Gemination and degemination in English affixation

Investigating the interplay between morphology, phonology and phonetics

Sonia Ben Hedia

Draft of May 27, 2019, 18:04

language science press

scientific study of all aspects of the phonetics and phonology

promoting the

lab phon

lab phon

of spoken and signed languages

through scholarly exchange across disciplines

Studies in Laboratory Phonology

In this series:

1. Cangemi, Francesco. Prosodic detail in Neapolitan Italian.

2. Drager, Katie. Linguistic variation, identity construction, and cognition.

3. Roettger, Timo B. Tonal placement in Tashlhiyt: How an intonation system accommodates to adverse phonological environments.

4. Mücke, Doris. Dynamische Modellierung von Artikulation und prosodischer Struktur: Eine Einführung in die Artikulatorische Phonologie.

5. Bergmann, Pia. Morphologisch komplexe Wörter im Deutschen: Prosodische Struktur und phonetische Realisierung.

6. Feldhausen, Ingo & Fliessbach, Jan & Maria del Mar Vanrell. Methods in prosody: A Romance language perspective.

7. Tilsen, Sam. Syntax with oscillators and energy levels.

8. Ben Hedia, Sonia. Gemination and degemination in English affixation: Investigating the interplay between morphology, phonology and phonetics.

Gemination and degemination in English affixation

Investigating the interplay between morphology, phonology and phonetics

Sonia Ben Hedia

Freie Universität Berlin

Contents

Contents

Contents

1 Introduction

In English, affixation may lead to the adjacency of two identical consonants across a morpheme boundary. When in a derivative the final consonant of the prefix and the first consonant of the base are the same, a phonological double consonant emerges (see examples 1–3). The same happens when the first consonant of the suffix and the last consonant of the base are the same (see example 4). I will call these phonological double consonants MORPHOLOGICAL GEMINATES.

(1) **un**-:   *un-natural, un-known*

(2) **in**-:   *in-numerous, im-mortal*

(3) **dis**-:   *dis-satisfy, dis-solution*

(4) **-ly**:   *real-ly, sole-ly*

There are two possibilities for the phonetic realization of morphological geminates: Either the phonological double is realized with a longer duration than a phonological singleton (gemination), or it is of the same duration as a singleton consonant (degemination). It is, however, yet unclear in which cases we find gemination, and in which we find degemination.

There are numerous claims about the pattern of gemination in English affixation in the literature (see, for example, Wijk 1966: 141; O'Connor 1973: 255; Mohanan 1986: 18; Ladefoged 1993: 251; Roach et al. 2011; Wells 2008; Cohen-Goldberg 2013: 1055f.), but there is hardly any evidence for these claims. Only four studies have empirically investigated gemination in English affixed words: Kaye (2005), Oh & Redford (2012), Oh (2013) and Kotzor et al. (2016). Due to methodological issues and the small scale of the studies, their empirical findings are not sufficient to explain the gemination pattern of English affixational geminates.

As gemination in English affixation can be regarded as a morpho-phonological process which is mirrored on the phonetic level, explaining its pattern is of high theoretical importance for morpho-phonological approaches which discuss the role of phonetics in phonology and morphology. Finding out which factors govern gemination in English affixation can reveal important insights about the interplay between morphology, phonology and phonetics .

One can distinguish between two major branches of morpho-phonological approaches. The first one can be categorized as rule based and categorical in nature, while the second one is founded on the assumption that processes are gradient and dependent on the properties of individual words. Both types of approaches assume morphological boundary strength to affect the phonetic realization of complex words. It is generally assumed that weaker boundaries lead to more phonetic reduction, while stronger boundaries lead to less reduction. The two types of approaches deviate, however, in how they conceptualize these boundaries. In turn, they differ in their predictions about how morphological boundaries affect the phonetic realization of complex words, including the phonetic realization of morphological geminates.

Categorical approaches like Lexical Phonology (cf., for example, Kiparsky 1982, Mohanan 1986) assume boundary strength to depend on affixes. Affixes belong to different lexical strata which determine the phonological relation between an affix and its base. This relation is reflected on the phonetic level. For the phenomenon of gemination it is predicted that level 1 affixes, such as *in-*, are separated from their base by a weak morphological boundary and hence degeminate. Level 2 affixes, such as *un-*, in contrast, geminate due to the strong morphological boundary which they feature.

Gradient probabilistic approaches, on the other hand, would expect factors which are related to individual derivatives to govern gemination. The Morphological Segmentability Hypothesis (Hay 2003), for example, claims that the decomposability of a word determines the boundary strength between the affix and its base. This strength is assumed to be mirrored in phonetic detail, such as the duration and reduction of boundary adjacent segments. Applied to gemination, one would thus expect that more decomposable words display longer consonant durations (gemination), while less decomposable words display shorter durations (degemination).

In this book, I will test the predictions for morphological gemination made by various approaches to the morpho-phonological and the morpho-phonetic interface. On the one hand, I will test the predictions made by formal linguistic theories, which are mostly categorical in nature. On the other, I will test predictions which are derived from psycholinguistic approaches, which are mostly gradient in nature. Furthermore, I will test some general assumptions about the realization of complex words, as proposed by different models of speech production.

I will investigate morphological gemination with the five English affixes *un-*, negative *in-*, locative *in-*, *dis-* and adverbial *-ly*. The gemination pattern of each

affix will be investigated in a corpus and an experimental study. By finding out which approach can account best for the gemination pattern of English affixed words, important implications about the interplay between morphology, phonology and phonetics can be drawn.

The book is structured as follows. In Chapter 2, I will give an overview of the phenomenon GEMINATION. I will introduce key terminology, discuss the phonological representation of geminates and summarize previous work on gemination. I will mainly focus on morphological gemination in English. In Chapter 3, I will turn to the five affixes investigated in this book. I will describe the characteristics of each affix and compare them in a qualitative analysis. In Chapter 4, I will discuss the three investigated fields of morpho-phonological and morpho-phonetic approaches: Formal linguistic theories, psycholinguistic approaches to morphological processing and theories of speech production. I will summarize the main aspects of each field, discuss the most important theories in the field, and deduce the predictions each theory makes for gemination with the five affixes under investigation. These predictions will then be tested in a corpus study and an experimental study. The studies will be discussed in Chapters 5–8. While in Chapter 5 the general methodology underlying both studies will be described, Chapter 6 will focus on the methodology, analyses and results of the corpus study, and Chapter 7 will focus on the methodology, analyses and results of the experimental study. In Chapter 8, the results of both studies will be summarized and discussed with regard to the approaches discussed in Chapter 4. In Chapter 9 a final conclusion will be given.[1]

---

[1]Earlier versions of parts of Chapters 2, 5 and 6 have been previously published in Ben Hedia & Plag (2017). They were only minimally altered for the present book. The pertinent chapters and sections will be identified by a footnote.

2 Gemination

In this chapter, I will introduce and clarify the key terminology and notions necessary to understand the theoretical implications of this book. I will discuss different types of geminates and thereby show how gemination is a phonological as well as a morphological phenomenon. I will also explain the important role of phonetics in investigating gemination. After clarifying some general notions on gemination, I will concentrate on gemination in English by reviewing assumptions and previous research.

2.1 Geminates

Geminates are taken to be double consonants which are articulated with a particularly long duration (e.g. Hartmann & Stork 1972; Catford 1988; Trask 1996; Matthews 1997; Crystal 2008; Davis 2011; Galea 2016). Lexical (or "true") geminates denote a phonemic difference, i.e. they make up minimal pairs with their singleton counterparts such as in the Japanese words *kona* 'powder' versus *konna* 'such'. A second type of geminate are double consonants arising across a morphological boundary from the concatenation of two morphemes, such as in the English prefixed word *unnatural* or the compound *fun name*. For this type of geminates various labels are found in the literature, among them *fake geminates* (for example used by Hayes 1986; Oh & Redford 2012 and Kotzor et al. 2016), *derived geminates* (for example used by Kubozono 2017), *concatenated geminates* (for example used by Ridouane 2010) and *surface geminates* (for example used by Lahiri & Hankamer 1988; Galea 2016). I will refer to them as *morphological geminates*.

The main feature of geminates, distinguishing them from singletons, is their longer duration. But what is the durational difference between geminates and singletons? Acoustic research has shown that there is no universal answer to this question. The singleton-geminate ratio depends on various factors, such as the language in which the geminate occurs, the type of segment the geminate consists of and the geminate's position. With regard to cross-linguistic differences a review of empirical work on the topic reveals quite a big range of singleton-geminate ratios between languages. Stop geminates in word-medial position are,

for example, found to range from 1:1.5 in Madurese (Cohn et al. 1999) to 1:2.9 in Turkish (Lahiri & Hankamer 1988) (see also Dmitrieva 2017: 38f. for discussion of language-specific differences). Furthermore, durational differences heavily depend on the type of segment involved. For instance, Aoyama & Reid (2006) find that for Guinaang Bontok the highest ratios, i.e. the longest geminates, can be found with nasals (ratios between 1:1.72 and 1:2.15), followed by lateral approximants (ratio: 1:2.0), stops (ratios between 1:1.81 and 1:1.90), approximants (ratios between 1:1.56 and 1:1.69) and fricatives (ratio for [s]: 1:1.56). The lowest ratio is found for glides (ratio: 1:1.39). Similarly, for Italian, Payne (2005) finds the longest geminates with nasals and laterals (ratio for nasals: 1:2.1, ratio for laterals: 1:2.3) and the shortest with fricatives (ratio: 1:1.5). The influence of position is yet unclear and seems to depend on the language investigated (see Galea 2016: chapter 3 and Dmitrieva 2017: 36f. for a discussion of cross-linguistic gemination in different positions). While Ridouane (2010), for example, finds that word-final geminates are longer than word-medial geminates in Tashlhiyt Berber, Kraehenmann (2001) finds the opposite for Swiss German. For Maltese, Galea (2016) finds, similarly to Kraehenmann, word-medial geminates to be longer than word-final geminates. Studies on word-initial and word-final geminates are, however, quite rare. One reason for the low number of investigations on the topic might be that most geminates occur in word-medial position (Dmitrieva 2017: 34; Topintzi & Davis 2017: 11). Whether there is a systematic difference in duration between geminates in different positions is to be determined in further research.

In addition to duration, there are some other possible acoustic correlates of gemination discussed in the literature. In a study on Tashlhiyt Berber, Ridouane (2010), for example, shows that lexical geminates differ in their amplitude, as well as in the duration of their preceding vowel from their singleton counterparts. Geminates feature a higher amplitude and are preceded by a shorter vowel than singletons. While amplitudinal features of geminates are not well researched, the shortening of a geminate-preceding vowel was also found in other studies, such as in Lahiri & Hankamer (1988) for Bengali, in Cohn et al. (1999) for Buginese, Madurese and Toba Batak and in Galea (2016) for Maltese (see also Maddieson 1985 for discussion). However, there are also studies which did not find the duration of the preceding vowel to be affected by gemination (cf. for example Lahiri & Hankamer 1988 on Turkish, Ham 2001 on Hungarian, see also Ridouane 2010: 6 for a review of temporal acoustic attributes of gemination in different languages).

To summarize, even though there is evidence that in some languages geminates might affect the duration of their preceding vowel, as well as other acous-

tic properties, such as amplitude, the core feature of geminates is their internal, longer duration . Geminates are significantly longer than their singleton counterpart. Importantly, the singleton-geminate ratio is not universal and may vary depending on language, geminate position and the segmental features of the segment.

2.2 Morphological geminates

As mentioned in the previous section, there are two different types of geminates: lexical and morphological geminates. In English, the language under investigation in this book, lexical geminates do not exist. However, English has morphological geminates. Two adjacent identical consonants may either emerge word-internally through affixation (e.g. *unnatural*), or across a word boundary in compounding (e.g. *book case*) and in phrases (e.g. *The man naps.*). In this book, I will concentrate on gemination in English affixation. Affixational geminates emerge in prefixed words when the final segment of a prefix and the first segment of the base are identical. In suffixed words a morphological geminate emerges when the first segment of the suffix and the last segment of the base are identical. Examples of geminates with prefixed and suffixed English words are given in (1) and (2). Note that while in most cases the phonological double consonant is represented by an orthographic double, there are also some words in which the two identical consonants are interrupted by an additional character (e.g. *unknown, solely*).

(1)   *unnatural, unknown, innumerous, immortal, dissatisfied*

(2)   *really, solely, cleanness, soulless*

   While the durational features of lexical geminates are clear in the sense that they are significantly longer than their singleton counterparts, facts are less clear with morphological geminates. Since morphological geminates do not denote a phonemic difference, there are essentially two possibilities for their phonetic realization: preservation and reduction. If the two consonants are preserved, I will speak of gemination. If the two consonants are reduced, I will speak of degemination. In case of preservation one should expect a significant durational difference between a double consonant and a singleton, with the double consonant being longer. In the case of reduction, i.e. degemination, two options are possible. The first is categorical in nature: one of the two underlying consonants would be deleted, to the effect that there would be no durational difference between a singleton and the degeminated double consonant. Another option is that

degemination is a gradient phenomenon. Under this view the potential reduction of two identical consonants straddling a morphological boundary is gradual and could depend on word-specific properties, for example the morphological decomposability of the word in question. While most theoretical approaches expect gemination to be categorical, the question is yet unanswered and needs to be addressed empirically (see §4.3.1 for further discussion).

In general, morphological geminates are investigated less than lexical geminates, and only a few studies are available which empirically investigated the matter. One prominent idea tested in the available studies is whether there is a difference in the realization of geminates with different types of morphological boundaries. For example, Bergmann (2017) conducted an experimental study on gemination in German nominal compounds (e.g. *Schifffenster*, Eng. 'ship window') and particle verbs (e.g. *auffallen*, Eng. 'notice'). She found that both constructions geminate and that the degree of gemination, i.e. the duration of the double consonant, depends on accentuation, as well as lexical frequency. Duration is enhanced with low frequency words, as well as when a word bears sentence accent. The study thus shows that the realization of morphological geminates is influenced by prosodic, as well as lexical factors. Bergmann's results do, however, not support the idea that the realization of geminates is influenced by the type of morphological boundaries across which they occur. In her study there was no difference in the realization of geminates across compound-internal boundaries and word-internal geminates in particle words.

Ridouane (2010) found similar results for the influence of different morphological boundaries on geminate duration in Tashlhiyt Berber. He compared the phonetic correlates of gemination in word-initial lexical geminates with the ones in word-initial morphological geminates. The morphological geminates display the same durational differences to singletons as the lexical geminates. In other words, with regard to duration, morphological and lexical geminates are alike. However, while lexical geminates also display shorter preceding vowel durations and higher amplitudes than singletons, these secondary cues of gemination were not found for morphological geminates. These results fit in with Bergmann's, as both studies do not find durational differences depending on the morphological boundary of the geminate. However, in contrast to Bergmann, Ridouane found additional phonetic differences between geminates with different boundary strengths, suggesting that boundary strength might indeed play a role in gemination. Ridouane (2010) interprets the acoustic differences between lexical and morphological geminates as arising from differences in the underlying representation of the two different types of geminates. According to him the rep-

resentations of lexical geminates are "stronger" than the ones of morphological geminates. Therefore, lexical geminates feature, in contrast to morphological geminates, enhancing correlates (such as higher amplitudes and shorter preceding vowel durations) – a suggestion which I will discuss in more detail in §2.3.

A study conducted on Maltese word-initial geminates by Galea et al. (2014) also supports the idea that lexical and morphological geminates differ in their underlying structure and phonetic realization. Galea et al. found shorter durations for morphological geminates than for lexical geminates. While the durational differences found do not fit in with Ridouane's (2010) results, the finding that morphological geminates generally differ from lexical geminates in their phonetic realization fits in with Ridouane's idea of lexical geminates being "stronger" than morphological geminates. In contrast to morphological geminates they are not affected by weakening processes, such as phonetic reduction.

As discussed above, in contrast to the results by Ridouane (2010) and Galea et al. (2014), no difference was found between the different types of geminates in Bergmann (2017), i.e. Bergmann (2017) did not find differences in the realization of word-internal geminates in particle words and word-boundary geminates in compounds. This opens up the question of whether differences only exist between the geminates of certain types of morphological boundaries, such as morphological vs. non-morphological. However, the three studies discussed deviate from each other in many respects, so that no firm conclusions can be drawn. First, the studies investigated geminates in different positions. While Bergmann looked at word-medial geminates, Ridouane and Galea et al. investigated word-initial geminates. Second, the three studies looked at different languages. As discussed in §2.1, the realization of geminates differs between languages. Therefore, differences in results might be due to language-specific factors. A third potential cause for the deviating results might be that different types of segments were investigated. Further studies which systematically look at the influence of morphological boundary strength on gemination in different languages are needed to clarify the matter. It is especially necessary to address the question of whether only a binary distinction between morphological vs. lexical geminates exists, or whether different types of morphological geminates show differences.

For English six studies on morphological gemination exist: Delattre (1969), Kaye (2005), Oh & Redford (2012), Oh (2013), Kotzor et al. (2016) and Ben Hedia & Plag (2017). While some insights about English geminates can be gleaned from these studies, due to methodological reasons, as well as sample size, many aspects of morphological gemination in English remain unknown. I will discuss

each study in detail in §2.4.[1] Before turning to morphological gemination in English though, I will discuss the phonological representation of geminates, particularly the representation of morphological geminates.

2.3 Phonological representation of geminates

The phonological representation of geminates has been a topic of discussion for decades and "will remain an area of theoretical controversy in the foreseeable future" Davis (2011: 22). The main question of dispute is whether geminates should be represented as one or two underlying phonological segments. To understand why this question is raised, we need to take a look at the phonological properties of geminates.

According to Hayes (1986) lexical geminates are characterized by three phonological properties: ambiguity, integrity and inalterability. Ambiguity refers to the ambiguous phonological behavior of geminates. In some respects they behave as if they were two segments (e.g. their duration and their ambisyllabicity), and in some they behave as if they were one (e.g. one feature bundle). Integrity refers to the fact that geminates cannot be split up by rules of epenthesis (see, for example, Abu-Salim 1980; Kenstowicz 1994). Inalterability alludes to the geminate's resistance to undergo phonological rules that are expected to apply to its singleton counterpart, such as for example spirantization (see, for example, Kenstowicz 1994 and Kirchner 2001: chapter 5 for discussion).

The phonological representation of geminates should accommodate all three properties, i.e. capture that geminates are like two segments in some respects and like one in others. This already poses a challenge for phonological theory. The fact that geminates do not display universal behavior across languages complicates the matter further. For example, Kenstowicz (1994) notes that Icelandic geminates violate the inalterability aspect. When part of a consonant cluster, the first part of the geminate undergoes a phonological rule which shifts its aspiration to the preceding segment, i.e. the geminate is altered. The variation in the phonological behavior of geminates across languages might suggest that there is no universal representation of geminates. This view is for example taken by Ham (2001), who suggests that representations are language-specific and may even differ within one language depending on geminate position in the word.

Interestingly, the discussion about the representation of geminates mainly revolves around lexical geminates and not around morphological geminates. One

---

[1]Ben Hedia & Plag (2017) is part of this book and will be discussed in Chapter 6.

Draft of May 27, 2019, 18:04

reason is that, according to the literature, morphological geminates behave differently than lexical geminates. They allow for epenthesis, i.e. are not characterized by integrity, and undergo phonological alternations, i.e. are alterable (see also Kenstowicz 1994; Kirchner 2001 and Ridouane 2010 for discussion). Therefore, one can state that morphological geminates are not characterized by the same three phonological properties as lexical geminates. As discussed above, there is also empirical evidence for different underlying representations of lexical and morphological geminates (cf. Ridouane 2010; Galea et al. 2014). While for lexical geminates there are arguments for a single underlying representation (such as their inalterability and integrity), morphological geminates behave like two adjacent segments and are therefore commonly represented by two segments. This view ties in with Ridouane's argument of lexical geminates having a stronger representation than morphological ones.

The two most-discussed ways of representing geminates are the autosegmental representation (see, for example, Leben 1980; Hayes 1986; Levin 1985; Ridouane 2010) and the moraic representation (see, for example, Hayes 1989; Davis & Ragheb 2014; Topintzi 2008). The autosegmental representation, first proposed by Leben (1980), uses two separate tiers to capture the ambiguous structure of lexical geminates – the skeletal tier[2] and the segmental tier. While the skeletal tier represents the prosody of a structure, the segmental tier represents its segments.
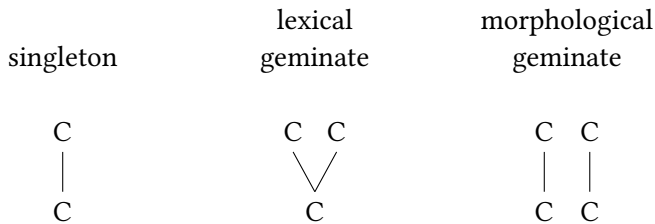
Figure 2.1: Autosegmental representation of geminates

Figure 2.1 shows the autosegmental representations of singletons, lexical geminates and morphological geminates (for similar analyses see, for example, Kenstowicz 1994: 413, Gussmann 2002: 26f. and Ridouane 2010: 62). The upper tier shows the skeletal tier and the lower one the segmental. While singletons only

---

[2]The skeletal tier is also referred to as *CV-tier* (cf., for example, Hayes 1986; Ridouane 2010; Ridouane & Hallé 2017), *X-tier* (cf., for example, Levin 1985) or *length-tier* (cf., for example, Vago & Ringen 2011). The different labels mirror differences in the approaches which are not relevant for the current book and will therefore not be discussed here.

take one slot at both levels, lexical geminates occupy one slot at the segmental tier and two slots on the skeletal tier. The two prosodic slots account for the long duration of the geminate, as well as for its ambisyllabicity. The single slot on the segmental tier represents the geminate's inalterability and integrity. Since morphological geminates differ from lexical geminates in terms of their integrity and inalterability, they take two slots on both tiers. This also mirrors their derivational nature, which naturally entails that a morphological geminate is made of two concatenated identical segments.

