's **trickier** than you think
MORETRICKY1	S0417: ex-wife and kind of erm er (.) in the place that 's **more tricky** for me to find work and
MORETRICKY2	S0517: >>oh that 's what --ANONnameM --ANONnameM said yesterday your --ANONnameM the lifts are **more tricky** than learning to ski
MORETRICKY3	S0470: well most people do when I mention it to them but actually getting them to to do it is **more tricky**
YELLOWER1	S0653: the **yellower** pictures
MOREYELLOW1	S0653: >>or what it does S0655: things S0654: that make your face *[...]* S0654: >> **more yellow**
ANGRIER1	S0687: >>exactly there 's always another nutter ready to step up (...) you know so it was like literally it was like cut off the head it grows another head and comes back **angrier** and more vengeance with more vengeance against the rest of the world
MOREANGRY1	S0144: and that 's the scary thing (.) that 's what 's growing in this bubbling in this country (.) these people are getting more and **more angry** but thinking that represents Islam but it does n't
MOREANGRY2	S0439: and she was like yeah I know I really have to work on it and I 'm really sorry and I kept getting **more angry** and angry because I realised that you were annoyed at me and I was like I was n't annoyed with you I was just did n't care
MOREANGRY3	S0013: erm but I think the other one 's even **more angry** is n't he ?
CLOUDIER1	S0243: I think we 're gon na end up with a darker **cloudier** before so long it 's
MORECLOUDY1	S0008: and er come out again and you could see where it was and it was sort of like *[...]* S0008: bit **more cloudy** bit less cloudy bit more
CRAZIER1	S0653: >>I do n't sunbathe (.) what ? my hair purple ? I do n't think that would make me look younger *[...]* S0653: **crazier** sure

MORECRAZY1	S0542: I 'd be like where is this coming from ? and then you ca- like ca n't find it and you 'd be like getting even **more crazy**
DENSER1	S0202: but no it erm because their bone density is erm denser than ours they are actually bigger it 's a scientific fact
MOREDENSE1	S0012: yeah but that 's that 's **more dense** now if you like that 's more sort of er strict it 's a lot stricter now (.) they 've all got to be ear tagged
GREYER1	S0083: >>it 's **greyer** than it was yesterday
MOREGREY1	S0083 : >>it 's even **more grey** than yesterday
MOREGREY2	S0268: oh it 's there well they were I I thought they were very pale if they 're alsatians they 're **more grey**
MOREGREY3	S0618: cos it was **more grey** and than brown erm other birds that I know really well red kite wrong colour wrong shape
HANDSOMER1	S0416: I mean he 's **handsomer** than me by far
MOREHANDSOME1	S0253: when it was cool he 's much **more handsome** the boy in the woods is much more handsome than Peeta
LEANER1	S0192: it will get the fat out of them it 's **leaner** meaner fat grilling machine
MORELEAN1	S0511: >>it would be **more lean** ?
MADDER1	S0415: girls are a bit S0415: **madder** like that
MOREMAD1	S0519: and then --ANONnameM was peeping at him through the holes of the slide which made him **more mad** so he just goes and he would n't stop
MOISTER1	S0370: it 'll only germinate if it 's going to be cool cos it knows it 's going to be **moister**
MOREMOIST1	S0439: you do n't really taste the beetroot it just gun- gives it (.) a heavier it 's just moist (.) much **more moist** than erm *[...]* S0439: normal cake yeah
NAUGHTIER1	S0342: was n't was n't --ANONnameF **naughtier** as a as a little as a little child ?
MORENAUGHTY1	S0653: he just keeps going --ANONnameF --ANONnameF like that I think --ANONnameF 's being more and **more naughty** as she gets older cos I think she was sitting on the sofa like however when we first moved here
ROOMIER1	*[...]* I was like the three-door actually feels like **roomier** in front *[...]*
MOREROOMY1	S0520: they 're a little bit **more roomy** on the bum
STICKIER1	S0337: yeah it 's a lot **stickier** than it used to be (.) yeah but there did n't used to be chunks of erm lime on it
MORESTICKY1	S0008: they get this sort of stuff and it 's a bit like Play-Doh *[...]* S0008: but **more sticky**
SUBTLER1	S0521: >>I 'm not sure about your sarcasm --ANONnameF ? *[...]* S0521: could be a bit **subtler** I think
MORESUBTLE1	S0015: >>I know but you you that 's she she does n't you do n't want her to know that d o you ? cos you 'll go well you 're not having it obviously you can pay and you can wha tever but you kind of do it in a much **more subtle** way as a you know I I 'm I 'm really pleased you found someone who can look after you and *[...]*
MORESUBTLE2	*[...]* ANONnameF 's was a great deal **more subtle** and more attractive *[...]*
MORESUBTLE3	S0426: >>probably have n't got one exactly and planing is probably much **more subtle** and difficult than it might be