Differently from the autosegmental approach, the moraic approach does not entail a segmental prosodic tier on which the geminate is represented as having two slots. Instead, the root-node is directly connected to a higher prosodic structure, i.e. the mora, which represents a segment's underlying weight. Figure 2.2 shows the moraic representation of singletons, lexical and morphological geminates. While lexical geminates are underlyingly heavy, i.e. moraic, singletons are light, i.e. not moraic. Morphological geminates are represented as two identical singletons. These singletons are regarded as independent from each other and are therefore not underlyingly moraic (for similar analyses see, for example, Ham 2001: 14, Davis & Ragheb 2014: 17 and Davis 2017).

<div align="center">

singleton   lexical geminate   morphological geminate

μ
|

C   C   C C

</div>
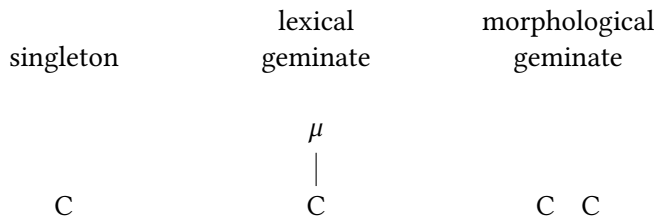
Figure 2.2: Moraic representation of geminates

Comparing the segmental and the moraic approach, it is striking that while the representation of lexical geminates differs between the two approaches, morphological geminates are represented as two underlying segments in both analyses. In this book I will adopt this view and assume that the underlying representation of morphological geminates consists of two segments. In word-medial position one of the segments is in coda position, the other forms the onset of the following syllable.

While the underlying representation of morphological geminates as two distinct segments seems undisputed, it is yet unclear how these two identical segments are realized at the phonetic level. As described in §2.2, only few studies on

morphological gemination exist. Most of these studies do not systematically investigate important lexical factors which might influence the realization of morphological geminates (e.g. morphological category and decomposability). These factors are, however, important to look at since their role in gemination can provide us with important insights about the morpho-phonological, as well as as the morpho-phonetic interface (see Chapter 4 for a thorough discussion). In this book, I will empirically investigate morphological gemination with English affixes. I will systematically test which factors influence the phonetic realization of morphological geminates with *un-*, locative *in-*, negative *in-*, *dis-* and *-ly*, and thereby contribute new insights to the understanding of morphological geminates and morpho-phonological theory.

2.4 Gemination in English

The theoretical literature only sparsely discusses morphological gemination in English. The phenomenon is mostly mentioned implicitly and rarely discussed in more than one sentence. Assumptions about gemination in English can, however, be gleaned from some theoretically oriented studies and from secondary sources such as handbooks, textbooks or pronunciation dictionaries. Additionally it is possible to deduce predictions about gemination behavior of English affixes from morpho-phonological theories and psycholinguistic approaches of morphological processing. In this section, I will restrict my discussion of English gemination to explicit mentions in the literature, as well as previous empirical studies on the topic. While I will concentrate on gemination with the affixes under investigation, I will also review general statements about gemination in English, i.e. I will take a look at mechanisms which are generally assumed to govern gemination in English, including gemination in compounds and phrases. In Chapter 4, I will then turn to prominent formal linguistic and psycholinguistic approaches, as well as prominent theories of speech production. I will discuss those approaches in detail, and deduce clear predictions about the gemination behavior of the affixes *un-*, *in-*, *dis-* and *-ly*.[3]

2.4.1 Assumptions

I will start this review by looking at how pronunciation dictionaries (e.g. Kenyon & Knott 1953; Roach et al. 2011; Wells 2008) treat morphological geminates in *un-*, *in-*, *dis-* and *-ly*-affixed words. There is a systematic difference between the

---

[3]An earlier version of this section was published in Ben Hedia & Plag (2017).

representations of the prefix *in-* and the prefix *un-* in the dictionaries. If the prefix *un-* is attached to a base starting in /n/, the word is transcribed with a long nasal (i.e. with [nː]). In contrast, if the prefix *in-* attaches to a base starting with /n/, the transcription only shows a short /n/ (i.e. [n]). The only exception is the word *innavigable* in Roach et al. (2011), where the word is transcribed with two [n]s. It is unclear what distinguishes this word from the other *in-*prefixed words. With *in-*, there is the complication that the prefix has three additional variants that may or may not involve gemination: *im-*, *ir-* and *il-*, as in *immobile*, *irresponsible* and *illegal*, respectively. In the dictionaries one consistently finds a short consonant in these cases, too. That is, all allomorphs of *in-* are taken to behave in the same way with regard to degemination.

For the prefix *dis-*, variation is found in Roach et al. (2011). While some types are transcribed with two fricatives (e.g. *dissatisfy, dissimulation*), most types are transcribed with only one [s] (e.g. *dissolution, dissemble*). It is unclear on which bases it is decided whether a type is transcribed as featuring one or two consonants. There even is variation among types of the same root. While *dissimulation* is transcribed with two [s]s, *dissimulate* is transcribed with only one. Interestingly, in Roach et al. (2011) a long fricative (i.e. [sː]) is never assigned to *dis-*prefixed words. This suggests that, according to Roach et al. (2011), morphological geminates are realized differently in *dis-*prefixed words than in *un-*prefixed words. In contrast to doubles with *dis-*, doubles with *un-* are transcribed with a long nasal (i.e. [nː]) instead of with two (i.e. [nn]). Differently from in Roach et al. (2011), in Wells (2008) there is no variation found for *dis-*prefixed words. All types are transcribed with only one /s/, suggesting that *dis-* always degeminates.

For *-ly-*suffixed words, one finds an interesting note on gemination in Wells (2008), stating that "after a stem ending in l, one l is usually lost" (451). Wells (2008) thus suggests that *-ly-*suffixed words degeminate.

Pronunciation dictionaries, which generally consider citation forms unaffected by context-specific or situation-specific influences, thus suggest gemination for *un-* and degemination for *in-* and *-ly*. For *dis-* the dictionaries do not agree. While one dictionary states that the prefix degeminates in all cases (Wells 2008), another one predicts cases in which at least some of the morphological geminates are pronounced as two consonants (Roach et al. 2011).

Turning to the pertinent phonological or morphological literature, a similar picture emerges for the prefixes *un-* and *in-*, which are the most prominent, and hence the most discussed, examples for morphological gemination in English. Wijk (1966: 141), O'Connor (1973: 255), Mohanan (1986: 18), Borowsky (1986:

119ff.), Catford (1988: 111), Kreidler (1989: 106), Ladefoged (1993: 251), Harris (1994: 18), Spencer (1996: 22), Cohen-Goldberg (2013: 1055f.), and Cruttenden & Gimson (2014) all agree that *un-* geminates. Remarks on *in-* are less frequent, and often only refer to isolated pertinent words, but those authors who mention the issue of double nasals with *in-* all agree that *in-* degeminates (Ladefoged 1993: 251; Mohanan 1986: 18; Harris 1994: 18ff. Cruttenden & Gimson 2014: 248; Cohen-Goldberg 2013: 1055f.).

The affixes *dis-* and *-ly* are far less frequently discussed in the literature. The adverbial suffix *-ly* is only mentioned implicitly in some of the literature mentioned above. In those works, derivatives with *-ly* are used as examples of words which geminate (cf. Wijk 1966: 141; Harris 1994: 23; Spencer 1996: 22). In contrast, Bauer (2001: 82), Giegerich (2012: 353) and Bauer et al. (2013: 169) claim variation with *-ly*. Some *-ly*-suffixed words are believed to geminate (e.g. *stalely* and *vilely*) and some are believed to degeminate (e.g. *fully* and *really*). For *-ly*-affixed words in which the suffix occurs after the suffix *-al*, degemination is claimed (e.g. *federally*, *globally*, *spiritually*). Bauer et al. (2013: 169) furthermore state that degemination is variable with yet some other words (e.g. *dully* and *wholly*). For the prefix *dis-*, there is no discussion which explicitly mentions the gemination behavior of the prefix. Assumptions about *dis-* can therefore only be gleaned from general statements about gemination in English, as well as from dictionary entries, which are, as described above, contradictory.

After looking at specific mentions of the affixes in the literature, let us now turn to more general thoughts on gemination in English, i.e. the general mechanisms believed to govern gemination in English. Most of the discussed literature claims that the affix involved is decisive for the phonetic realization of a double consonant. They assume gemination to be a categorical, i.e. not a gradient, phenomenon. The majority of approaches accounts for the alleged difference in gemination behavior between affixes by positing two different kinds of morphological boundary. Mohanan (1986: 18) and Borowsky (1986: 119ff.), in the framework of Kiparskian lexical phonology (Kiparsky 1982 et seq.), for example, assign *in-* to level 1 and *un-* to level 2. In this theory, level 1 affixes have weak morphological boundaries which go along with greater phonological integration with their base, including assimilation and degemination. Level 2 affixes, in contrast, form strong boundaries with their base and are phonologically less integrated. Hence, gemination is expected for level 2 affixes. Similar in spirit is Harris' (1994) account, in which the author distinguishes between root affixation (for *in-*) and word affixation (for *un-* and *-ly*). In root affixation, generally one phoneme is deleted when two identical segments immediately follow each other. Cohen-Goldberg

(2013) attributes the alleged difference in gemination between *in-* and *un-* to their difference in productivity: the less productive prefix *in-* degeminates, while the more productive *un-* geminates. Giegerich (2012: 354) proposes that gemination depends on the status of the derivative. The phonological word status, similar to productivity, mirrors morphological boundary strength. Only derivatives with weak morphological boundaries form prosodic words (see §4.2.3 for discussion). According to Giegerich (2012), geminates only occur across prosodic word boundaries, which is why most *-ly*-derivatives, which form a prosodic word with their base, degeminate.

Bauer et al. (2013: 169) describe gemination as less predictable, i.e. not solely predictable by the affix involved. They, in contrast to the approaches discussed above, assume variation within the gemination pattern of one affix. Furthermore, the authors point at possible variables, other than the affix, which could have an effect on gemination (e.g. speech tempo and the speaker). Similarly Giegerich (1992: 191, 288) notes the effect of speech mode by stating that geminates are usually simplified in connected speech.

One can summarize that most of the theoretical literature, as well as pronunciation dictionaries, assume gemination to be affix-dependent. Different boundary strengths between affixes are assumed to cause differences in gemination. While the nature of the boundary deviates between different approaches, the main idea is that stronger boundaries lead to gemination and weaker boundaries lead to reduction, i.e. degemination. Only few sources assume that factors other than the affix involved influence gemination, and that variation in gemination can be found in words compromising the same affix. In Chapter 4, I will return to the different ideas proposed and discuss them in more detail.

Summarizing the affix-specific predictions, one can state that for *un-* and *in-*, pronunciation dictionaries, as well as the majority of the theoretical literature, agree that the former geminates, while the latter degeminates. Less is said about *dis-* and *-ly*. For *dis-*, dictionaries make contradicting assumptions and the theoretical literature is silent about its gemination behavior. For *-ly*, one also finds contradicting predictions. While Wells (2008) suggest degemination takes place, the majority of the theoretical literature claims gemination with *-ly*. Some of the literature predicts variation.

2.4.2 Previous empirical work

Apart from Ben Hedia & Plag (2017), which is given in Chapter 6, there are five studies on gemination in English: Delattre (1969), Kaye (2005), Oh & Redford (2012), Oh (2013) and Kotzor et al. (2016). The first study by Delattre (1969)

looked at word-boundary geminates. Delattre compared the duration of double consonants at word boundaries (such as the double nasal in the sentence "*I've seen Nelly*") with word-final singletons (such as the nasal in "*I've seen Elly*") and word-initial singletons (such as the nasal in "*We see Nelly*"). He investigated three different segments: /n/, /l/ and /s/. For all three he found that word-boundary geminates are longer than singletons. The nasal showed the highest degree of gemination (singleton-geminate ratio: 1:1.5). The fricative and the lateral had a singleton-geminate ratio of 1:1.3. Interestingly, the duration of the preceding segment did not vary depending on the number of consonants. In other words, gemination did not affect preceding vowel duration.

Even though this study shows that morphological gemination in English may lead to longer durations, it can only be regarded as a first clue to understand gemination in English. Since Delattre solely looked at word-boundary geminates, i.e. not at other types of morphological boundary, it remains unclear which role morphology, and consequently different boundary strengths, actually plays in the realization of geminates. Furthermore, there are major methodological issues. Delattre's results are based on only a few types and it is thus unclear which role type-specific effects might have played. The study did furthermore not account for possibly intervening factors such as, for example, speech rate. Another drawback of the study is its lack of appropriate statistics. Taking all of these drawbacks into account one can nevertheless assume that morphological gemination at word boundaries in English does, at least in some cases, lead to gemination.

Kaye (2005) and Oh & Redford (2012) both empirically investigated gemination with the two English prefixes *un-* and *in-*. In both studies the gemination of *in*-prefixed words was investigated by looking at words that featured the allomorph *im-*. The reason for this is that there are very few *in*-prefixed words with a base starting in /n/ (such as *innumerous*), i.e. there are not enough different types to empirically investigate gemination with *in-* (see §3.1.2 for further discussion).

Kaye (2005) investigated only two *un*-prefixed types (*unknown, unnamed*) and one *in-* prefixed type (*immature*). In an elicitation task, ten speakers produced these words, as well as the words' bases in isolation. Kaye then compared the durations of the nasals in the different words. The results indicate that both prefixes geminate. The [n] in *unknown* is longer than the [n] in *known*, the [n] in *unnamed* is longer than the [n] in *named* and the [m] in *immature* is longer than the [m] in *mature*. Kaye notes, however, that whether an *in*-prefixed word geminates or not depends on the individual speaker. Not all speakers produced the prefixed words with a longer nasal than the base. However, since Kaye did not apply any statistical analyses (beyond computing averages) and only investigated a very

limited number of types, the results are somewhat inconclusive. What we can see, however, is that Kaye's empirical data go against the claim that *in-* always degeminates.

Oh & Redford (2012) also investigated gemination with *un-* and *in-*, as well as gemination at word boundaries (e.g. *dim morning, one nail*). With regard to gemination with *in-* and *un-*, they compared the duration of morphological geminates with the duration of assumed phonological singletons in words starting with similar phonemic strings. The authors investigated 16 different words which contained two consonants in the orthographic representation. The items were categorized by Korean speakers (i.e. speakers of a language that has phonological geminates) who rated the duration of the nasals as either single or double, based on an English native speaker's pronunciation of these words. The words *immovable, immoral, immemorial, immeasured, unnoticed, unnamed, unnerve, unnail* were categorized as containing a double nasal, while *ammonia, immensely, immunity, immigrational, annex, innate, annoyed, innerve* were categorized as words containing a single nasal. Additionally, Oh and Redford included word-boundary geminates (e.g. *dim morning, one nail*) in their data set to investigate potential differences between word-internal and word-boundary geminates. All items were put into carrier sentences and read by eight participants in two different conditions (normal speech vs. careful speech).

With regard to word-internal geminates the analyses showed that the items rated by Korean speakers as having double nasals were longer in duration than items rated as having single nasals. This indicates that at least some words with the prefix *in-* show gemination. However, there is variation in the gemination pattern of *in-* found by Oh & Redford (2012): the set of words with singletons mainly contains words that are morphologically simplex, but some words are not simplex. The word *immigrational*, for example, is prefixed (compare *migration, immigration*), which in turn means that in this word, *in-* degeminates, while in the other prefixed words it geminates. Note also that *in-* in the word *immigrational* (like, arguably, *innate* 'existing in a person […] from birth', *OED online*, s.v. 'innate, adj.'), has a locative meaning. Incidentally, both words in which we find *in-* as a locative prefix ended up in the set of words that do not geminate, while the words with negative *in-* showed gemination. This might hint at a systematic difference between locative and negative *in-*, an issue that has so far never been discussed in the theoretical literature.

The study also reveals general differences between the two prefixes *un-* and *in-*. Oh & Redford (2012) found a difference in absolute nasal duration between the two prefixes. The nasal in *in*-prefixed words is significantly shorter than the

nasal in *un*-prefixed words. This difference is more prominent in careful speech than in normal speech. The durational difference between the prefixes vanishes, however, in relative duration. Relative duration refers to the nasal duration relative to the duration of the preceding vowel. This means that not only nasal duration differs between the two prefixes but that there is also a difference in preceding vowel duration. The prefix *un*- features a longer vowel than the prefix *in*-. In other words, *un*- is generally longer than *in*-. In addition to prefix duration, Oh & Redford (2012) also found non-durational differences between *un*- and *in*-. In careful speech, speakers sometimes inserted a pause between the two nasals of *un*-prefixed words, whereas a pause was never inserted in *in*-prefixed words. The authors interpret the inserted pause as a boundary cue.

Turning to word-boundary geminates, results revealed that there was no durational difference in absolute duration between word-internal and word-boundary geminates. There was, however, a difference between the two types of geminates in relative duration. Word-boundary geminates were shorter than word-internal geminates, and as long as singletons in relative duration. In other words, while geminates in *un*- and *in*-prefixed words are longer than singletons in absolute and relative duration, word-boundary geminates are only longer than singletons in absolute duration.

Oh and Redford interpret their results as evidence that word-boundary geminates are represented differently than word-internal geminates, i.e. that gemination is influenced by boundary strength. Furthermore, they argue that the prefix *un*- might be represented differently than the prefix *in*-. They base their argument on the finding that, even though there does not seem to be a systematic difference in the gemination behavior of the two prefixes, the two prefixes display differences in their overall duration, as well as differences with regard to non-durational boundary cues. Oh & Redford (2012) venture the idea that different affix representations emerge from differences in boundary strength (cf. Kiparsky 1982; Mohanan 1986, see §4.2.1 for discussion), or from differences in productivity and segmentability (cf. Hay 2003, see §4.3 for discussion). Since the prefix *un*- has a stronger boundary than *in*-, and since it is more productive and segmentable, there is less reduction with *un*- than with *in*-, i.e. it is longer and features boundary cues.

To investigate the relation of different morphological boundaries and gemination further, in a follow-up study Oh (2013) investigated whether there is a difference between geminates at compound-internal boundaries (e.g. *homemade*) and geminates at word boundaries (e.g. *room maid*). The comparison of geminate durations and the duration of singletons at word boundaries (e.g. *dough*

made) showed that both types of geminates are longer than singletons. A significant difference between compound-internal and word-boundary geminates is only found in careful speech, and only in absolute consonant duration. In accordance with the idea that word-boundary geminates have a stronger morphological boundary than compound-internal geminates, in careful speech word-boundary geminates are longer. In normal speech the difference vanishes. For relative duration, there never is a difference between compound-internal and word-boundary geminates. Oh (2013) interprets the results as there not being a difference in the realization of compound-internal and word-boundary geminates in English.

Thus, while in Oh & Redford (2012) there might be some indication for effects of boundary strength on gemination (difference between word-internal and word-boundary geminates in relative duration), we do not find these effects in Oh (2013). There are various possible explanations. It might for example be that only certain boundary differences lead to differences in gemination. This explanation ties in with the results on morphological gemination in German, Tashlhiyt Berber and Maltese discussed in §2.2. These studies also showed that, while gemination differed between some types of boundaries, it did not between others. Another explanation for the deviating results might be related to the studies' methodologies. Both studies only investigated a very small number of types, which made it impossible to test the influence of type-specific factors on duration. Furthermore, there might be other intervening factors such as, for example, speech rate and speaker, which influence duration, and which were not systematically taken into account. With regard to methodology, another crucial aspect must be considered when interpreting the results. The studies show differences in effects depending on the acoustic correlate used as the measure of gemination. Effects are different for absolute and relative duration. There are also differences depending on speech condition, i.e. careful vs. normal speech. Especially the different outcomes depending on speech condition suggest that factors related to the experimental set-up influence results of durational studies to a great degree. It is yet unclear how to interpret the relation between the different outcomes and the different conditions. To shed light on the matter, it is necessary to conduct further studies which systematically compare experimental data with conversational speech, and which control for intervening factors in a systematic way. More advanced statistics, which are able to tease apart different effects, are necessary to explain which factors cause which durational differences, and whether it is indeed boundary strength which leads to differences in gemination between different words and constructions.

The most recent study on gemination in English looked at gemination in suffixation and compounding. Using a reading experiment, Kotzor et al. (2016) investigated the two suffixes *-ly* and *-ness*, as well as compounds with sonorant ([l] and [n]) and stop geminates ([p], [t] and [k]). With respect to the suffixed data, they compared the double consonant with the singleton in the pertinent base word to which the suffix *-er* was added. Hence, the double consonants in *-ness-* and *-ly-*suffixed words (e.g. /nn/ and /ll/ in *coolly* and *meanness*) were compared to the pertinent singletons in *-er*-suffixed words (e.g. /n/ and /l/ in *meaner* and *cooler*). The doubles in compounds (e.g. /nn/ in *pine nut*) were compared to singletons in similar compound words (e.g. /n/ in *pineapple*). The results reveal that both types of geminates, i.e. suffixational and compound geminates, are longer than the pertinent singletons. There was no effect of the geminate on the preceding vowel duration, neither for the suffixes, nor for the compounds. Thus, Kotzor et al. (2016) did not find a significant difference between the gemination of suffixes and compounds.

Unfortunately, Kotzor et al. (2016) do not provide a separate analysis for each of the two suffixes. In other words, their analysis does not provide the possibility to state whether the suffixes *-ness* and *-ly* behave differently with regard to their gemination behavior. This can be regarded as a major drawback of the study since, as described in §2.4, it is commonly assumed that gemination is affix-dependent. From a theoretical point of view, there is good reason to assume that the two suffixes *-ly* and *-ness* do not behave identically. This idea is further supported by the fact that geminate duration varies among different types of consonants (cf. §2.1). The lateral in *-ly*-words is expected to inherently show different durations than the nasal in *-ness*-words. Furthermore, there might be structural differences between the two affixes, such as their segmentability and boundary strength, which might lead to different behavior regarding gemination. An additional problem with the study is the limited number of types investigated. For each affix only six types with a morphological geminate were included, making it impossible to investigate type-specific effects such as word-form frequency or a word's individual decomposability. Thus, even though the study provides some evidence for the gemination of *-ly* and *-ness*, one must be cautious to interpret the results. Further research on the individual suffixes incorporating a greater number of different types is necessary to make reliable statements about their gemination behavior.

To summarize, previous research on gemination in English leaves us with a number of unsolved problems. First, there is only little empirical evidence available and the few studies which do exist differ significantly in their methodol-

ogy and the constructions they investigated. Hence, the facts essentially are unclear. Only three studies looked at affixational gemination in English, i.e. the phenomenon under investigation in this book (Kaye 2005; Oh & Redford 2012; Kotzor et al. 2016). Two of these studies (Kaye 2005; Oh & Redford 2012) call the assumption that *in-* degeminates into question, and hence demonstrate the need for further testing of widely-held beliefs about gemination in English.

Second, existing empirical studies are rather limited in their data sets and consider only words spoken under experimental conditions, i.e. in isolation or in carrier sentences. What is lacking is data from natural speech. As pointed out in the literature (e.g. Giegerich 1992; Bauer et al. 2013), and as evidenced by Oh & Redford (2012) and Oh (2013), the mode of speech might significantly influence the realization of morphological geminates. Therefore, it is necessary to investigate the phenomenon in various conditions, making it possible to compare natural speech with experimental data. This will allow us to find out which factors influence gemination on which level.

Third, existing studies have not simultaneously considered different influences which might affect gemination, but rather concentrated on one specific aspect, i.e. they neglected other possibly intervening factors. None of the studies described above looked at word-specific factors, such as word-form frequency or a word's individual decomposability. Even though, as shown in §2.4, most claims about gemination are based on morphological categories, morphological factors were only considered sparsely in previous research. While most studies pointed out that different morphological categories, such as different affixes or derivatives with varying morphological boundary strength, might differ in their gemination behavior, their methodology was insufficient to shed light on the matter. The studies comparing *un-* and *in-*, for example, just assumed a categorical difference in boundary strength between the two affixes. This assumption needs to be empirically investigated. Furthermore, the investigation of *in-* did not consider that there are two different *in*-prefixes, i.e. locative and negative, and that there are potential differences between them. The study on the two suffixes *-ly* and *-ness* did not differentiate between the two suffixes at all, just assuming a similar behavior of both of them.

To conclude, previous research hinted at some interesting effects, such as the gemination of affixes which are assumed to degeminate, and differences in gemination depending on the type of morphological boundary involved. However, due to mainly methodological reasons further research is needed to clarify the facts.

Draft of May 27, 2019, 18:04

3 The Affixes under investigation

In this chapter I will describe the five affixes *un-*, negative *in-*, locative *in-*, *dis-* and *-ly* using the relevant morphological and phonological literature. I will discuss their phonological behavior, as well as important morphological, semantic and lexical properties, and compare them with regard to these factors.

Before discussing the affixes, I will give a brief overview of the factors which lead to their inclusion in this study. The prefixes *un-* and *in-* were included because of two reasons. First, they are the most prominent examples of gemination/degemination given in the literature (see discussion in §2.4). Second, they are investigated empirically in two previous studies, i.e. a comparison to previous results is possible. To investigate an additional prefix, and to also take a look at gemination in suffixes, the affixes *dis-* and *-ly* were added to the data set. The choice to include these two affixes was on the one hand due to their comparatively high type frequency, which allows for testing type-specific effects on gemination, and on the other due to their phonological and morphological make-up. The five affixes partly overlap in their characteristics, such as their semantics, their prosodic make-up and their segmentability. Importantly, they also show some major differences in their features. This combination of similarities and differences between the affixes makes it possible to test various factors which potentially affect across affixes.

In the following, I will describe the formal and structural characteristics of each affix, lay out their phonological and prosodic behavior, and discuss their semantics. Since the literature is often not very specific, or even contradictory when discussing certain aspects of affixation (e.g. stress pattern or productivity), it is often not possible to give a clear-cut description of a certain affix and its behavior. I will remain neutral regarding most of the controversial issues but lay out the different possibilities, as found in the literature.

After looking at each affix in isolation, I will compare the five affixes with each other. It is of prime importance to look at the differences between the affixes, since these differences lead, according to the theories discussed, to different predictions for their gemination behavior. I will pay special attention to boundary strength as it is one of the most important factors for predicting gemination.

At the end of this section, I will provide some insightful figures regarding the scope of the phenomenon across affixes. In other words, I will lay out how many types with a morphological geminate exist for each affix.

3.1 Description of the affixes

3.1.1 The prefix *un-*

The affix *un-* is a native prefix which takes native and non-native bases. According to Bauer et al. (2013: 355, 361, 371ff.), its bases can be verbs, nouns and adjectives. The prefix rarely changes the category of its base and does not take bound roots. It is one of the most productive prefixes in English and has a clearly negative meaning, which varies slightly depending on the base it attaches to. On verbs its meaning is reversive (e.g. *screw* vs. *unscrew*) and sometimes privative (e.g. *dress* vs. *undress*), on adjectives its meaning is contrary or contradicting (e.g. *cool* vs. *uncool*), and on nouns it mostly is privative (e.g. *faith* vs. *unfaith*). In general, the meaning of *un-*derivatives is very transparent. Only few forms exist in which one cannot deduce a derivative's meaning by adding the negative meaning of *un-* to that of its base (e.g. *unorthodox*).

Turning to the prefix's phonological attributes, *un-* shows what can be called 'optional assimilation'. This assimilation affects *un-*derivatives with a base starting in a bilabial (e.g. *unplugged, unbreakable, unmarried*) and *un-*derivatives with a base starting in a velar plosive (e.g. *ungrateful, uncool*). Before bilabials the prefix-final nasal sometimes is realized as [m]. Before velar plosives the prefixal nasal sometimes is velarized, i.e. realized as [ŋ] (cf. Hanote et al. 2010: 5f. Bauer et al. 2013: 180; Okada 2013: 125). This optional assimilation is not mirrored in orthography and, according to Okada (2013: 125), is purely phonetic, i.e. not phonological, in nature. Stockwell & Minkova (2001: 87f.) explain the optional assimilation of *un-* by a conflict between "ease of pronunciation" and "transparency". While the assimilation of the nasal makes pronunciation easier, the transparent nature of the prefix blocks its total assimilation. Except for one study (Hanote et al. 2010), there is, according to my knowledge, no empirical work on the assimilation of English *un-*. Hanote et al. (2010) found no systematic pattern with regard to assimilation with *un-*. However, since the authors restricted their study on dictionary data it is questionable how generalizable their results are. As Raffelsiefen (1999: 138), for example, proposes, assimilation with *un-* might be sensitive to register. According to her, assimilation is most likely to occur in fast, casual speech. This would mean that dictionary data is not suited to investigate the matter. To

conclude, *un-* sometimes assimilates but the pattern of its assimilation is yet unclear.

Before discussing the stress status of *un-*, a general note on prefix stress is in order. The stress of prefixes is not well researched and observations are mostly anecdotal, i.e. rely on individuals' intuitions rather than on empirical studies. Furthermore, as evidenced in Videau & Hanote (2015), prefixal stress depends on various factors, such as semantic transparency, syntactic context and extra-linguistic context. It is yet unclear how these factors interact, i.e. it is unclear how exactly they influence stress with prefixes. Even when leaving contextual influences aside, i.e. when concentrating on lexical stress in isolated words, it is quite difficult to determine the stress pattern of prefixed words.

The determination of stress for monosyllabic prefixes, such as *un-*, *in-* and *dis-*, is of special difficulty. When followed by an unstressed syllable, they are, due to prosodic constraints, usually taken to be stressed. Matters are, however, more complicated when the prefix is followed by a stressed syllable. In those cases it is very challenging to determine the relative prominence relations between the prefix and the base-initial syllable, i.e. it is very hard to determine whether the prefix is stressed or not. While it is generally assumed that prefixes can bear stress, it is unclear how stress patterns in those cases. In other words, it is unclear whether the prefix is unstressed, whether it bears primary stress or whether it bears secondary stress. As discussed in Okada (2013: 126) for *un-*, *in-* and *non-*, there is variation within prefixes, and this variation "leads linguists to different descriptions". To summarize, due to difficulties in determining stress and the not well researched variation in prefix stress, the descriptions of prefixal stress deviate between sources for the prefixes *un-*, *in-* and *dis-*.

The available descriptions of stress with *un-* reflect the problems just pointed out. While Allen (1978: 4) notes that the prefix never bears stress, other sources such as Jespersen (1965: 464f.) and Okada (2013: 126) state that *un-* is a stress-preserving prefix that may carry stress. Similarly in Wells (2008) *un-* can be stressed. Except for Allen (1978), all sources agree that *un-* is stressed if the base-initial syllable of the pertinent derivative is unstressed (e.g. ˌunconˈventional[1]). Assumptions deviate for those cases in which the first syllable of the derivative's base is stressed (e.g. *unˈjust*). Jespersen (1965: 464f.) notes for these cases that, while in some of the most common *un-*derivatives the prefix is unstressed (e.g. *unˈcommon, unˈhappy*), in most cases the prefix is stressed (e.g. *ˈunˈaided, ˈunˈjust*). In Wells (2008: 808) it is noted that, while generally the prefix may be either

---

[1]Throughout this book I will use ˈ to mark primary stressed syllables and ˌ to mark secondary stressed syllables.

stressed or unstressed in those cases, verbs always have a stressed prefix (e.g.
ˌun'coil). Furthermore, it is noted that *un-* "is unstressed particularly where it is
not a true prefix (*un'wieldy*)" (808). Wells does not give a definition of true pre-
fixhood though, i.e. this statement is quite unclear. Divided or uncertain usage
is annotated for some of the *un*-prefixed words with a stressed base-initial sylla-
ble (e.g. ₍ˌ₎*un'bearable*, ₍ˌ₎*un'leash*). As evidenced in a study by Hanote et al. (2010:
2ff.), the assignment of prefixal stress with *un-* in Wells (2008) does not follow
any systematic pattern. One can thus state that the literature does not provide
clear, systematic and empirically-based criteria for stress in derivatives with a
stressed base-initial syllable. It therefore remains unclear when *un-* is stressed
and when it is unstressed.

One can summarize that *un-* is a very segmentable and very productive prefix.
Its meaning is clearly negative and transparent in the vast majority of derivatives.
The fact that *un-* only optionally assimilates can be interpreted as a result of its
high segmentability which blocks total assimilation, and which hence ensures
the prefix's phonological independence from its base. With regard to stress, one
can state that *un-* can bear stress but that the stress pattern of the prefix is yet
unclear.

3.1.2 The prefix *in-*

When investigating the prefix *in-*, one must acknowledge the existence of two
different *in*-prefixes in English: negative *in-* and locative *in-*. While the existence
of negative *in-* is uncontroversial, the idea of locative *in-* may not be as straight-
forward. The reason is that locative *in-* often occurs in derivatives with bound
roots, such as *inject* or *infuse*. In these words the discrete meaning of *in-*, as well
as the discrete meaning of the base, is often unclear. This might lead to the view
that locative *in-* is not a prefix but some sort of unit below the word level with
no clear semantic content. This view, however, neglects that locative *in-* does
indeed have a stable meaning.

Let us take the standard methodological approach to morphological categories
according to which an affix should have an identifiable, stable meaning across
different words (cf., for example, Plag 1999: chapter 5.2.2; Stockwell & Minkova
2001: 63ff. Schulte 2015: 68). Under this approach, we would consider *in-* a loca-
tive prefix in all those words (and only in those) where the word-initial string *in-*
can be assigned some locative meaning and where at the same time the remain-
ing string is also attested outside that word with a stable, identifiable meaning.
Implementing this method, we would be able to assign some locative meaning
to the string *in-* in words such as *infuse* 'to pour in', *implant* 'to plant in' and *im-*

port 'to bring in' (*OED online* paraphrases). The remaining strings, i.e. the bases, in these words are all attested (either as words or as bound roots) outside these words with sufficiently similar meaning (cf. *transfuse*, *plant*, *export*). This small sample thus shows that, at least in some words, there is a locative prefix *in-*.

In this study, all words in which the affix and the base carry a stable meaning are considered as morphologically complex. Since, complex words with both *in*-prefixes exist, i.e. negative and locative *in-*, both types of *in-* are included in the study. Below I will take a closer look at each type of *in-*.

3.1.3 Negative *in*-

Differently from *un-*, negative *in-* is a non-native (or Latinate) prefix that takes non-native bases. It mostly takes adjectives as its base (e.g. *intolerant, immortal*). Only sometimes nouns are taken as its base (e.g. *inexperience*). In most cases, negative *in-* does not change the word class of its derivative. Usually free bases are found with negative *in-* (e.g. *intolerant, impossible*), but in some derivatives bound roots are found (e.g. *inept, innocent*). In these words it is, however, questionable whether *in-* can even be considered a prefix (cf. Bauer et al. 2013: 356f., 611).

The majority of the literature claims that negative *in-* is not productive (cf. Bauer & Huddleston 2002: 1688). A search of hapax legomena in the Corpus of Contemporary American English (Davies 2008–2014), as carried out by Bauer et al. (2013: 361), reveals however that there is indeed a number of new derivatives with negative *in-*. Examples are *immedical, inactual, inconservative* and *inextractable*. One should note, however, that negative *in-* is far less productive than *un-*, which carries similar meaning. Both prefixes denote contrary or contradictory meaning on adjectives. Because of these similarities in meaning the prefixes are often treated as rivals. It is often claimed that the existence of a derivative featuring one of the two prefixes blocks the existence of a derivative with the other. According to this view, the existence of the word **inhappy*, for example, is blocked by the existence of the word *unhappy* (see, for example, Jespersen 1965: 467ff. Bauer & Huddleston 2002: 1688f. and Bauer et al. 2013: 377ff. for discussion). As reported by Bauer et al. (2013: 377), the idea of rivalry between the two prefixes is, however, invalidated by the existence of derivatives such as *inaccessible* and *unaccessible*. There is no noticeable difference in meaning between the two derivatives, and in turn there is no difference in meaning between the two prefixes *un-* and *in-*. However, while the two prefixes often are identical in meaning, Bauer et al. (2013: 379), similar to Jespersen (1965) and Bauer & Huddleston (2002), note that *in*-derivatives are more often lexicalized than derivatives

with *un-*. To summarize, while negative *in-* is very similar to *un-* in its meaning, it does not share its high degree of productivity and its consistency with regard to semantic transparency.

Turning to the phonological features of negative *in-*, one must mention its obligatory assimilation. Before a bilabial the prefixal nasal becomes /m/. Before a base-initial lateral the nasal becomes /l/, and before an alveolar rhotic it becomes /r/. Examples are *impossible*, *immortal*, *illogical* and *irregular* (cf. Bauer & Huddleston 2002: 1687; Bauer et al. 2013: 359; Okada 2013: 123f.). This assimilation is also mirrored in orthography.

With regard to stress, *in-*, similar to *un-*, is generally assumed to be stressed when followed by an unstressed syllable (cf. Jespersen 1965: 473; Wells 2008: 381, 384; Bauer et al. 2013: 183; Okada 2013: 126). Examples are ˌinexˈhaustible, ˌindeˈscribable and ˌimmeˈmorial. In those cases *in-* bears secondary stress. There are also some derivatives, such as ˈimpotent and ˈimpious, in which *in-* is primarily stressed. Bauer et al. (2013: 183) note that while stress with *in-* seems unsystematic in general, the prefix seems to only bear primary stress when the base of the derivative is bound. In case of a stressed base-initial syllable, *in-* is, in contrast to *un-*, generally taken to be unstressed (cf. Wells 2008; Bauer et al. 2013: 183; Okada 2013: 126).

We can summarize that, while negative *in-* is similar in meaning to *un-*, it is less semantically and phonologically transparent. It is also less productive. Negative *in-* has more bound roots than *un-*, is sometimes semantically opaque, and derivatives are more often lexicalized. The prefix is phonologically more integrated in its base than *un-*. This is indicated by the obligatory assimilation of the prefix. The stress pattern of *in-* can also be interpreted as an indicator of its phonological integration. While in some respects the stress pattern is similar to the one of *un-*, the prefix, contrary to *un-*, also bears primary stress in some cases. Primary stress emerges with bound roots and denotes stress shift, i.e. in those derivatives the position of primary stress has shifted from the base to the prefix. This shows that the boundary between prefix and base is rather weak.

3.1.4 Locative *in-*

Locative *in-* differs in interesting respects from negative *in-*. Differently from negative *in-*, the origin of locative *in-* is twofold. This leads to a prominent distinction in the literature. Traditional approaches differentiate between non-native (or Latinate) locative *in-* and native locative *in-*. While the first one is commonly analyzed as a prefix which takes non-native bases, the second is often analyzed

as a locative particle (cf. Jespersen 1965: 497ff. Marchand 1969: 115,163f. Bauer & Huddleston 2002: 1685). Examples are given below:

(1)   non-native locative *in-*:   *inject, immigrate, impress*

(2)   native locative *in-*:   *inborn, inbreed, indoors*

Bauer et al. (2013) also make a distinction between the two different cases but regard both types of locative *in-* as prefixes. Since the authors only discuss productive prefixes, and since they consider non-native locative *in-* unproductive, only native *in-* is discussed in their work. It is said to be a productive prefix that attaches to nouns, adjectives and verbs. It takes native and non-native bases (Bauer et al. 2013: 334, 340).

It is generally assumed that the two types of locative *in-* differ in their type of base, their productivity and in the degree of phonological assimilation they display. While the non-native prefix is believed to mostly take non-native bases (see Jespersen 1965: 499), native *in-* also attaches to native bases (see Bauer et al. 2013: 334). While non-native locative *in-* is not productive, the native form is believed to be productive. With regard to phonological assimilation, non-native locative *in-* undergoes the same obligatory assimilation as negative *in-*, while native locative *in-* is described as only optionally assimilating (see Jespersen 1965: 499; Bauer et al. 2013: 335). The common assumption thus seems to be that non-native locative *in-* is phonologically and semantically less transparent than native locative *in-*. There are, however, no empirical studies investigating whether these criteria mirror the distinction between the two types of locative *in-*.

It is not easy to distinguish between native and non-native locative *in-*. This problem is already discussed in early works such as Jespersen (1965: 499) and Marchand (1969: 164). The semantics of both types of locative *in-* overlap to a high degree. Both prefixes denote a locative meaning, which is formulated as 'into, in, within; on, upon; towards, against' in the *OED online*. The *OED* further mentions that

> "[s]ince *in-* prefix 1 [non-native *in-*] and *in-* prefix 2 [native *in-*] are identical in form, and to a great extent in sense, they come in later use to be felt as one and the same prefix; and it is this resulting prefix which appears in many words of later formation."

It is thus suggested that, while locative *in-* has two origins, the distinction between the two formerly distinct types of locative *in-* is no longer clear. This is in line with the observation that the criteria used to distinguish native from non-native locative *in-* are rather blurry. In this study, I will not explicitly distinguish

between the two types of locative *in-*. I will, however, account for differences in semantic transparency, phonological make-up and type of base in my studies. I therefore will be able to implicitly differentiate between more decomposable and less decomposable locative *in-*derivatives, which in turn, according the literature, may mirror the distinction between native and non-native locative *in-*.

With regard to stress, the literature is mostly silent about a distinctive stress pattern for locative *in-*. Except for Wells (2008: 384), who notes that *in-* is "generally stressed only (i) if meaning 'in' rather than 'not'", there are no explicit mentions about the stress pattern of locative *in-*. While, according to Wells (2008), locative *in-* is stressed and negative *in-* is not, there are also sources which state that negative *in-* can be stressed (see description of negative *in-* above). It is thus unclear whether there is a systematic difference between the stress pattern of locative and negative *in-*. One could venture the idea that locative *in-*, due to its assumingly higher frequency of bound roots, more often bears primary stress than negative *in-*. This would fit in with Wells's statement. It, however, is an assumption that calls for empirical evidence.

To summarize, locative *in-* is of two different origins which makes it rather difficult to make clear statements about the prefix. It seems that it occurs quite frequently with bound bases, is often semantically opaque and less productive than negative *in-*. In other words, locative *in-* seems to have a weaker morphological boundary than negative *in-* and *un-*. This impression is, however, to be tested empirically.

3.1.5 The prefix *dis-*

The prefix *dis-* is a non-native (or Latinate) prefix that takes mostly native bases. Occasionally non-native bases are found in *dis-*derivatives (e.g. in the word *disbelief*). The prefix takes mostly verbs and nouns as its base, sometimes adjectives. It is rarely category-changing and mostly takes words as its bases (see Jespersen 1965: 481; Marchand 1969: 158ff. Bauer et al. 2013: 355, 357). According to Bauer et al. (2013: 358), the few *dis-*derivatives with bound roots, such as *distort*, came to English as borrowings. Furthermore, the authors state that these borrowed words differ in their phonological structure from the other *dis-*words (e.g. in their stress pattern).

As negative *in-* and *un-*, the prefix *dis-* denotes negativity. With adjectives *dis-* is, similar to *un-* and *in-*, contrary or contradictory (e.g. *dishonest*). It is only in this category that *dis-* is productive (see Bauer et al. 2013: 358). With nominal bases *dis-* is privative (e.g. *disbelief*). With verbs it often is privative or reversative (e.g. *disconnect* or *disarm*) (see Bauer et al. 2013: 372, 375). Because of the

similarities in meaning between *un-*, negative *in-* and *dis-*, the discussion of affix rivalry can be extended to the prefix *dis-*. As with *un-* and *in-*, there is, however, no evidence for the blocking of one derivative by the existence of another. This is indicated by the co-existence of *un-* and *dis-*derivatives with the same base and identical meaning (see Bauer et al. 2013: 380). Examples are *unarm / disarm* and *uncharge / discharge*. Bauer et al. (2013: 380) note that semantic differences between *un-* and *dis-* are restricted to specific derivatives and are due to lexicalization. The prefix *dis-* is more often lexicalized than the prefix *un-* (e.g. *unbar* 'remove the bar' vs. *disbar* 'remove from the bar, i.e. to be disqualified from practicing law').

Different from *in-*, there is no systematic assimilation of the prefix *dis-*. However, as mentioned by Jespersen (1965: 480), in some *dis-*derivatives the prefix-final fricative has assimilated to its base-initial vowel by becoming voiced (e.g in the words *disease* and *dissolve*). Note that in these words assimilation goes together with resyllabification. Marchand (1969: 158) claims that this assimilation with *dis-* can only be found in monomorphemic words, such as in the word *disaster*. However, while most of the *dis-*words with a voiced fricative are indeed simplex, at least some can be categorized as complex using to the methodology described above (e.g. *disease* and *dissolve*). Note that in these assimilated words the meaning of the derivative cannot be deduced from the meaning of its parts, i.e. they are semantically opaque. Hence, in these words semantic opacity goes together with phonological opacity.

The stress pattern of *dis-*prefixed words is similar to the stress pattern of *in-*prefixed words. In derivatives with an unstressed base-initial syllable *dis-* is assumed to be stressed (e.g. ˌdisalˈlow, ˌdisenˈdow). In derivatives with bases starting in a stressed syllable, such as *disˈown*, the prefix is usually unstressed (see Jespersen 1965: 479f. Bauer et al. 2013: 183). There are also some *dis-*derivatives which bear primary stress on the prefix (e.g. ˈdisparate, ˈdiscount, ˈdissipate). Bauer et al. (2013: 183, 360) state, however, that primary stress with *dis-* is very rare and mostly restricted to semantically opaque derivatives with bound roots. On a more general note, the authors furthermore note that stress with *dis-* does not seem to follow a systematic pattern. In Wells (2008: 223) it is stated that *dis-* is "stressed when followed by an unstressed syllable, and often even when not". This statement also mirrors that there does not seem to be a systematic and accurate description of the stress behavior of *dis-*.

All in all, *dis-* is less productive and less transparent than *un-*. Some derivatives are semantically opaque and feature bound roots. There is no systematic phonological assimilation. Only in a few derivatives assimilation can be found.

These derivatives are, however, often regarded as simplex. As stated by Bauer et al. (2013: 440), in most derivatives *dis-* behaves like a prosodically independent word, i.e. there is no phonological integration of the prefix with its base. There are, however, a few cases in which *dis-* bears primary stress (e.g. *discount*), as well as a few derivatives with an assimilated fricative (e.g. *disease*). This indicates that for some *dis-*derivatives there is some phonological integration of the prefix with its base.

3.1.6 The suffix *-ly*

The fourth affix investigated in this study is the native suffix *-ly*. There are two different types of *-ly* in English, adjectival *-ly* as in *friendly* and adverbial *-ly* as in *coolly*. In this study, I will concentrate on adverbial *-ly*.

Adverbial *-ly* has been a quite frequent topic of discussion in the morphological literature (see for example Zwicky 1995; Plag 2003; Giegerich 2012; Bauer et al. 2013). Debates revolve around the question of whether the suffix is inflectional or derivational. The main argument for the derivational status of adverbial *-ly* is that the suffix is category changing, i.e. it takes adjectives as its base and turns them into adverbs. However, in this context the question of whether adverbs and adjectives should be seen as two distinct syntactic categories must be raised. If one does not adhere to the adjective-adverb distinction, the category-changing argument is no longer valid (cf. Plag 2003: 195; Giegerich 2012).

There are various arguments for the inflectional status of *-ly*. First, adverbial *-ly*, in contrast to the majority of derivational suffixes, cannot be followed by other derivational suffixes. In some cases it is even preceded by an inflectional suffix (e.g. in the word *interestingly*) (cf. Giegerich 2012). Second, *-ly* does not denote a clear lexical meaning (cf. Plag 2003: 195; Giegerich 2012; Bauer et al. 2013: 324). The suffix itself just expresses manner, means, duration, frequency and temporal relations. It follows that the meaning of *-ly*-derivatives is mostly determined by the meaning of their base and not by the meaning of the suffix (see Bauer et al. 2013: 326f.). Third, the suffix *-ly* is, similar to inflectional suffixes, very transparent and productive. Plag (2003: 196) mentions that only a few semantically opaque *-ly*-words exist. Examples are *hardly* and *shortly*. The suffix is very productive and can more or less freely attach to any adjective. The only restriction to its productivity, mentioned in Bauer et al. (2013: 334), is its reluctance to attach to bases already ending in *-ly*. While its high productivity and the lack of a clear semantic meaning might indeed suggest that *-ly* is inflectional, Bauer et al. (2013: 324) point out that there are also other English derivational

suffixes with similar features (e.g. *-al* and *-ment*) whose derivational status is not doubted.

To summarize the discussion about the derivational status of adverbial *-ly*, there are arguments for both assumptions, i.e. in some respects adverbial *-ly* behaves like an inflectional suffix and in some it behaves like a derivational suffix. As pointed out by Plag (2003: 196) the distinction between inflection and derivation might not be categorical, and *-ly* might be an example of an affix which lies in between.

Turning to phonology, some *-ly*-derivatives, such as *markedly* and *amazedly*, display base allomorphy. In those derivatives an epenthetic vowel is inserted between base and suffix to meet the prosodic requirement of *-ly*. As reported by Bauer et al. (2013: 172), *-ly* requires bases with more than one syllable to form a trochee before the suffix. If the base does not meet this requirement, vowel epenthesis, i.e. base allomorphy, is used change the prosodic structure of the base. The suffix itself does, however, not display any allomorphy.

Bauer et al. (2013: 163) report that some *-ly*-derivatives display resyllabification. They provide the word *frailly* as the only example. Note that the morphological geminate /ll/ in this example is assumed to be realized as a single syllable-initial consonant, i.e. the authors assume degemination. This demonstrates that resyllabification is closely connected to gemination, and that it should, if possible, be taken into account when investigating gemination. However, resyllabification is very difficult to investigate. Bauer et al. (2013: 168) attribute this difficulty to the variability found in resyllabification. Furthermore, syllable boundaries are very difficult to determine and empirical evidence is needed to make valid statements. Up to now this empirical evidence is lacking, and therefore, no statement about resyllabification with *-ly* can be made. With regard to stress, one can state that the suffix is never stressed and that it does not cause stress shifts in its base (cf. Bauer & Huddleston 2002: 1670; Bauer et al. 2013: 323).

To summarize, *-ly* is a very productive and semantically transparent suffix. Because of its regularity, formal properties and its rather weak semantic contribution, it is often argued to be inflectional. The suffix is, however, category-changing and therefore usually categorized as derivational. It is unstressed, sometimes causes allomorphy in its base and might resyllabify.

3.2 Comparison of the affixes

The analysis above has revealed that the affixes *un-*, negative *in-*, locative *in-*, *dis-* and *-ly* are characterized by different phonological, morphological and seman-

tic properties. These differences lead to differences in boundary strength, which in turn lead to different predictions for gemination. Because of this relation of boundary strength and predictions for gemination, it is necessary to compare the affixes with regard to their boundary strength.

The notion of boundary strength generally refers to the morphological boundary which separates the affix from its base. This strength can, however, be defined in different ways. One can look at it from a prosodic point of view by asking to which degree an affix is phonologically integrated in its base. Alternatively, one can look at it from a lexical point of view, taking factors such as semantic transparency and productivity into account. In the following comparison I will take different viewpoints, i.e. I will compare the affixes and their boundary strengths from a phonological/prosodic point of view and from a lexical point of view. In addition to the term boundary strength, I will use the terms DECOMPOSABILITY and SEGMENTABILITY. While I will use the term decomposability when talking about derivatives, I will use the term segmentability when talking about affixes. The stronger a boundary, the more decomposable is the derivative and the more segmentable is the affix.

Table 3.1 summarizes the characteristics of each affix as described in the previous sections. Looking at phonological and prosodic factors, it is striking that *-ly* is different from all other affixes in the set. It is never stressed, claimed to resyllabify and may lead to base assimilation. All other affixes in the set may be stressed and do not resyllabify. For those affixes which undergo assimilation, the assimilation always affects the affix, i.e. not the base. One major reason for the differences between *-ly* and the other affixes might of course be the fact that *-ly* is the only suffix in the set. With regard to prosodic boundary strength, one can assume that the possible resyllabification and the fact that *-ly* is always unstressed hint at a rather weak boundary between base and suffix.

Comparing the four prefixes with each other, the two *in*-prefixes show the most integration with their base. This is indicated by their obligatory assimilation. Even though there might be some derivatives with locative *in-* which do not undergo assimilation (some derivatives with native locative *in-*), the number of these derivatives is assumed to be very small. It is therefore assumed that locative *in-* assimilates to a similar degree as negative *in-*. The prefixes *dis-* and *un-* integrate less with their base. For *dis-* only a few derivatives assimilate. For *un-* assimilation is optional. A valid comparison of stress with the four prefixes is difficult to make. This is due to the general problems with determining prefixal stress. However, based on the descriptions above, it seems that while primary stress is possible on *in-* and *dis-*, it is not on *un-*. Prefixal primary stress can be

Table 3.1: The affixes under investigation: a comparison

	un-	*negative in-*	*locative in-*	*dis-*	*-ly*
Assimilation	non-obligatory affix assimilation	obligatory affix assimilation	mostly obligatory affix assimilation	no assimilation	base assimilation possibly
Resyllabification	no	no	no	no	
Stress	possible	possible	possible	possible	never stressed
Type of Base	free bases	mostly free bases	frequently bound bases	mostly free bases	free bases
Semantics	clearly negative	clearly negative	locative	clearly negative	no clear meaning
Semantic Transparency	transparent	mostly transparent	mostly opaque	mostly transparent	transparent
Productivity	very productive	productive	rarely productive	productive	very productive

regarded as part of a prefix's phonological integration with the base. Only in derivatives with a rather weak prosodic boundary the base's stress pattern can be changed in affixation, i.e. primary stress can be shifted from the base to the affix. In other words, prefixes with primary stress are less segmentable and feature a weaker boundary. Thus, with regard to stress, *un-*, which is never primarily stressed, is more segmentable than the other three prefixes.

Overall the following picture emerges. The strongest prosodic boundary can be found in *un*-derivatives, followed by *dis*-derivatives. Both *in*-prefixes have a rather weak prosodic boundary. Because of general differences between prefixes and suffixes a comparison of *-ly* with the other affixes is difficult. However, since *-ly* features a rather weak prosodic boundary, its prosodic segmentability can be assumed to be comparable to the one of *in-*, which also features a weak prosodic boundary.

Turning to lexical factors, the prefix *un-* is the most transparent and most segmentable affix. In other words, it has the strongest boundary. This is indicated by its high productivity, its clear semantics and the fact that it takes only words as its base. The suffix *-ly* is similar to *un-* in its productivity, transparency and the type of base it takes. A very important difference is, however, that *-ly* does not denote a clear lexical meaning. Therefore, the boundary between base and affix is assumed to be weaker for *-ly* than for *un-*. Negative *in-* and *dis-* are very similar with regard to their lexical attributes. Both of them are productive, have a preference for words as their base, have mostly transparent derivatives and denote a clear lexical meaning. They are assumed to be less segmentable than *un-*, which is even more transparent and productive. The comparison to *-ly* is difficult. On the one hand, *-ly* is more productive and more transparent than negative *in-* and *dis-*, on the other, *-ly*, in contrast to negative *in-* and *dis-*, does not denote a clear meaning. Whether the two prefixes or the suffix have a stronger boundary therefore depends on how one ranks the importance of a clear lexical meaning for determining boundary strength. Locative *in-* has the weakest boundary of the four prefixes. Its meaning is mostly opaque, it is rarely productive and very often takes bound roots. In contrast to *-ly*, locative *in-* does, however, denote a clear lexical meaning when semantically transparent.

Figure 3.1 summarizes the comparison of affixes by displaying segmentability hierarchies. In these segmentability hierarchies the affixes are sorted by their boundary strength. Since boundary strength can be defined in different ways, three hierarchies are displayed. The first one orders the affixes by prosodic boundary strength (Prosodic Hierarchy), the second and the third by lexical boundary strength (Semantic Hierarchy, Non-Semantic Hierarchy). The difference

between the Semantic Hierarchy and the Non-Semantic Hierarchy is that in the Semantic Hierarchy lexical meaning is regarded as a more important criterion for boundary strength than productivity, semantic transparency and type of base. The opposite is assumed in the Non-Semantic Hierarchy. This leads to different positions of the suffix *-ly*. While in the Semantic Hierarchy *-ly* is ranked as the least segmentable affix, in the Non-Semantic Hierarchy it is positioned between *dis-*, negative *in-* and *un-*.

Overall one can see similar patterns in all hierarchies. This shows that prosodic and lexical boundary strength cannot be regarded as two independent concepts but that they rather go hand in hand. For example, in all hierarchies *un-* is very segmentable and locative *in-* is rather less segmentable, i.e. features a rather weak boundary. However, there are also differences between the hierarchies. Especially the ranking of the suffix *-ly* highly depends on which criteria one applies to determine boundary strength.

	Segmentability hierarchy	Additional assumption
Prosodic Hierarchy	*un-* > *dis-* > {*in-*$_{\text{NEG}}$, *in-*$_{\text{LOC}}$, *-ly*}	
Semantic Hierarchy	*un-* > {*dis-*, *in-*$_{\text{NEG}}$}> *in-*$_{\text{LOC}}$ > *-ly*	lexical meaning over productivity, transparency and type of base
Non-Semantic Hierarchy	*un-* > *-ly* > {*dis-*, *in-*$_{\text{NEG}}$}> *in-*$_{\text{LOC}}$	productivity, transparency and type of base over lexical meaning

Figure 3.1: Segmentability hierarchies of affixes

There are two important things to note with regard to the hierarchies just proposed. First, they are based on theoretical assumptions found in the morphological literature on affixes. These assumptions are only partly supported by empirical evidence, and it is therefore necessary to empirically test whether they are borne out by the data. This will be done in this book. Second, the hierarchies proposed are based on categorical distinctions. In other words, they assume very

similar behavior across the derivatives of one affix. This means that predictions on gemination based on these hierarchies can only predict the behavior of all derivatives of an affix, not the behavior of individual words. Some of the theoretical approaches discussed in the next sections are categorical in nature. For these approaches I will return to the segmentability hierarchies in order to make predictions about gemination. For gradient approaches, i.e. approaches that do not assume the uniform behavior of the derivatives of a certain affix, the segmentability hierarchies are of less relevance.

3.3 Scope of gemination across affixes

Let us now have a look at the scope of gemination for each affix. An easy way to get an idea of how many types with morphological geminates exist for each affix is conducting a corpus study. Using the query tool Coquery (Kunter 2016), I searched the DVD version of the Corpus of Contemporary American English (COCA) (Davies 2008–2014) for all *un-*, *in-*, *dis-* and *-ly*-affixed words with an orthographic double consonant at the morphological boundary. With more than 520 million words COCA is one of the largest available corpora for English. It can therefore be assumed that the number of types found in the corpus is representative of the actual number of types.

I searched for morphological geminates by querying orthographic strings. For prefixes I searched for the orthographic string the prefix is made of followed by the last segment of the prefix, i.e. <unn> for *un-*, <inn> for *in-* and <diss> for *dis-*. Note that I also conducted a search for the allomorphs /ɪm/, /ɪr/ and /ɪl/ of *in-*, i.e. I also searched for the orthographic strings <imm>, <irr> and <ill>. For the suffix *-ly*, I searched for the sequence <lly>. I then checked the morphological status of the words found using the CELEX database (Baayen et al. 1995) and the English Lexicon Project (ELP) (Balota et al. 2007). All words which, according to at least one of the two databases, featured the affix in question were included in the count. Note that, due to the fact that the two databases do not feature all existing affixed types, and that not all existing types with a morphological geminate are attested in COCA, the count presented here is not exhaustive. Furthermore, morphological geminates which are not represented by the sequence of two identical orthographic segments (e.g. *solely, unknown*) are not taken into consideration. Nevertheless, the count conducted is useful to get a general impression of the scope of morphological gemination across the affixes.

Figure 3.2 shows a bar plot displaying the number of types of morphological geminates for each affix. Each bar represents one affix. The number of types is

given next to each bar. The bar for the prefix *in-* represents the number of types with all four allomorphs of *in-*, i.e. /ɪn/, /ɪm/, /ɪr/ and /ɪl/.
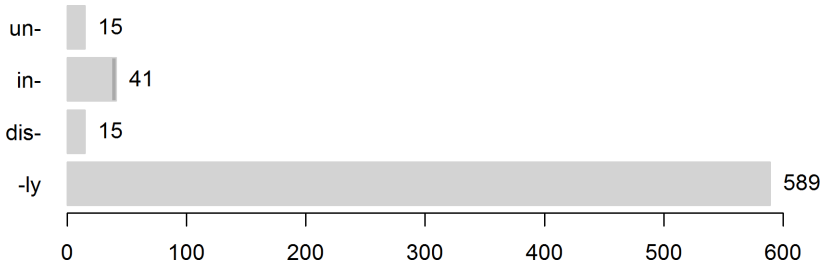
Figure 3.2: Number of types with morphological geminates for each affix

A comparison of the affixes reveals that while there are quite a lot of types with morphological geminates for the suffix *-ly*, the number of types is much smaller for all of the prefixes. In other words, there are not many *un-*, *in-* and *dis-*prefixed words with a morphological geminate. Especially the number of morphological geminates with *un-* and *dis-* is very small and might raise methodological problems with regard to the investigation of some of the factors possibly influencing gemination. Some variables, such as frequencies, might not show enough variation to reliably be investigated. I will return to these issues in the pertinent sections of this book.

For the prefix *in-*, the number of attested types with morphological geminates is much higher than for the prefixes *un-* and *dis-*. However, out of the 105 attested types only one features the allomorph /ɪn/ (e.g. *innumerable*). 33 types start with the allomorph /ɪm/ (e.g. *immortal*), 51 with /ɪr/ (e.g. *irregular*), and 20 with /ɪl/ (e.g. *illogical*). This distribution has important consequences for the investigation of the prefix. Since the number of types in /ɪn/ is so limited, my investigation of gemination with *in-* will mainly focus on the allomorph /ɪm/, for which more types exist. This is, as discussed in §2.4.2, in line with previous empirical work on gemination with *in-*. The investigation of /ɪr/ and /ɪl/ is not reasonable as derivatives with these allomorphs always feature a morphological geminate, i.e. there are no derivatives with /ɪr/ and /ɪl/ which feature a singleton consonant at their morphological boundary. This makes a comparison of the duration of a phonological double with a phonological singleton impossible.

4 Morphological gemination: implications for theory

In this chapter, I will show how the gemination pattern of English affixes can shed light on various theories of the morpho-phonological and morpho-phonetic interface. I will discuss three different fields of theory which directly or indirectly predict the gemination pattern of English: Formal linguistic theories, psycholinguistic approaches of morphological processing and theories of speech production. After giving a general overview, I will discuss each field separately, explain its main branches and deduce the predictions made for gemination with the five affixes investigated in this study. At the end of the chapter, I will summarize the predictions of all theories discussed.

4.1 Overview

One can distinguish two major branches of morpho-phonological (or morpho-phonetic) theories. The first set of theories is formal, categorical and often generative in nature. Generally, in this type of theory it is assumed that phonological entities are abstract. Furthermore, these theories do not allow for a direct morpho-phonetic interface. Phonetic detail is thus believed to not mirror morphological structure directly. Depending on the theory, phonological rules and prosodic structure can, however, indirectly mirror the morphological structure of complex words and lead to phonetic effects. In these formal theories, morphological gemination is regarded as a categorical morpho-phonological and rule-based process, i.e. not as a gradient, word-specific process of the direct morpho-phonetic interface. Degemination is categorical and occurs when, due to some phonological rule or process, one segment of the double consonant is deleted. This deletion is then mirrored in phonetics. Depending on the theory, the predictions of which structures undergo degemination, and which structures display gemination, deviate. I will look at three prominent formal linguistic approaches in this book, i.e. Lexical Phonology, newer stratal approaches such as Stratal

Optimality Theory, and Prosodic Phonology, and discuss their predictions for gemination with the five affixes looked at in this study.

The second branch of theories is psycholinguistic in nature and incorporates gradient, probabilistic factors. Experience-based effects, such as probabilities and frequencies of occurrence, play a major role in this type of theory. In contrast to formal approaches, theories of this kind do not assume gemination to be categorical and rule-based, but rather to be gradient and word-specific. In general these psycholinguistic approaches assume a more direct relation between morphological structure and phonetics than formal theories by, for example, assuming that fine phonetic detail may be stored in the mental lexicon. In this study, I will concentrate on two prominent factors in psycholinguistic theory: decomposability and morphological informativeness. These two factors are believed to influence the phonetic realization of complex words, and should thus also play a prominent role in morphological gemination.

The binary classification of theories into formal, categorical approaches and gradient, probabilistic approaches must not be regarded as absolute. Formal linguistic theories may incorporate gradient, probabilistic factors, and in psycholinguistic approaches categorical concepts may be found. However, even though the categorization is not absolute, it helps to get a systematic overview of approaches which propose explanations for morpho-phonological and morpho-phonetic variation. The categorization simplifies the comparison of similar approaches, and helps with the identification of changes and developments of theories.

In addition to the two types of theories outlined above, I will also discuss speech production theories. There are various ways of how the morpho-phonological (or the morpho-phonetic) interface is modeled in speech production. I will present two different types of models. In the first type, a strict feed-forward structure with no explicit morpho-phonetic interface is assumed. The second type of model is based on exemplars, i.e. concrete mental entities which are formed from experience. While the former fits in better with the categorical approaches outlined above, the latter can be regarded to be more in the spirit of gradient approaches. However, I will desist from drawing explicit relations between speech production models and the different formal linguistic and psycholinguistic approaches discussed. The reason is that these relations are never explicitly mentioned in the literature, and that, due to the underspecification of morpho-phonological aspects in speech production models, a lot of information is not available that would be needed to relate the different theories with each other. I will therefore review the two speech production models on their own. I will discuss them with respect to their understanding of morpho-phonological

processes (such as, for example, morphological gemination), and will especially focus on the already mentioned underspecified aspects of the processing of complex words. Because of these underspecifications no explicit predictions for gemination in English affixation will be drawn from these models. However, I will discuss more general assumptions the models make with regard to the phonetic realization of complex words. Some of these assumption can and will be tested in this study.

4.2 Formal linguistic theories

In this section, I will discuss the following formal linguistic approaches: Lexical Phonology, newer stratal approaches such as Stratal Optimality Theory, and Prosodic Phonology. Within the framework of Prosodic Phonology, I will concentrate on the PROSODIC WORD. In all three approaches morpho-phonological processes are assumed to be categorical in nature. Furthermore, in all of them affixes are believed to play a crucial role in morpho-phonological processes. The three approaches assume gemination to be a categorical process, i.e. either the two consonants of a morphological geminate remain and we find gemination, or one consonant is deleted and we find degemination. In all of the approaches boundary strength is one of the most important determiners for phonological behavior, i.e. also for gemination. One can thus see that there are quite a few similarities in the approaches. However, there are also important differences between them. One of these differences is the type of boundary assumed to be decisive for the phonological behavior of a derivative. While Lexical Phonology, for example, assumes affix-specific lexical differences in boundary strength, Prosodic Phonology defines boundary strength mainly in phonological terms. Another difference between the approaches is the assumed role of gradiency in morpho-phonological processes. Crucially, these differences between approaches lead to different predictions for gemination. In the following, I will discuss each approach and present the predictions it makes for gemination with the five affixes investigated in this study.

4.2.1 Lexical Phonology

Lexical Phonology is one of the first theories which discussed the morpho-phonological interface and which makes clear predictions about gemination in English. According to Lexical Phonology (cf. Kiparsky 1982; Mohanan 1986), all morphological processes are carried out by modules in the lexicon. These modules are

called levels or strata. A complex word is formed by retrieving entries from the lexicon which then undergo the morpho-phonological processes of the different strata. The strata are linked to specific affixes and are connected to phonological rules, which are carried out at the level they belong to. At each level one or more affixes can be added to a word. After affixation, the phonological rules adjust the form of the derivative so that the phonotactics of the language the word is formed in are met. During the word formation process an item can pass each level several times. A word does not leave a stratum until all operations the stratum holds for the word are carried out, i.e. all derivational or inflectional processes of a level are completed. This concept is called CYCLICITY. After a word has passed all strata, it leaves the lexicon and undergoes post-lexical rules. These rules adjust the word's sub-phonemic and word-external features. At the post-lexical stage the complex word does not carry any morphological information. This is due to the principle of BRACKET ERASURE, which erases all morphological boundaries after each lexical level. Hence, according to Lexical Phonology the phonetics of a complex word does not show traces of its morphological structure.

Figure 4.1 depicts the word formation process for the complex word *imperfect-ness*, which consists of the level 1 prefix *in-*, the root *perfect* and the level 2 suffix *-ness*.[1] First, the root *perfect* is retrieved from the lexicon and passed to the first level. After the level 1 affix *in-* is added to the root, level 2 morphology passes the concatenated form to level 2 phonology, which applies the pertinent lexical rules. In this case, the adjacency of the two phonemes /n/ and /p/ triggers the application of an assimilation rule which changes /n/ to /m/. After this change in form, the word is passed back to level 2 morphology to check whether further affixes need to be added. Since this is not the case, the word leaves level 2 and is sent to the second level. Before entering level 2, all morphological boundaries of the derivative are erased (due to bracket erasure). When *imperfect* undergoes level 2 processes it is thus treated the same way as a simplex word. In level 2, level 2 morphology adds the suffix *-ness* to the derivative. The attachment of *-ness* does not change the phonological form of the word, i.e. no phonological rules are applied. The derivative goes through level 2 morphology and phonology again to check whether additional morphemes need to be added. It then leaves the lexicon, i.e. it is passed to the post-lexical stage.

---

[1]Note that the figure only displays two levels, i.e. level 2 and level 2. These are the two levels relevant for affixation in English, i.e. the two levels relevant for this study. The number of assumed levels differs between different variants of Lexical Phonology (see Giegerich 1999 for an overview), the level 1- level 2-distinction, as displayed in Figure 4.1, is, however, agreed on by all variants.
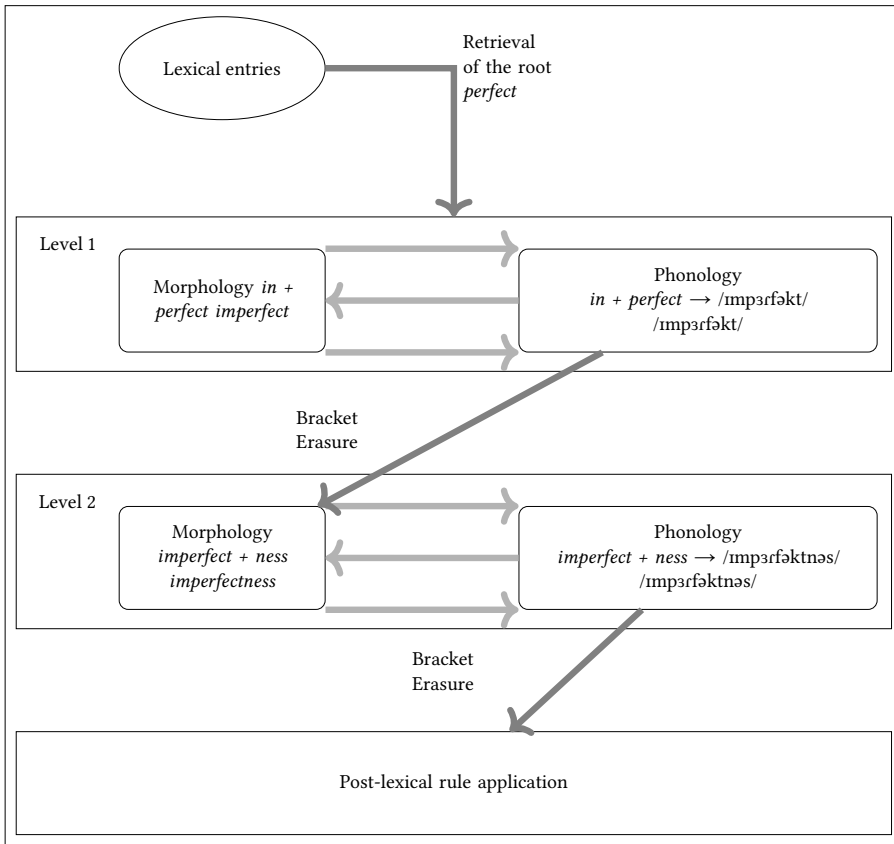
Figure 4.1: Word formation process of *imperfectness* in Lexical Phonology

As suggested by this example, level 1 affixes differ from level 2 affixes with regard to their boundary strength. The notion of boundary strength, as applied here, goes back to Chomsky & Halle (1968). They assumed two different types of morphological boundaries: strong ones and weak ones. While level 1 affixes trigger weak morphological boundaries, which go along with a high degree of phonological integration, level 2 affixes form strong boundaries and integrate less. The example *imperfectness* illustrates the difference. While the adding of the level 1 affix *in-* leads to assimilation, neither the phonological form of the base, nor the one of the suffix changes when the level 2 affix *-ness* is added. One should, however, note that, even though the notion of boundary strength corresponds well with the theory of Lexical Phonology, Kiparsky (1985: 239) explicitly states

that by introducing the different strata the notion of boundary strength becomes unnecessary in Lexical Phonology.

Transferring the concept of lexical strata and its implications for the phonological relation between affix and base to gemination, the following picture emerges: a morphological geminate will degeminate when part of a level 1 affix, and it will geminate when part of a level 2 affix (see Mohanan 1986: 18). To find out what is predicted for the affixes under investigation, we need to determine which level each affix belongs to. Based on phonological and lexical properties, as well as on affix origin, Lexical Phonology assumes that the affixes *dis-* and *in-* belong to level 2, and that the affixes *un-* and *-ly* belong to level 2. Note that no differentiation between locative and negative *in-* is made. Both belong to level 2. Comparing the distinction between level 1 and level 2 affixes with the segmentability hierarchies proposed in the previous chapter, one can see that the distinction resembles the Non-Semantic Segmentability Hierarchy (*un-* > *-ly* > {*dis-*, *in-*$_{\text{NEG}}$} > *in-*$_{\text{LOC}}$). This hierarchy is based on lexical factors and ranks productivity and transparency above lexical meaning.

Figure 4.2 depicts Lexical Phonology's prediction for gemination (and degemination) with the five affixes. Due to a morphological process a morphological geminate, i.e. a double consonant, emerges. For the level 1 affixes negative *in-*, locative *in-* and *dis-* the theory predicts the deletion of one of the two adjacent consonants by some kind of phonological process. This deletion is expected to lead to a short duration of the morphological geminate, i.e. to degemination. For the affixes *un-* and *-ly* no deletion is expected, i.e. no phonological process is applied. The double consonant remains and the morphological geminate is therefore realized with a long duration. The affixes *un-* and *-ly* geminate. To summarize, according to Lexical Phonology, the level 2 affixes *un-* and *-ly* are expected to geminate, and the level 1 affixes negative *in-*, locative *in-* and *dis-* are expected to degeminate.

As described above, Lexical Phonology assumes gemination, i.e. the duration of the double consonant, to be mostly influenced by the type of affix involved. In other words, the affix involved affects consonant duration the most. However, there are additional factors which are assumed to also influence duration, namely post-lexical factors (e.g. speech rate, the preceding segment). Importantly, these factors are expected to be found on top of the expected effect of the affix. Lexical factors which mirror an individual word's morphological structure are not expected to play a role. All traces of morphological structure are erased by bracket erasure. The predictions of Lexical Phonology are summarized below Figure 4.2.
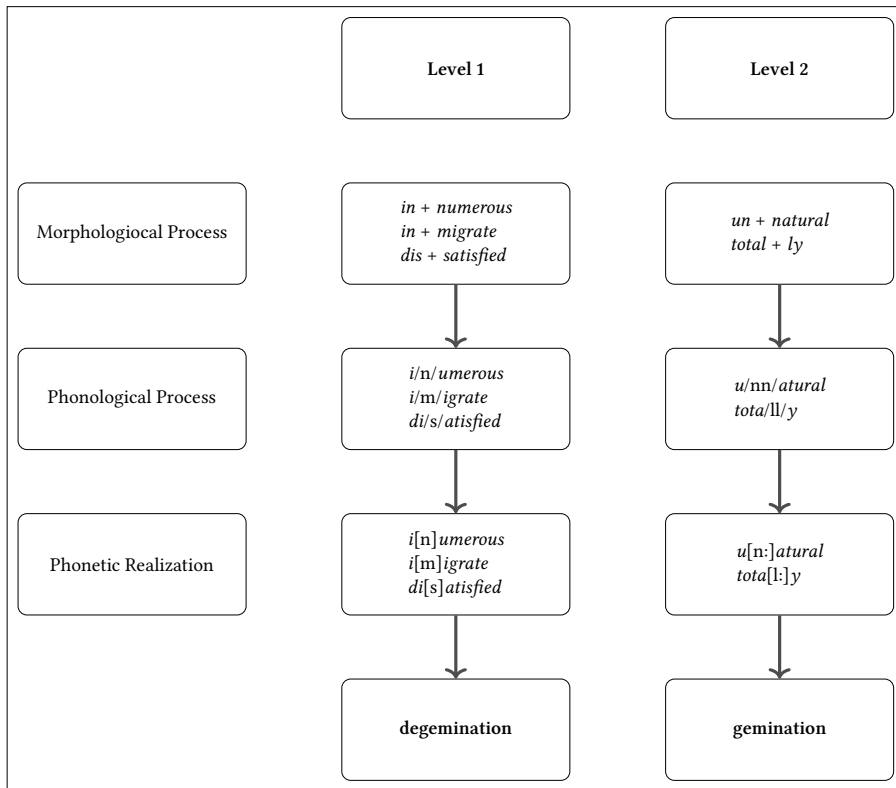
Figure 4.2: Lexical Phonology: predictions for gemination with *un-*, negative *in-*, locative *in-*, *dis-* and-*ly*

Lexical Phonology: Predictions

1. *un-* geminates
   The nasal in *un*-prefixed words with a phonological single consonant will be shorter than the nasal in *un*-prefixed words with a phonological double consonant.

2. *in-* degeminates
   The nasal in *in*-prefixed words with a phonological single consonant will be as long as the nasal in *in*-prefixed words with a phonological double consonant.

3. *dis-* degeminates
   The fricative in *dis*-prefixed words with a phonological single consonant

will be as long as the fricative in *dis*-prefixed words with a phonological double consonant.

4. *-ly* geminates
   The lateral in *-ly*-suffixed words with a phonological single consonant will be shorter than the lateral in *-ly*-suffixed words with a phonological double consonant.

5. Post-lexical factors affect consonant duration.

6. Factors which mirror the morphological structure of individual words do not affect consonant duration.

4.2.2 Newer stratal approaches

During the last decades various problems with Lexical Phonology have been revealed. These problems are closely related to the strict feed-forward structure of the model, as well as the strict division of labor between the strata (see also Giegerich 1999: chapter 2 and Plag 2003: chapter 7 for a thorough discussion). The criticism of Lexical Phonology has led to modifications of the theory and the development of newer stratal approaches, which are argued to solve the problems found with the original approach. Two of these approaches will be discussed in this section: the base-driven approach proposed by Giegerich (1999) and Stratal Optimality Theory. To understand the changes made in these newer approaches, it is necessary to take a closer look at the problems with the original approach.

One of the most prominent problems with Lexical Phonology is affix ordering. According to Lexical Phonology, level 1 affixes are attached to a base before level 2 affixes, and level 2 affixes are attached before inflectional affixes.[2] Therefore, level 2 affixes are not expected to be found within level 1 affixes, and inflectional affixes are not expected to be found within level 2 or level 2 affixes. However, in some derivatives these unexpected structures are found. For example, in the derivative *interestingly* the inflectional suffix *-ing* precedes the level 2 suffix *-ly*,[3] and in the derivative *ungrammaticality* the level 1 affix *-ity* is attached after the level 2 affix *un-*. Examples like these call into question whether the strict feed-forward structure proposed by Lexical Phonology is valid (see also Plag (1999: chapter 4) for discussion).

---

[2]According to most varieties of Lexical Phonology (e.g. Kiparsky 1982; Mohanan 1986) inflectional affixes belong to an additional stratum which follows derivation.

[3]Note that this structure does not pose a problem to Lexical Phonology if *-ly* is regarded as an inflectional suffix. However, Lexical Phonology adapts the view that *-ly* is derivational.

Another problem with Lexical Phonology is the variation found within affixes. According to the theory, all affixes of one stratum, and thus all derivatives of one affix, should display the same (or at least very similar) behavior with regard to morpho-phonological processes. However, recent literature has pointed out that there is variation within the strata and even within the derivatives of one affix (cf., for example, Raffelsiefen 1999; Bauer et al. 2013; Plag 2014; Bermúdez-Otero 2017). One prominent example of this variation is stress shift. The suffix -*able*, for instance, is expected to preserve stress. There are, however derivatives with -*able* in which stress shift is possible (e.g. *ˈanalyze – ˈanalyzable ~anaˈlyzable*). With some derivatives stress shift even seems to be consistent (e.g. *ˈcategorize – categoˈrizable*) (cf. Plag 2014: 213f.). Empirical studies on stress confirm variation within affixes, i.e. different derivatives of the same affix deviate in their stress pattern (see for example Collie 2008 for -*ion* and -*ity*, Sanz 2017 for -*ory*). The categorization of affixes as either level 1 or level 2 is insufficient to explain the variation found.

A related, and maybe more essential, problem is the categorization of an affix as either level 1 or level 2. As summarized by Raffelsiefen (1999: 134), level 1 affixes feature bound roots, are unproductive, determine stress, trigger assimilation and yield idiosyncratic meaning. Level 2 affixes attach to words, are productive, are stress-neutral, block assimilation and yield compositional meaning. However, as can be seen in Table 3.1 in §3.2, affixes often simultaneously feature level 1 and level 2 properties. Negative *in*-, for example, assimilates, i.e. features a level 2 property, and is productive and semantically transparent, i.e. simultaneously features level 2 properties. This shows that the assignment of an affix to one or the other level is often not clear-cut.

The three problems just discussed, i.e. affix ordering, intra-affix variation and the categorization of affixes, demonstrate that the strict division between level 1 and level 2 affixes, as well as the strict sequential order of level 2 and level 2 processes, are insufficient to explain the variation found in English derivatives. While this insufficiency has led some linguists to abandon the idea of lexical strata in general (cf., for example, Johnson 1997; Bybee & Hopper 2001; Pierre-humbert 2001; Hay 2001), others have modified the stratal approach. Giegerich (1999), for example, proposes a base-driven stratal approach. For English, he assumes two lexical strata: a root stratum and a word stratum. In the first stratum, i.e. the root stratum, all structure-changing morpho-phonological operations are carried out. In the second stratum, i.e. the word stratum, all structure-building processes take place. The important difference between Giegerich's approach and Lexical Phonology is that in Giegerich's approach, it is not the affix which is de-

cisive for the stratum a derivative is formed in. Instead, the distinction between stratum 1 and stratum 2 is based on a derivative's base. If the base is a word, the derivative is formed in stratum 2. If the base is a root, the derivative is formed in stratum 1. In light of categorical approaches of morpho-phonology, degemination can be classified as a structure-changing process, i.e. a stratum 1 process. Thus, only derivatives featuring a root, i.e. stratum 1 derivatives, are expected to display degemination. Derivatives with words as their base are expected to geminate.

There are two major problems with Giegerich's approach and the predictions it makes for gemination in English affixation. The first problem is that prefixation is not discussed, and that the approach is thus restricted to suffixation. One could assume that the predictions for suffixes can be extended to prefixes, i.e. one could predict prefixed words with words as bases to geminate and prefixed words with roots as bases to degeminate. It is, however, unclear whether this extension would comply with the approach. The second problem concerns the categorization of bases as roots or words. Giegerich does not give clear criteria for this categorization. This makes the testing of predictions for gemination which are based on the distinction of word vs. root almost impossible. Furthermore, Giegerich states speaker-dependent differences, i.e. the same base might be a root for one speaker and a word for another (cf. Giegerich 1999: chapter 3.2.1.). Due to these assumed differences among speakers, post-hoc explanations are available for all cases. If a word geminates, the speaker can be argued to have stored the base as a word. If a word degeminates, the opposite can be argued, i.e. the speaker has stored the base as a root. Because of these post-hoc explanations, the base-driven approach suggested by Giegerich might not be falsifiable. It is therefore not tested in this study.

Stratal Optimality Theory (Stratal OT) is another, newer development of stratal approaches. It combines the modular feed-forward structure of Lexical Phonology with OT mechanisms (cf. Bermúdez-Otero 2012; 2013; Kiparsky 2015; Bermúdez-Otero 2017). As in classical Lexical Phonology, the grammar is organized into three levels: two lexical levels and one post-lexical level. The first lexical level is the stem-level, the second is the word-level. The categorization of an affix as level 1 or 2 largely depends on the type of base it attaches to. If an affix attaches to a stem, it belongs to level 2, if it attaches to a word, it belongs to level 2 (cf. Kiparsky 2015: 7; Bermúdez-Otero 2017: 9f.). Thus, as in Giegerich's approach, the base of a derivative plays a crucial role in Stratal OT. Also similar to Giegerich's approach, the categorization of a base as a stem or a word is not always clear. However, it can be assumed that all bound roots are classified as stems. Free-

standing words might be classified as words. Crucially, in contrast to Giegerich's approach, speaker-dependent variation is not discussed in Stratal OT, i.e. post-hoc classifications are not an issue.

Stratal OT assumes each stratum to have its own constraint system. This system is responsible for phonological operations, i.e. also for degemination. As noted by Kiparsky (2015: 5), the constraint system at each level consists solely of input/output and markedness constraints. No additional constraints are assumed. The order of constraints may deviate between the levels. For morphological geminates one can assume a different ranking of constraints at each of the two lexical levels. This difference leads to degemination on the stem-level and gemination on the word-level. It is thus assumed that level 1 affixes degeminate, while level 2 affixes geminate. This prediction is very similar to the prediction made by Lexical Phonology. There is, however, an important difference between the predictions of the two approaches. This difference is rooted in the assumption of dual-level affixes.

Different from classical stratal approaches, Stratal OT allows for dual-level affixes. According to Stratal OT, these affixes behave like level 1 affixes when attached to a stem and like level 2 affixes when attached to a word (cf. Bermúdez-Otero 2017: 15, 33). Bermúdez-Otero (2017: 33) notes the prefix *in-* to be one of these affixes. He states that while in derivatives like *importune in-* behaves like a stem-level prefix, in derivatives like *impolite* it behaves like a word-level prefix. The comparison of the five affixes investigated in this study revealed that the prefix *dis-* behaves similarly to the prefix *in-* (cf. §3.2). Like *in-*, *dis-* also features level 2 and level 2 properties, e.g. it attaches to words and to bound roots. It can therefore be claimed that *dis-* also is a dual-level affix.

Since the behavior of dual-level affixes depends on the type of base they attach to, variation in the gemination of *in-* and *dis-* is expected. The prefixes are expected to geminate when attached to a word. They are expected to degeminate when attached to a stem. As discussed above, the distinction between stem and word is, however, not always clear. In turn, the predictions are not always clear. However, for some words clear predictions can be formed. Since all bound roots are classified as stems, *in-* and *dis-* are expected to degeminate in all derivatives with a bound root. Only in derivatives with a word as a base *in-* and *dis-* can geminate. The predictions for *un-* and *-ly* are very straightforward. According to Stratal OT, the two affixes belong to the second stratum. The prediction for *un-* and *-ly* therefore remains the same as in Lexical Phonology. Both affixes are expected to geminate.

Apart from dual-level affixes, there is another important difference between

Lexical Phonology and Stratal OT. This difference is related to the lexical storage of complex words. While according to Lexical Phonology all complex words are computed online, Stratal OT assumes stem-level derivatives to be stored as a whole (cf. Bermúdez-Otero 2012: chapter 3). Therefore, according to Stratal OT, frequency effects on level 2 derivatives are expected. Words with a high frequency are expected to be reduced phonetically, i.e. should also display shorter durations of the affixational consonant. This reduction process is, however, not expected for level 2 derivatives which are only stored analytically (cf. Bermúdez-Otero 2012: chapter 3.3).

One can summarize that while some of the basic assumptions of Lexical Phonology remained in newer approaches (e.g. modular feed-forward structure, lexical vs. post-lexical level), there are also some important differences (e.g. dual-level affixes, whole-word storage). These differences lead to different predictions for gemination with the five affixes of this study. While the prediction for the level 2 affixes *un-* and *-ly* remains the same, i.e. they are expected to geminate, variability is expected for the other two affixes. Another important difference to Lexical Phonology is that Stratal OT includes psycholinguistic, gradient factors, such as frequencies, by assuming whole-word storage of stem-level derivatives. Below the predictions of Stratal OT are summarized.

Stratal Optimality Theory: Predictions

1. *un-* geminates
   The nasal in *un*-prefixed words with a phonological single consonant will be shorter than the nasal in *un*-prefixed words with a phonological double consonant.

2. *in-* degeminates in derivatives with a bound root
   The nasal in *in*-prefixed words with a phonological single consonant will be as long as the nasal in *in*-prefixed words with a phonological double consonant if the derivative has a bound root as its base.

3. *in-* geminates in derivatives with a word as a base
   The nasal in *in*-prefixed words with a phonological single consonant is shorter than the nasal in *in*-prefixed words with a phonological double consonant if the derivative has a word as its base.

4. *dis-* degeminates in derivatives with a bound root
   The fricative in *dis*-prefixed words with a phonological single consonant

Draft of May 27, 2019, 18:04

will be as long as the fricative in *dis*-prefixed words with a phonological double consonant if the derivative has a bound root as its base.

5. *dis-* geminates in derivatives with a word as a base
   The fricative in *dis*-prefixed words with a phonological single consonant is shorter than the fricative in *dis*-prefixed words with a phonological double consonant if the derivative has a word as its base.

6. *-ly* geminates
   The lateral in *-ly*-suffixed words with a phonological single consonant will be shorter than the lateral in *-ly*-suffixed words with a phonological double consonant.

7. Post-lexical factors affect consonant duration.

8. Word frequency influences the duration of level 1 derivatives, i.e. *in-* and *dis*-prefixed words with a bound root.

4.2.3 The prosodic word

Empirical work has shown that the morphological structure of a derivative may be directly mirrored in its acoustic realization. Sproat & Fujimura (1993), for example, found differences in the phonetic realization of /l/ depending on the strength of the morphological boundary the sound occurred at. Stronger morphological boundaries were found to feature a darker and longer /l/ than weaker morphological boundaries. Similarly, Lee-Kim et al. (2013) found that the phonetic realization of /l/ depends on the morphological structure of the derivative. For Dutch homophones, Schuppler et al. (2012) found more reduction with complex words than with simplex words, i.e. word-final /t/ was more often deleted in simplex than in complex words.

With regard to duration, Cho (2001), Sugahara & Turk (2009), Hanique et al. (2011), Smith et al. (2012) and Plag et al. (2017) found systematic differences between monomorphemic and morphemic words with similar phonological structures. For Korean, Cho (2001) found articulatory evidence on the variability of intergestural timing in monomorphemic and complex words. In an EPG study, the timing of the gestures for [ti] and [ni] shows more variation when the sequence is heteromorphemic (i.e. across a morpheme-boundary) than when it is tautomorphemic (i.e. without straddling a boundary). For English, Sugahara & Turk (2009) found phonetic differences between the final segments of a monomorphemic stem as against the final segments of the same stem if followed by a

suffix. Stems followed by level 2 suffixes had slightly longer rhymes than their monomorphemic counterparts. Smith et al. (2012) discovered systematic phonetic differences in the realization of the first three segments between prefixed words and what they call *pseudo-prefixed words* (such as *mis-time* versus *mistake*, respectively). Similarly Plag et al. (2017) found a difference between morphemic and non-morphemic /s/ in English. They found non-morphemic /s/ to be longer than morphemic /s/.

The studies above raise the question of how to model the relation between morphology and phonetics. Even though the studies yield somehow contradicting results, i.e. in some of the studies stronger morphological boundaries lead to more reduction (e.g. Sugahara & Turk 2009; Smith et al. 2012) and in some the opposite is the case (e.g. Schuppler et al. 2012; Plag et al. 2017), it seems certain that there is a stable effect of morphology on the acoustic realization of a word. One might thus suggest a morpho-phonetic interface which allows phonetics to directly access morphological information. This idea is picked up mainly by psycholinguistic approaches of morphological processing, which completely abandon the strictly feed-forward structure of stratal theories (cf. §4.3). In contrast, most formal theories hold on to the modular structure. These theories do not feature a direct interface between morphology and phonetics. Instead, they explain the effects of morphology on phonetics by referring to the prosodic structure of complex words (cf., for example, Booij 1983; Sproat 1993; Nespor & Vogel 2007; Sugahara & Turk 2009; Bergmann 2014). The prosodic structure of a derivative is closely connected to its morphological structure and is believed to directly influence the phonetic realization of a word.

A very important concept in prosodic approaches, especially with regard to duration, is prosodic boundary strength. The common assumption is that the stronger the prosodic boundary, the less reduction is found. The strength of a boundary in prosodic terms highly depends on the prosodic domain it is adjacent to. Figure 4.3 shows the prosodic hierarchy, which depicts the different prosodic domains.[4] The higher the domain in the hierarchy, the stronger the boundary and the less reduction is expected. Segments followed by an intonational phrase boundary are, for example, expected to be less reduced than segments followed by a phonological word boundary.[5] The effects of prosodic boundary strength are assumed to be additive, i.e. the more boundaries are present, the less reduction

---

[4]Note that the constituents of the hierarchy differ slightly depending on author and approach. The hierarchy displayed here is taken from Hall (1999: 9).

[5]The phonological word is also called *prosodic word* or *p-word*. In this book, I will use the terms interchangeably.

is expected.

<div align="center">

phonological utterance (U)

|

intonational phrase (IP)

|

phonological phrase (ϕ)

|

phonological word (ω)

|

foot (F)

|

syllable (σ)

</div>
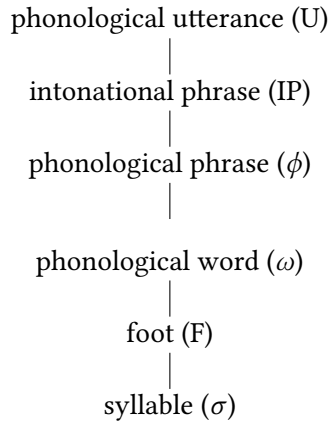
Figure 4.3: Prosodic Hierarchy

Morphological geminates in English affixed words can occur at phonological word boundaries, foot boundaries and syllable boundaries. Since "[t]he phonological word (ω) represents the interaction between the phonological and the morphological components of the grammar" (Nespor & Vogel 2007: 109), the phonological word boundary is the most important for this study. In other words, gemination might be determined by phonological word boundaries, which mirror the morphological structure of a derivative and influence the acoustic realization of the complex word. The phonological word can be regarded as a mediator between the morphology and the phonetics of a complex word.

A grammatical word can consist of one or more prosodic words (cf., for example, Booij 1983: 29; Booij 1985: 267; Hall 1999: 2). Importantly, only complex words can consist of several prosodic words. This is due to the fact that prosodic word boundaries must align with morphosyntactic boundaries (see for example Hall 1999: 2). Crucially, not every morphological boundary corresponds to a p-word boundary. In other words, while some affixes form separate prosodic words, and thus occur at a p-word boundary, others form a prosodic word together with their base, i.e. are not adjacent to a p-word boundary. Examples of complex words with differing prosodic word structures are given below. The affixes in the words on the left constitute independent prosodic words. The affixes in the words on the right do not form independent p-words. They form one prosodic word together with their base.

(1) $(un)_\omega(natural)_\omega$, $(un)_\omega(told)_\omega$ (2) $(really)_\omega$, $(inject)_\omega$

As stated by Hall (1999: 3), prosodic words can form the domain of phonological and prosodic rules. It can thus be assumed that prosodic words also form the domain of gemination. While derivatives in which the affix forms a prosodic word on its own (e.g. $(un)_\omega$ $(natural)_\omega$) geminate, affixes which form a prosodic word together with their base (e.g. $(really)_\omega$) degeminate (cf. Giegerich 2012: 3543; Bergmann 2014). In other words, reduction, i.e. degemination, is only found when the morphological geminate does not occur across a prosodic word boundary. The crucial question now is how to determine whether a morphological geminate occurs across such a boundary, i.e. whether an affix constitutes a prosodic word on its own.

It is generally assumed that prosodic word structure corresponds to morphological boundary strength. Stronger morphological boundaries form phonological word boundaries. However, the specific criteria for determining prosodic words status are still debated in the literature (cf. Raffelsiefen 1999; Hall 1999 for an overview).[6] In earlier approaches, prosodic word status was merely a mirror image of the level 1-level 2 distinction made in stratal theory. While level-1 affixes were assumed to be integrated in the p-word of their derivative, level 2 affixes were assumed to form a p-word on their own (see, for example, Aronoff & Sridhar 1983; Booij 1983; Szpyra 1989). According to this view, the predictions for gemination made by Prosodic Phonology would be identical to the ones made by Lexical Phonology. Level 2 affixes form an independent p-word and thus geminate, level 1 affixes do not form an independent p-word and thus degeminate.

More recent approaches to the prosodic word deviate from the stratal categorization. The prosodic word status of English suffixes, for example, is debated fairly often in the literature. The general assumption is that suffixes do not form prosodic words on their own (see, for example, Wennerstrom 1993: 311; Raffelsiefen 1999: 184; Hall 2001: 401; Sugahara & Turk 2009). There are, however, different views with regard to the question of whether suffixes follow the prosodic word formed by their base (e.g. $(run)_\omega$-*ing*), or whether they are integrated in the prosodic word of their base (e.g. $(running)_\omega$). While Sugahara & Turk (2009), for example, suggest that all level 2 suffixes follow the p-word formed by their base, Raffelsiefen (1999) and Hall (2001) propose that this is only true for some of those suffixes. They suggest that only consonant-initial suffixes, such as *-ness* and *-ly*, follow the prosodic word which is formed by their base. Vowel-initial suffixes,

---

[6]Note that the criteria also deviate between languages. The overview given here is limited to the prosodic word in English.

such as *-ing* or *-er*, are integrated in the prosodic word of their base. Examples of the p-word structures according to the two different approaches are given below. As can be seen, the suffix *-ly*, which is under investigation in this study, is preceded by a p-word boundary in both approaches. It is never integrated in the p-word of its base.

(3)

derivative	Raffelsiefen	Sugahara & Turk
run-ing	*(running)$_\omega$*	*(run)$_\omega$-ing*
cool-er	*(cooler)$_\omega$*	*(cool)$_\omega$-er*
cool-ness	*(cool)$_\omega$-ness*	*(cool)$_\omega$-ness*
cool-ly	*(cool)$_\omega$-ly*	*(cool)$_\omega$-ly*

Turning to the prosodic word status of English prefixes, two approaches are to be discussed, Wennerstrom (1993) and Raffelsiefen (1999). Wennerstrom (1993) suggests the distinction between analyzable and non-analyzable prefixes. Prefixes which are analyzable form prosodic words, prefixes which are not analyzable do not. The key criterion for analyzability is focusability. If a prefix can be focused, it is analyzable and forms a prosodic word. Importantly, prosodic word status depends on individual derivatives and not on the prefix involved. This means that the same prefix might be considered a prosodic word in one derivative but not in another (Wennerstrom 1993: 314). As an example, Wennerstrom (1993: 311) presents the word *external*, in which the prefix *ex-* can be focused in a sentence such as "*The country has both INternal and EXternal problems*". As shown in this example, analyzability is independent of whether the prefix's base is a bound root or a word.

One problem with Wennerstrom's approach, as pointed out by Raffelsiefen (1999: 161f.), is that focusability does not correlate with phonological characteristics of p-words. Furthermore, it seems that almost every prefix can be focused under certain conditions (see also Plag 2003: chapter 4). In normal, conversational speech there might be the additional problem of determining whether a prefix is focused or not. For these reasons focusability is not considered a useful, reliable criterion for determining prosodic word status.

Raffelsiefen (1999) proposes that segmentability correlates with prosodic word status. In more decomposable words, i.e. in derivatives in which the prefix is more segmentable, prefixes form independent prosodic words. In less decomposable words, they do not. She suggests that morphological structure, i.e. decomposability, is translated into prosodic structure, which in turn is mirrored in the phonetic realization of a derivative. Derivatives made up of only one prosodic word,

i.e. derivatives with less segmentable prefixes, are realized similarly to simplex words. The phonetic realization of prefixes which form prosodic words on their own, i.e. highly segmentable prefixes, is not affected by their base. A prefix's degree of segmentability in a given derivative is defined by the semantic and phonological properties of the derivative. Semantic and phonological aspects of segmentability are assumed to correlate with each other, an assumption which seems to be true for the affixes of this study (cf. §3.2). On top of decomposability, Raffelsiefen (1999: 175f.) assumes an influence of frequency on prosodic word structure. She suggests that derivatives of very high frequency are more likely to be parsed as a whole. In other words, high frequency derivatives are processed as single prosodic words irrespective of decomposability.

Table 4.1: Criteria for prosodic word status of English prefixes (Raffelsiefen 1999)

Criterion	Prosodic word $(prefix)_\omega(base)_\omega$	No prosodic word $(prefix\ base)_\omega$
Stress	Secondary stress on prefix	Primary stress on prefix or unstressed prefix
	(e.g. ˌinˈtolerant, ˌdisˈhonor)	(e.g. ˈimpotent, inˈdifferent)
Syllabification	No syllabification of prefixal coda	Syllabification of prefixal coda
	(e.g. *dis.integrate*)	(e.g. *di.sease*)
Aspiration	Aspiration of base-initial stops	No aspiration of base-initial stops
	(e.g. *dis[kʰ]olor*)	(e.g. *dis[k]over*)
Flapping	No flapping of base-initial stop	Flapping of base-initial stop
	(e.g. *in[tʰ]olerant*)	(e.g. *in[ɾ]egrate*)
Type of derivative	Mostly of native origin	Mostly loanwords
	(e.g. *unpleasant, unjust*)	(e.g. *inject, impotent*)
Meaning	Strictly compositional	Not compositional
	(e.g. *unpleasant, impolite*)	(e.g. *inject, impotent*)
Type of base	Word as a base	Bound root as a base
	(e.g. *unpleasant, impolite*)	(e.g. *inject, innocent*)

To determine the segmentability of an affix, i.e. prosodic word status of a prefix, Raffelsiefen (1999) proposes a number of criteria. The criteria relevant for the

prefixes under investigation are displayed in Table 4.1. Unfortunately, not all of the criteria turn out to be useful. Since prefixal stress is very hard to determine (especially in derivatives with base-initial primary stress) the stress-criterion is not suited to reliably determine prosodic word status (cf. §3.1.1 for discussion of prefixal stress). Similarly, syllabification cannot always easily be determined and is thus not helpful. With regard to aspiration and flapping, these criteria are not applicable to all derivatives since they are restricted to specific sounds. In all derivatives with a double consonant a vowel follows the consonant, i.e. for none of these words the aspiration or the flapping criterion is applicable. Therefore, these two criteria cannot be used in this study.

The type-of-derivative-criterion merely displays a tendency and can therefore not be applied. This leaves us with two useful criteria: meaning and type of base. Prefixes which are semantically transparent and feature a word as a root are assumed to form independent prosodic words. Prefixes which are semantically opaque and feature a bound root are assumed to form prosodic words together with their base.

There are two problems with the prosodic word approach and its predictions for gemination. The first problem is that the predictions made are restricted to certain combinations. No predictions are made for semantically opaque words with words as bases, or for semantically transparent words with bound roots. The second problem refers to the fact that only two criteria are applicable to determine prosodic word status, meaning and type of base. This is problematic because both criteria are lexical, i.e. not prosodic. Neither the meaning of a word nor its type of base can serve as direct evidence for a word's prosodic structure, and one must acknowledge the possibility that possible effects of the two factors on gemination are not due to prosodic word structure but to other lexical mechanisms. Nonetheless, it is justified to apply the two criteria *meaning* and *type of base* to test prosodic word status (as defined by Raffelsiefen).

According to Raffelsiefen's (1999) approach, meaning and type of base correlate with prosodic aspects of phonological word structure, e.g. stress and resyllabification. Therefore, these two lexical criteria can be applied to determine prosodic word status. One must, however, keep in mind that determining prosodic word status by applying lexical measures of decomposability is based on the prosodic word approach by Raffelsiefen (1999), i.e. by the assumption of a close connection between prosodic and lexical word structure. There is the possibility that effects of meaning and type of base are independent from effects of prosodic structure. Prosodic criteria, such as syllabicity and stress, would be more direct determiners of phonological word structure. As described above, they are, how-

ever, not applicable, and further research on these prosodic criteria is needed to validly test phonological word status.

Table 4.2: Prosodic word statuses of *un-*, *in-*, *dis-* and *-ly*

Affix	*(base)$_\omega$ suffix*	*(prefix)$_\omega$(base)$_\omega$*	*(prefix base)$_\omega$*
un-		*(un)$_\omega$(natural)$_\omega$*	
in-		*(im)$_\omega$(polite)$_\omega$*	*(inject)$_\omega$*
dis-		*(dis)$_\omega$(trust)$_\omega$*	*(dissipate)$_\omega$*
-ly	*(cool)$_\omega$ -ly*		

Table 4.2 summarizes the analysis of the prosodic word status of *un-*, *in-*, *dis-* and *-ly*. The predictions for gemination made by the prosodic word approach are based on this analysis. Affixes forming independent prosodic words are expected to geminate, affixes which do not form independent prosodic words are predicted to degeminate. According to Raffelsiefen's approach, the three prefixes *un-*, *in-* and *dis-* form prosodic words when they are part of a semantically transparent derivative with a word as a base. The prefixes are integrated in the prosodic word of their base when they are part of a semantically opaque derivative with a bound root. Note that, as discussed in §3.1.1, derivatives with *un-* are always semantically transparent and feature words as a base. The prefix *un-* is thus expected to always form a prosodic word and to always geminate. Gemination with *in-* and *dis-* is predicted to depend on the type of base the prefix takes in a given derivative. For semantically transparent derivatives with words as bases, gemination is predicted. For opaque derivatives with bound roots, degemination is predicted. The suffix *-ly* is always preceded by a prosodic word boundary. It does, however, never form a prosodic word on its own. It is thus expected to degeminate. As noted above, Raffelsiefen (1999) additionally states that very frequent derivatives are parsed as a single prosodic word. Therefore, an effect of frequency on gemination is assumed. More frequent derivatives are more likely to degeminate. The predictions are summarized below.

Prosodic Word (Raffelsiefen 1999): Predictions

1. *un-* geminates
   The nasal in *un-*-prefixed words with a phonological single consonant will be shorter than the nasal in *un-*-prefixed words with a phonological double consonant.

2. *in-* degeminates in semantically opaque derivatives with a bound root
The nasal in *in*-prefixed words with a phonological single consonant will be as long as the nasal in *in*-prefixed words with a phonological double consonant if the derivative is semantically opaque and has a bound root as its base.

3. *in-* geminates in semantically transparent derivatives with a word as a base
The nasal in *in*-prefixed words with a phonological single consonant will be shorter than the nasal in *in*-prefixed words with a phonological double consonant if the derivative is semantically transparent and has a word as its base.

4. *dis-* degeminates in semantically opaque derivatives with a bound root
The fricative in *dis*-prefixed words with a phonological single consonant will be as long as the fricative in *dis*-prefixed words with a phonological double consonant if the derivative is semantically opaque and has a bound root as its base.

5. *dis-* geminates in semantically transparent derivatives with a word as a base
The fricative in *dis*-prefixed words with a phonological single consonant will be shorter than the fricative in *dis*-prefixed words with a phonological double consonant if the derivative issemantically transparent and has a word as its base.

6. *-ly* geminates
The lateral in *-ly*-suffixed words with a phonological single consonant will be shorter than the lateral in *-ly*-suffixed words with a phonological double consonant.

7. Post-lexical factors affect consonant duration.

8. Derivatives with higher token frequency will more likely degeminate.

4.3 Psycholinguistic approaches to morphological processing

Empirical studies have found that the morphological structure of a derivative influences its acoustic realization, and that there thus is an effect of morphol-

ogy on fine phonetic detail (see, for example, Sproat & Fujimura 1993; Cho 2001; Sugahara & Turk 2009; Pluymaekers et al. 2010; Smith et al. 2012; Lee-Kim et al. 2013; Plag et al. 2017, see also discussion in §4.2.3). In the previous section, this interaction between morphological and phonetic structure was explained by assuming a mediator between the two levels, i.e. the phonological word. However, there are certain restrictions connected with this assumption. Phonological word boundaries are restricted to mirroring categorical differences in morphological structure (for example simplex versus complex words, or weak versus strong boundaries). It follows that the prosodic word approach can only explain categorical, i.e. non-gradient, effects of morphological structure on phonetic detail. In contrast, some psycholinguistic approaches to morphological processing expect gradient, probabilistic effects of morphological structure on phonetic detail. In general, these approaches assume morphological structure to be gradient and to directly influence the phonetic realization of derivatives, i.e. they assume a direct morpho-phonetic interface.

Different from formal linguistic theories, psycholinguistic approaches are mainly based on empirical studies. On the one hand, these studies provide empirical support for the assumptions made by the approaches; on the other, differences in the studies' methodologies and outcomes preclude general, uniform theoretical assumptions about the morpho-phonetic interface. Assumptions and predictions are less formalized, less stream-lined and less precise than the ones of formal linguistic approaches. The predictions made are rather probabilistic, often revolve around one theoretical concept and often focus on the properties of individual words. Hay (2001; 2003), for example, investigated the influence of decomposability on the phonetic realization of complex words. She found more reduction with less decomposable derivatives than with more decomposable derivatives. Pluymaekers et al. (2010) investigated the influence of an affix's informativeness on its phonetic realization and found that the more informative a linguistic unit is, the less phonetic reduction is found. Cohen (2014) conducted a study on the influence of an affix's paradigmatic probability on its duration. Her results showed that the more probable a suffix is, the shorter the preceding stem is pronounced.

In this study, I will focus on two frequently investigated and discussed concepts in psycholinguistic approaches: DECOMPOSABILITY and MORPHOLOGICAL INFORMATIVENESS. Below I will lay out how and why they are assumed to influence the phonetic realization of affixed words. This discussion includes the discussion of models of word storage and morphological processing. Furthermore, I will provide an overview of previous empirical work. At the end of each section, I will present how each of the two factors is expected to influence gemination with *un-*,

in-, *dis-* and *-ly*.

4.3.1 Decomposability

The idea that decomposability influences the phonetic realization of complex words is closely connected to dual-route models of morphological processing. These models assume that complex words are simultaneously stored as a whole and in their parts (cf., for example, Frauenfelder & Schreuder 1992; Schreuder & Baayen 2015; deVaan et al. 2011; Caselli et al. 2016). According to these models, an affixed word like *unnatural* has two separate entries in the mental lexicon, the whole-word form (*unnatural*) and the decomposed form, which consists of two separate entries (*un-* and *natural*). Dual-route models stand in opposition to models which assume that a complex word is exclusively stored in its parts (cf., for example, Prasada & Pinker 1993; Marcus et al. 1995; Clahsen 1999; Pinker & Ullman 2002). In these models, only irregular and non-transparent lexicalized forms are stored as a whole. Regular complex words, such as *unnatural*, are always stored in their decomposed form and computed online.

It is generally assumed that the way a word is accessed, i.e. as a whole or via computation, influences its phonetic realization. A word that is accessed as a whole will more likely show phonetic reduction than a word which is computed online. This is due to the fact that in speech processing morphemes are recognized via their phonological segments, which in turn means that segments that represent morphemes should be more resistant to reduction. It follows that models that assume all complex words to be computed online, and all simplex words to be stored, predict systematic differences in the phonetic realization of complex and simplex words. Simplex words are expected to show more reduction. As discussed above, some empirical studies have indeed found the expected differences (see, for example, Cho 2001; Sugahara & Turk 2009; Smith et al. 2012). These purely computational models, i.e. the models predicting all complex words to be accessed via their parts, do, however, not predict systematic differences between different complex words. Dual-route models in contrast predict these differences. Complex words accessed via the whole word route are expected to be phonetically more reduced than complex words accessed via the decomposed route.

Hay (2001; 2003) proposes that when accessing a complex word, both routes are activated. She argues that decomposability is the main factor governing which route is faster in accessing the complex word. The assumption is that the more decomposable a complex word is, i.e. the more easily segmentable it is, the more likely it is accessed via its individual parts. The less decomposable a complex

word is, i.e. the less easily it is to segment, the more likely it is accessed as a whole. Hay (2001; 2003) suggests RELATIVE FREQUENCY as the central measurement for decomposability. Relative frequency is defined as the frequency of a derivative relative to the frequency of its base. If a derivative is less frequent than its base, the derivative is rather decomposable and thus likely to be accessed via the decomposed route (e.g. *impossible, softly*). If a derivative is more frequent than its base, it is less decomposable and thus likely to be accessed as a whole (e.g. *inject, swiftly*). The influence of relative frequency on the access route is explained by resting activations. A lexical entry of high frequency has a higher resting activation than an entry that is less frequently accessed. Higher resting activations lead to faster access, which in turn means that the route which is more frequently used will win the race when accessing a complex word. In other words, if the derivative is more frequent than its base, its resting activation is higher and the whole word route is faster. If the base frequency is higher than the whole word frequency, the decomposed route has a higher resting activation and wins the race.

Figure 4.4 schematically depicts the race between the two access routes for the word *unnatural*. The dashed arrow indicates the direct route and the solid arrows depict the decomposed route. The nodes represent the lexical entries for *unnatural*, and the line width of a node indicates its activation level[7]. A very frequent lexeme has a high resting activation (indicated by a thick line width). In our example, the frequency of *natural* is much higher than the frequency of *unnatural* (72.451 vs. 2.025 in COCA). This indicates that the resting activation of *natural* is higher than the one of *unnatural*. Therefore, it is assumed that the base *natural* is accessed faster than the derivative *unnatural*. In other words, the decomposed route is accessed faster than the whole word route. Note that the difference in frequencies, i.e. a much higher base frequency than derivative frequency, indicates that *unnatural* is a highly decomposable derivative. We can thus summarize that, according to the decomposability approach, the highly decomposable word *unnatural* is predicted to be accessed via its parts and not via the whole word route.

As explained above, the route of access is assumed to be mirrored in the phonetic realization of a complex word. A word accessed via its parts is expected to display less reduction than a word accessed as a whole. Reduction is particularly expected at the morphological boundary. With regard to morphological gemination, a cross-boundary phenomenon, one can thus expect words which

---

[7]Note that the activation level for the prefix *un-* is not depicted in the picture. The role of the affix in the model will be discussed later in the chapter.

Draft of May 27, 2019, 18:04
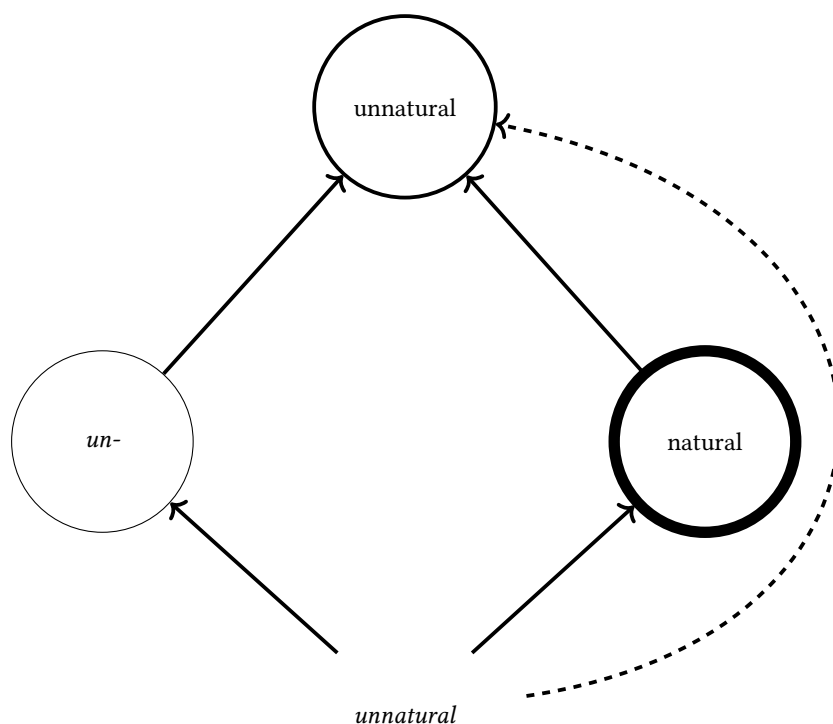
Figure 4.4: Dual-route model as in Hay (2001)

are less decomposable, and which are thus accessed as a whole, to be likely to degeminate, i.e. to show reduction. More decomposable words are expected to geminate, i.e. show no or less reduction. Turning back to the example above, one would thus expect gemination with *unnatural*. Its high decomposability, and consequently the decomposed route via which it is accessed, suggests a low degree of reduction of the morphological geminate.

Empirical evidence for the effect of decomposability on the realization of complex words is contradictory and not conclusive. In the following, I will discuss six studies which investigated decomposability and found mixed results. The studies point at several complications with the approach. These complications have to be addressed in order to test the role of decomposability in morphological gemination, and in morphological processing in general.

Hay (2003) investigated base-final /t/-reduction in -*ly*-suffixed words. By comparing word pairs, she tested whether the stop is more reduced in less decom-

posable words than in more decomposable words. In the five word pairs she compared, one derivative was more decomposable than the other, i.e. one derivative was less frequent than its base (e.g. *softly*) and one was more frequent than its base (e.g. *swiftly*). The less decomposable words were expected to show more deletion and shorter durations. Hay found the expected effect. However, as pointed out by Hanique & Ernestus (2012), Hay's methodology poses some problems. The number of investigated types is very small and hence hardly sufficient to draw general conclusions. Furthermore, instead of testing the direct effect of decomposability on duration, Hay compared average durations using a ranking system. This comparison does not allow for a straightforward interpretation. It is thus questionable how robust the effect found really is.

In Hay (2007), *un*-prefixed words were investigated. It was tested whether more decomposable derivatives have a longer prefix duration than less decomposable derivatives. In contrast to Hay (2003), this corpus study tested the influence of relative frequency on duration directly. The speech of two groups of speakers was investigated in the study: the speech of *early speakers*, who were born before 1920, and the speech of *late speakers*, who were born after 1920. While the expected effect of relative frequency on prefix duration was found for the early speakers, there was no effect for the late speakers. Hay's explanation for the null result with late speakers is that those speakers use *un-* less productively. Therefore, she claims, the prefix is less decomposable for late speakers in general. It is thus assumed that the segmentability of the prefix, which is not directly accounted for in relative frequency, interferes with the effect of relative frequency. In other words, the gradient measure of decomposability alone, i.e. relative frequency, might not be able to capture a derivative's decomposability, and it might be necessary to also take an affix's average segmentability into account.

Collie (2008) conducted a study on the effect of relative frequency on stress preservation with suffixes. While it was expected that derivatives which are accessed via the whole word route are less likely to maintain their prosodic structure, complex words accessed via the decomposed route were assumed to preserve their prosodic structure by demoting primary stress to secondary stress (e.g. *acˈcelerate – acˌceleˈration*). Derivatives which are more frequent than their base, i.e. derivatives of low decomposability, were thus expected to exhibit non-preserving behavior. While Collie (2008) found the expected effect for the suffix *-ion*, the effect was not found for the suffix *-ity*. This might mean that only some affixes are affected by gradient decomposability, i.e. relative frequency.

For Dutch, Hanique et al. (2011) investigated the influence of relative frequency on the duration of schwa in prefix-final position. Three different prefixes (*ge-*,

be- and *ver-*) were investigated. Only for one of the three prefixes relative frequency showed a significant effect. For *ge-* a higher relative frequency, i.e. a higher derivative frequency and a lower base frequency, led to shorter schwa durations. Hanique et al. (2011) explain the differences in results between the prefixes by referring to differences in semantic opacity. While they analyze *ge-* to be semantically transparent, the other two prefixes are analyzed as semantically opaque. It is suggested that relative frequency only affects transparent affixes since opaque derivatives are always retrieved as a whole from the mental lexicon. Similar to Collie (2008), this study thus reveals that not all affixes are affected by relative frequency and that categorical factors of decomposability, such as semantic transparency, might interfere with gradient effects of decomposability.

That relative frequency does not have the same effect for all affixes is also shown by Schuppler et al. (2012). In their data set on Dutch complex words, derivatives with a higher relative frequency, i.e. a lower decomposability, show less word-final /t/-deletion than derivatives with a lower relative frequency, i.e. a higher decomposability. In other words, less reduction is found with less decomposable derivatives. This result goes counter the assumption that less decomposable words show more reduction. Schuppler et al. (2012) explain their findings with reference to the informativeness of the affix, an idea further discussed in the next section.

The five studies mentioned above primarily concentrated on relative frequency as a measurement of a derivative's decomposability. The results reveal, however, that other measures of decomposability, i.e. semantic transparency (cf. Hanique et al. 2011) and the productivity of the affix (cf. Hay 2007), might also be predictive for the realization of complex words. Bürki et al. (2011) investigated yet another measure of decomposability. They used a decomposability rating to predict schwa reduction in French complex words. Words rated as being less decomposable were expected to show more reduction. No effects were found. There are several possible explanations for the null result. The first (and most straightforward) explanation is that there is no effect of decomposability on schwa reduction. This explanation needs to be tested by further research. The second explanation is that a rating is not a suitable measure of decomposability. However, contra this explanation, Hay (2001; 2003) showed that there is a significant correlation between decomposability ratings and relative frequency. If relative frequency is a suitable measure of decomposability, so should be decomposability ratings. It might, however, be the case that one of the two measures is a better predictor for durations in a particular type of model. Differences in distributions and fine-grainedness

of the measure might cause different effect sizes, as well as differences in significance between measures. The third explanation is also related to methodological aspects. In Bürki et al. (2011), mean ratings were used and the authors explicitly mention the skewed distribution of the ratings, which, as just explained, might have influenced the results. Furthermore, only the ratings of five participants were included. This might have affected the validity of the rating in general. The question whether decomposability ratings are suitable to operationalize decomposability thus remains open at this point.

The review of the empirical work on decomposability has revealed that the empirical facts are unclear and inconclusive. Some studies have found the expected effect of decomposability and some have not, one has even found the opposite effect (Schuppler et al. 2012). One major reason for the incongruities between studies can be seen in their diverging methodologies. Two important issues, which are also relevant for the present investigation, have to be addressed here. First, the studies investigated different phenomena in different domains (e.g. segment deletion in the base, affix duration, stress shift), and it might be that there are differences with regard to the influence of decomposability depending on the domain investigated. Second, the operationalization of decomposability varied across studies. This problem is also discussed by Hanique & Ernestus (2012: 16), who state there is no uniform definition of decomposability and that "[f]urther studies have to provide a better definition of morphological decomposability before we further investigate the role of morphological structure in speech production". The most frequently used measure of decomposability is relative frequency, defined as the frequency of the derivative relative to its base. But even this seemingly well-defined measure was applied in different ways across studies. While some studies used it as a gradient measure (e.g. Hanique et al. 2011; Schuppler et al. 2012), in others it was used in categorical terms. Hay (2001) and Collie (2008), for example, compared more decomposable with less decomposable words, i.e. instead of testing the gradient effect of relative frequency on phonetic realization, they redefined relative frequency as a binary measure. In addition to relative frequency, the studies referred to semantic transparency (cf. Schuppler et al. 2012), productivity (cf. Hay 2007) and a decomposability rating (cf. Bürki et al. 2011) as measures of decomposability. Furthermore, one can think of additional operationalizations of decomposability. Hay (2003), for example, shows that phonological transparency correlates with relative frequency. Additionally, the type of base of a derivative or semantic similarity measures might be used.

A second reason for differences in results might be related to the affix under investigation, as also suggested by some of the authors themselves (e.g. Hanique

et al. 2011; Schuppler et al. 2012). It might be the case that only certain affixes are affected by certain decomposability measures (cf. Collie 2008; Hanique et al. 2011; Schuppler et al. 2012). It might, for example, be that only very productive, transparent affixes are affected by the a derivative's particular decomposability (as suggested by Hanique et al. 2011). This idea is supported by Hay's (2007) data, in which prefix duration is only affected by relative frequency when the prefix is productive. The informativeness of an affix might also play a role, as suggested by Schuppler et al. (2012) (see also next section fur further discussion). The idea that only transparent affixes are influenced by decomposability entails the assumption that only derivatives with transparent affixes can be accessed via the decomposed route, i.e. that derivatives with opaque affixes will always be accessed as a whole. This in turn would mean that the predictions made by the decomposability approach are not only word-specific but must also concern the affixes involved.

At first sight, the predictions made by the decomposability approach seemed quite clear: in complex words of high decomposability the morphological geminate will geminate, and in complex words of low decomposability the morphological geminate will degeminate. The discussion above has, however, shown that after all, the predictions are not that straightforward. How can one test the influence of decomposability, i.e. operationalize the concept? Which role does the affix play? Furthermore, one might raise the question of whether, according to the decomposability approach, degemination is predicted to be categorical or gradient.

To address the problem of how to operationalize decomposability, I will include five possible measures of decomposability in my studies: relative frequency, semantic transparency, type of base, a decomposability rating and semantic similarity scores. I will first investigate their relation to each other in order to validate their assumed correlation, i.e. ensure that they all tap into the same underlying concept. I will then test the predictions for gemination using the different measures. I will thus find out which variable is best suited to predict phonetic reduction in terms of duration.

Let us now turn to the role of the affix. In addition to investigating word-specific decomposability, I will also test affix-specific decomposability. There are two possible explanations of why we might find affix-specific effects. First, it might be the case that derivatives of one affix are so similar in their decomposability that practically all derivatives behave uniformly with regard to gemination. The second possibility is that, as suggested by previous research, only some of the affixes under investigation are affected by decomposability. To explore these

possibilities, it is necessary to investigate the segmentability of the five affixes under investigation.

In §3.2 we have already seen that there seem to be systematic differences in segmentability between the affixes. The two segmentability hierarchies which refer to lexical decomposability, i.e. the two hierarchies relevant for the decomposability approach, are depicted in Figure 4.5. As discussed earlier, the two hierarchies deviate in the placing of the suffix *-ly*, which is debatable and which depends on the role of semantics in decomposability – an issue which will be discussed further in the next section. According to the decomposability approach, affixes which are more segmentable, i.e. affixes which are higher on the segmentability hierarchy, are expected to geminate. Affixes which are less segmentable, i.e. affixes which are lower on the hierarchy, are expected to degeminate. Furthermore, one might expect differences in the degree of gemination depending on the affix's position in the hierarchies. Gemination with more segmentable affixes is expected to be stronger than gemination with less segmentable affixes. The strength of gemination is expected to be mainly indicated by the durational differences between phonological doubles and phonological singletons. Stronger gemination goes together with larger singleton-double ratios.[8]

I will test the affix-specific decomposability predictions by first validating the segmentability status of the five affixes, i.e. I will look at the distributions of the

---

[8]Note that I will use the terms STRENGTH OF GEMINATION and DEGREE OF GEMINATION interchangeably in this book. An affix which shows strong gemination, geminates to a high degree. An affix which shows weak gemination, geminates to a low degree.

	Segmentability hierarchy	Additional assumption
Semantic Hierarchy	*un-* > {*dis-*, *in-*$_{\text{NEG}}$}> *in-*$_{\text{LOC}}$ > *-ly*	lexical meaning over productivity, transparency and type of base
Non-Semantic Hierarchy	*un-* > *-ly* > {*dis-*, *in-*$_{\text{NEG}}$}> *in-*$_{\text{LOC}}$	productivity, transparency and type of base over lexical meaning

Figure 4.5: Lexical segmentability hierarchies of affixes

different decomposability measures across affixes, and will thereby test whether the theoretically-based hierarchies are borne out by the data. I will then test whether there are significant differences in gemination behavior between affixes. If so, I will compare these differences with the segmentability hierarchies proposed. In other words, I will check whether the differences in the degree of gemination between affixes mirror the differences in their segmentability.

Let us now turn the question of the nature of (de)gemination. The decomposability approach does not make clear predictions about the nature of degemination. Degemination can thus either be gradient, i.e. the double consonant shows more or less reduction, or categorical, i.e. in case of degemination there is no durational difference between singletons and doubles. If one assumes degemination to be categorical, one simultaneously assumes gemination to be governed by categorical factors. For certain categories, phonological doubles are categorically longer than phonological singletons. For others, there is no durational difference between doubles and singletons. If one assumes degemination to be gradient, one assumes gemination to be governed by gradient, word-specific factors.

In this study, I will explore both possibilities by investigating the distribution of durations across singletons and doubles, and by investigating which kind of factors govern gemination. If gemination is categorical, the durations of singletons and doubles should show a bimodal distribution. Doubles should be longer than the singletons. If gemination is gradient, one would expect a gradient increase in duration from singletons to doubles, i.e. no binary distribution (cf. Hanique et al. (2013) for a similar analysis of distributions to investigate the nature of schwa reduction in Dutch). Furthermore, in case of gradient gemination, the durational difference between doubles and singletons should be affected by word-specific factors, such as, for example, relative frequency. In case of categorical gemination, the durational difference between doubles and singletons should be affected by categorical factors, such as, for example, the affix.

To summarize, one can state that up to this point the decomposability approach does not make clear, spelled-out predictions for gemination. Before testing the effect of decomposability on gemination, it is necessary to explore the concept of decomposability with its possible operationalizations and domains. Also, the approach does not make assumptions about the nature of gemination, i.e. it is yet to explore whether gemination is gradient or categorical. For these reasons the predictions uttered at this point remain relatively vague.

Two different predictions with regard to the effect of decomposability are formulated. One concerns the decomposability of the individual derivative, and one the segmentability of the affix. They will be tested using different decompos-

ability measures. i.e. relative frequency, semantic transparency, type of base, a decomposability rating and semantic similarity scores. The two predictions are spelled out below.

Morphological Segmentability Hypothesis: Predictions

A: The decomposability of an individual word influences gemination
The more decomposable a derivative is, the higher is its degree of gemination.

The less decomposable a derivative is, the lower is its degree of gemination.

B: The segmentability of an affix influences gemination
The more segmentable an affix is, the higher is the degree of gemination with words containing that affix.
The less segmentable an affix is, the lower is the degree of gemination with words containing that affix.

In addition to the two decomposability predictions, two predictions with regard to the nature of gemination are formulated. One predicts gemination to be categorical, the other predicts gemination to be gradient. They will be tested by investigating the distribution of durations in the data sets and by investigating the type of effects governing gemination. Note that the two predictions concerning the nature of gemination are not exclusive to the decomposability approach but concern all approaches discussed. While formal approaches explicitly predict gemination to be categorical, the psycholinguistic approaches leave the question open. The two predictions are displayed below.

Nature of gemination: Predictions

A: Gemination is categorical

The duration of the affixational consonant(s) in the data set shows a bimodal distribution with one mode representing doubles and one mode being singletons. Doubles are longer than singletons.
Gemination is governed by categorical factors.

B: Gemination is gradient

The duration of the affixational consonant(s) in the data set does not show a bimodal distribution with one mode representing doubles and one mode being singletons.

The duration of the affixational consonant(s) increases gradually from singletons to doubles.

Gemination is governed by word-specific factors.

4.3.2 Morphological informativeness

The idea that linguistic units with high information load are less prone to phonetic reduction than linguistic units with low information load is well established in psycholinguistic approaches (cf., for example, Aylett & Turk 2004; Kuperman et al. 2007; Pluymaekers et al. 2010; Hanique & Ernestus 2012). Following Lindblom (1990), in speech production two forces work against each other: economy of articulatory effort and discriminability of the speech signal. On the one hand, speakers want to put as little effort as possible in producing speech, on the other, they want to ensure the intelligibility of the speech signal. As a result, they only put as much effort in pronunciation as they estimate to be necessary for the listener to discriminate the speech signal. A speaker's amount of effort is mirrored in the degree of reduction found in the speech signal, and depends on the information load of the linguistic unit. The degree of information load in turn is influenced by various factors, such as probability of occurrence, semantic information load and redundancy (see also Kuperman et al. (2007) for discussion). Elements with low information load, i.e. redundant elements with higher probabilities of occurrence and less semantic information load, are realized with less effort and are hence more reduced than elements with higher information load. They are shorter and less salient.

The elements under investigation in this study are affixes. Following the approach just described, one can assume affixes with higher information load to show less reduction than affixes with lower information load. Degemination can be defined as some sort of reduction. Even though the nature of gemination is yet unclear (see §4.3.1 for discussion on gradient vs. categorical gemination), it is certain that degemination results in some kind of phonetic reduction. It is thus predicted that affixes with higher information load are less prone to degemination than affixes with lower information load, and that the degree of (de)gemination depends on the degree informativeness of the affix.

To test this prediction, it is necessary to measure the information load of the affixes. In this study, the predictability of the affix and the semantic information load of the affix are used as indicators of informativeness. The predictability of

an affix is closely related to its probability of occurrence. If an affix is probable to occur, it is very predictable and thus not very informative.

Probability can be operationalized in various ways. Pluymaekers et al. (2010), for example, tested the effect paradigmatic probability on duration to investigate the effect of informativeness on phonetic reduction. They investigated Dutch -*igheid* (/əxhɛit/), which represents three different morphological structures. In some words -*igheid* represents one single suffix (-*igheid*-words), in some -*ig* belongs to the base word and only -*heid* is the suffix (-*heid*-words), and for some words both parsings are possible (*ambiguous words*). Pluymaekers et al. (2010) measured the informativeness of the investigated structures by counting the cohort of competitors in the morphological paradigm for each structure. The more competitor words in a paradigm, the less probable, and thus the more informative, the structure is. Since the paradigm for -*heid*-words is the least dense, i.e. this suffix has the least competitors, it is the least informative and most reduction is expected with words of this kind. Reduction was measured in terms of the duration of the /xh/ cluster. Pluymaekers et al. (2010) found the expected effect. The structure with least competitors in the morphological paradigm, i.e. the least informative structure, showed most reduction.

Similarly Schuppler et al. (2012) found a relation between informativeness in terms of number of competitors in the morphological paradigm and reduction. As laid out in the previous section, in their data complex words with a low relative frequency, i.e. highly decomposable words, showed more base-final /t/ reduction than less decomposable words. This goes against the assumptions made by the decomposability approach, and is the opposite of what was found for English adverbial -*ly* in Hay (2003).

Schuppler et al. (2012) explain the difference between their and Hay's results by referring to differences in informativeness between the two investigated structures. They hypothesize that relative frequency might play a different role for suffixes with higher information load, i.e. less probable suffixes, than for suffixes with lower information load. According to Schuppler et al. (2012), English -*ly* is more probable and less informative than the Dutch suffix -*t*. English adverbs always end in the suffix -*ly*, i.e. it has a high paradigmatic probability. The suffix is thus expected and does not feature a high information load. It differs in this respect from the inflectional Dutch suffix -*t*, which is only one of three possible forms in the inflectional paradigm. According to Schuppler et al. (2012), the suffix -*t* is therefore generally less probable than -*ly*. It has a higher information load and is not reduced. In other words, because Dutch inflectional -*t* is very informative, lower decomposability does not lead to more reduction with this

affix. Lower decomposability leads to more reduction with the less informative English adverbial suffix *-ly*.

The two studies above suggest that paradigmatic probabilities might be a good measure of predictability. However, this measure is not applicable in this study. The reason is that it is yet unclear how to measure the paradigmatic probability of derivational prefixes. To nevertheless investigate the predictability of the affixes in this study, another type of probability was looked at: syntagmatic probability.

As shown in Hanique et al. (2013), syntagmatic probability can also successfully be used as an indicator of predictability and informativeness. In Hanique et al. (2013), the deletion of word-final /t/ in Dutch past participles and in Dutch simplex words was investigated. The segment was more often deleted in complex than in simplex forms. Hanique & Ernestus (2012) argue that the results can best be accounted for by reference to the informativeness of the suffix *-t* in comparison to the informativeness of the segment /t/ in corresponding simplex structures. Since most Dutch participles end in /t/, the segment is highly predictable in complex words. Therefore, it is less informative than /t/ in corresponding simplex words. The higher reduction rate of /t/ in complex words can thus be explained by the low degree of informativeness of the suffix *-t*.

Let us now turn to the affixes under investigation in this study. Out of the five investigated affixes, the suffix *-ly* features the highest syntagmatic probability. Due to its function to create adverbs, its syntagmatic probability is very high, and, in turn, its predictability is very high. As the prefixes are not associated with a specific function, they are much less predictable. Furthermore, the suffix *-ly* is more predictable than the prefixes because of its position at the end of the word. Prefixes precede their base, which means that the base of prefixes does not serve as a cue for the occurrence of the prefix. Prefixes can basically occur after any word after which its base can occur. For example, the prefix *un-* as in *uncool* can occur after any word after which the word *cool* can occur. In contrast, suffixes follow their base. Their occurrence is restricted by the amount of base words they can take, and their base serves as a cue for their occurrence. It follows that the syntagmatic probability of prefixes, and in turn their predictability, is generally lower than the one of suffixes.

Overall, the discussion of predictability has shown that the suffix *-ly* is the most predictable affix in this study. It is much more predictable than the prefixes in this study. The differences in predictability between the prefixes is less clear. As explained above, paradigmatic probabilities are not useful in order to measure a prefix's predictability. Furthermore, it seems quite challenging to assess differences in the syntagmatic probability of derivational prefixes. As prefixes

occur before their base, and as they are not associated with a specific function, it is unclear on which base one can compare their probability of occurrence.

In addition to predictability, the semantic information load of an affix can also serve as a measure of its informativeness. An affix that contributes more to the meaning of a derivative is more informative than an affix which does not feature clear semantic content. Semantic information load is particularly interesting as a measure of informativeness for derivational prefixes (as the ones in this study), for which the application of other informativeness measures is quite problematic.

The semantics of the affixes under investigation were discussed thoroughly in §3.2. While some affixes, such as *un-*, feature a stable, transparent meaning, others, like locative *in-* and *-ly*, do not. The Semantic Segmentability Hierarchy in Figure 4.5 depicts the decline in semantic information load of the five investigated affixes. The lower the affix is positioned on the hierarchy, the less semantic information it conveys, and the less informative it is. The hierarchy shows that the suffix *-ly* is the least informative affix with regard to its semantic information load. As discussed earlier, the suffix does not contribute any lexical meaning to the derivative. For the prefixes, *un-* is the prefix with the most transparent and stable meaning, i.e. the affix with the highest information load. Negative *in-* and *dis-* denote a stable, negative meaning in most derivatives but there are also some derivatives in which the affix does not contribute a clear lexical meaning. Locative *in-* features the least semantic information of the prefixes (see §3 for detailed discussion of the semantics of all five affixes).

One can summarize that, due to its high predictability and low degree of semantic contribution to a derivative's meaning, the suffix *-ly* is the least informative affix in this study. It is thus expected to show the weakest degree of gemination. The analysis of the prefixes' informativeness is mainly based on semantic factors. The analysis revealed that *un-* is the most informative prefix. It is thus expected to show the highest degree of gemination. The informativeness of locative *in-*, negative *in-* and *dis-* is less clear and might vary among types. In derivatives with transparent meaning, the affixes are more informative than in derivatives with opaque meaning. This means there are two possible predictions for the gemination with these three prefixes: an affix-specific one and a word-specific one.

According to the affix-specific prediction, one would predict the overall informativeness of the affix to govern gemination. This means one would predict gemination to pattern according to the Semantic Segmentability Hierarchy. While gemination with *un-* is expected to be the strongest, and gemination with *-ly* is expected to be the weakest, the other three prefixes are expected to pattern in between. Importantly, one would not predict word-specific effects according to

this prediction.

According to the word-specific prediction, one would predict informativeness to be word-specific, i.e. in semantically transparent words a prefix is more informative than in semantically opaque words. Since *un-* is always semantically transparent, it is predicted to always display a high degree of gemination. For the other three prefixes the degree of gemination is expected to depend on semantic transparency. Opaque derivatives should display weaker gemination than transparent derivatives.

In this study, I will test both predictions made by the morphological informativeness approach, i.e. the affix-specific and the word-specific prediction. The predictions are summarized below.

Morphological Informativeness: Predictions

A: Affix-specific informativeness influences gemination
The more informative an affix is, the higher is the degree of gemination with words containing that affix.

The less informative an affix is, the lower is the degree of gemination with words containing that affix.

- Gemination patterns according to the Semantic Segmentability Hierarchy.

B: Word-specific informativeness influences gemination
The more informative an affix is in a given derivative, the higher is the degree of gemination in the derivative.

The less informative an affix is in a given derivative, the lower is the degree of gemination in the derivative.

- Derivatives with the prefix *un-* geminate to a high degree.

- Derivatives with the suffix *-ly* geminate to the lowest degree.

- Derivatives with transparent semantics and the prefixes *dis-*, negative *in-* and locative *in-* geminate to a high degree.

- Derivatives with opaque semantics and the prefixes *dis-*, negative *in-* and locative *in-* geminate to a low degree.

4.4 Speech production models

Speech production models are closely connected to the approaches discussed in the previous sections. As formal linguistic and psycholinguistic approaches, they also make assumptions about the morpho-phonological interface and are thus relevant for this work. However, speech production models are broader than the approaches previously discussed in that they are not solely concerned with the morpho-phonological interface but with speech production as a whole. In other words, the morpho-phonological interface, and the processing of complex words, form just a small part of these models.

Two types of speech production models can be distinguished: modular feed-forward models (cf. Levelt 1999; Levelt et al. 1999; Levelt 2000) and usage-based models (cf. Johnson 1997; Bybee 2002; Pierrehumbert 2001; 2002). With regard to morphological processing, the two types differ in their assumptions of how words are stored, as well as which factors influence the phonetic realization of morphemes. Crucially, neither is very explicit with regard to the morpho-phonological interface and no specific claims about the interplay of phonetics and morphology are made. Therefore, no explicit predictions about gemination can be drawn from the models. The data in this study might nevertheless provide evidence for the theoretical modeling of speech production by displaying general effects on duration. By investigating which morphological factors influence segmental duration at morpheme boundaries, general assumptions about the morpho-phonological interface made by different types of speech production models can be tested.

In traditional speech production models, such as Levelt et al. (1999), two main stages of processing can be distinguished: the lexical stage and the post-lexical stage. At the lexical stage lemmas are retrieved and grammatically encoded. At the post-lexical stage the morpho-phonological and the phonetic encoding take place, i.e. the relation of morphological and phonetic structure is defined at this stage. After all morphemes of a lemma are activated and assembled, their morpho-phonological code is spelled out. This code is segmental in nature. The spelled-out segments are then syllabified to form phonological words.[9] In a last step utterance prosody is generated to compute the phonological score, which serves as the basis for articulation. In other words, the phonological score is translated into phonetics, which is in turn translated into concrete instructions for the articulators (articulatory score).

---

[9]Note that the term *phonological word* in Levelt et al. (1999) is not synonymous with the term in the prosodic word approach (cf. §4.2.3).

Draft of May 27, 2019, 18:04

The crucial point with regard to morpho-phonological processing is that articulation is based on phonemic representations, i.e. the segmental morpho-phonological code. The question is whether morphological structure is present at this stage, and if so, how it is mirrored in the acoustic realization of complex words. As discussed by Cohen-Goldberg (2013: 1037), most traditional models, such as the one just described or the one suggested by Dell (1986), are largely silent about the post-lexical processing of multi-morphemic words. It is simply not stated whether the assembly of morphemes leaves traces in the phonological make-up of a word. However, since none of the models mentions any process which suggests that morphological structure is preserved in morpho-phonological encoding, it can be assumed that phonemic representations do not feature any traces of morphological structure. This also means that, according to traditional models of speech production, there is no difference in the phonemic representation of morphologically complex and morphologically simplex words, i.e. there is no difference in their acoustic realization.

Recent studies have challenged this assumption by showing that phonemically identical strings of different morphological status vary systematically in their phonetic realization. For example, Kemps et al. (2005) and Blazej & Cohen-Goldberg (2015) found that phonologically identical free and bound variants of a base (e.g. *clue* without a suffix compared to *clue* in *clueless*) differ acoustically. As already discussed in §4.2.3, other studies demonstrate that the realization of segments can vary systematically depending on the type of boundary (affix, compound, phrase) they are adjacent to (e.g. Sproat & Fujimura 1993; Smith et al. 2012; Lee-Kim et al. 2013). Furthermore, empirical work found systematic durational differences between homophonous affixes. Plag et al. (2017) and Godfrey (2016), for example, found homophonous English suffixes to display systematic differences in duration.

The studies above thus challenge standard models of speech production and demand for modifications. These modifications must explain the effect of morphological structure on the acoustic realization of complex words. Cohen-Goldberg (2013), for example, suggests an extension of standard theories by proposing the HETEROGENEITY OF PROCESSING HYPOTHESIS. As a consequence of the morpheme assembly, the hypothesis predicts structural weaknesses at morphological boundaries of phonemic representations. Each morpheme acts as an independent domain for post-lexical processes, which will therefore apply more strongly to tautomorphemic phonemes than heteromorphemic phonemes. Heteromorphemic phonemes are predicted to be less integrated with each other than tautomorphemic phonemes. Furthermore, it is proposed that the phonemes in multimor-

phemic words will inherit the lexical properties of the morpheme they belong to, i.e. there will be differences in activation levels between different morphemes of one word (e.g. because of different frequencies of the constituents of the complex word). It is therefore predicted that those aspects of post-lexical processing that are influenced by lexical properties (e.g. duration, vowel space) will vary by morphemes. However, as noted by the author himself, the hypothesis is not fully specified yet. It calls for more empirical work, and must be elaborated to not only account for differences between simplex and complex words but also for differences between morphemes of varying boundary strength (cf. Cohen-Goldberg 2013: 1057f.).

Instead of modifying traditional models, one could also assume another type of speech production model, e.g. usage-based models that assume a direct morphophonetic interface. These models might be better suited to explain the phonetic implementation of morphologically complex words than traditional models. Exemplarbased models (cf. Johnson 1997; Pierrehumbert 2001; 2002; Bybee 2002), for example, assume that the phonetic realization of a word, simplex and complex, is determined by exemplars experienced by the speaker. It is assumed that all phonetic variants of a word are stored as exemplars in a speaker's memory. These exemplars are organized in a network structure. When producing a word similar exemplars are activated, and every activated exemplar influences the target production. More frequent exemplars have stronger representations and are therefore expected to influence pronunciation to a higher degree than less frequent exemplars. Since, in contrast to traditional models, word-specific, fine phonetic information is stored in these models, differences in the realization of words with different morphological structure are expected. Systematic differences might be explained by the proposed network structure. However, usage-based models are not very specific with regard to the network structure, i.e. they do not explicitly define at which levels similarities play a role and how these different levels interact. Furthermore, according to exemplar-based models, speech production is a speaker-specific process, i.e. differences in phonetic realizations can always be explained by differences in exemplar-structure between speakers. To conclude, up to now usage-based models do not make explicit predictions for the realization of complex words. Further specifications of the models are necessary to test their validity.

One can summarize that currently none of the proposed speech production models is able to model the phonetic realization of complex words, i.e. the relation of morphological structure and fine phonetic detail. Traditional models (e.g. Dell 1986; Levelt et al. 1999) are silent about the post-lexical processing of com-

plex words. Hypotheses which try to specify post-lexical processing, such as the heterogeneity of processing hypothesis, need further development to be able to account for different acoustic realizations of complex words. The same is true for usage-based models, which also need further specification. To accurately model the processing of complex words, it is necessary to conduct further studies. These studies might provide some new insight about which factors influence the phonetic realization of complex words, i.e. which factors must be incorporated in models of speech production.

The present study can contribute new empirical facts about the role of morphological structure in phonetic realization by investigating the acoustic realization of the two homophonous prefixes locative and negative *in-*. According to standard models of speech production, they should behave similarly with regard to their phonetic implementation. There should be no systematic difference in duration between the two affixes, and they should display the same gemination behavior. Systematic differences between the two prefixes would provide another piece of evidence for the presence of morphological structure in phonetic detail, and would thus support the claim for a revision of standard speech production models. These revisions could, for example, pick up and further develop Cohen-Goldberg's heterogeneity of processing hypothesis. The results could also be used to further specify usage-based models of speech production.

4.5 Summary: Theoretical implications

In this chapter we have seen that the pattern of morphological gemination in English affixation has important implications for various theories of the morphophonological and morpho-phonetic interface. Even though the different approaches deviate from each other in important respects, one can state that all of them are based on the assumption that morphological boundary strength influences the phonetic implementation of complex words. Derivatives with stronger boundaries, i.e. more decomposable derivatives, are less likely to be reduced, i.e. are likely to geminate. Derivatives with weaker boundaries, i.e. less decomposable derivatives, are more likely to be reduced, i.e. are more likely to degeminate.

The conceptualization of boundary strength deviates, however, vastly between the approaches. While some assume a categorical difference between affixes (e.g. Lexical Phonology, Stratal OT), others assume boundary strength to be a gradient, probabilistic word-specific concept (e.g. the Decomposability Approach, the Morphological Informativeness Approach). While some approaches define boundary strength by means of mainly lexical factors, such as the type of base an

affix takes or an affix's productivity (e.g. Lexical Phonology, Stratal OT), others mainly focus on prosodic aspects (e.g. Prosodic Phonology) and others concentrate on semantics (e.g. Morphological Informativeness). The differences in the conceptualization of boundary strength mirror general differences in theoretical assumptions about the morpho-phonological interface. As described in detail in the previous sections, these differences lead to different predictions about gemination with the affixes under investigation in this study. Testing which approach makes the most accurate predictions can therefore provide an important theoretical contribution. By conducting empirical studies, and thus finding out which factors govern gemination in English affixation, the predictions made by some of the approaches will be falsified, while others will be supported. Note that some outcomes might simultaneously support two approaches. For example, the degemination of the suffix *-ly* is predicted by the prosodic word approach, as well as by the morphological informativeness approach. And the gemination of *un-* is predicted by almost all of the approaches.

Table 4.3 summarizes the approaches discussed. The first column names the approach, the second the theoretical concept assumed to govern gemination and the third the main factor(s) assumed to influence gemination. To test the predictions of each approach, it is crucial to empirically investigate the factors in the third column, i.e. to test their effect on morphological gemination. For the two psycholinguistic approaches discussed, i.e. Morphological Segmentability and Morphological Informativeness, two different predictions were formed: a word-specific one and an affix-specific one. Both are summarized in the table. Note that the affix-specific predictions will be tested by consulting the lexical segmentability hierarchies formed in §3.2. These hierarchies are based on qualitative analyses of the affixes' features and display the affixes' segmentability, as well as their degree of informativeness in terms of their semantics. In the course of this book these theoretically formed hierarchies will be verified by empirical data. For convenience, the pertinent hierarchies are repeated in Figure 4.6.

In addition to the factors influencing gemination and degemination, the nature of gemination will be investigated in this study. While the formal linguistic approaches assume gemination to be categorical, psycholinguistic approaches do not make assumptions about the nature of the phenomenon, i.e. according to psycholinguistic approaches degemination might be gradient. As laid out in detail in §4.3.1, I will test both possibilities, i.e. I will test the prediction that gemination is categorical and the prediction that degemination is gradient. This will be done by investigating the distribution of duration across doubles and singletons, and by investigating which type of factors govern gemination.

Draft of May 27, 2019, 18:04

Table 4.3: Summary of concepts and factors predicting gemination according to different theoretical approaches

Approach	Concept	Factor(s)
Lexical Phonology	stratum of affix	affix
Stratal OT	stratum of affix type of base for dual-level affixes	affix type of base
Prosodic Word	prosodic word status of affix	affix semantic transparency type of base
Morphological Segment-ability (word-specific)	decomposability of derivative	relative frequency semantic transparency type of base decomposability rating semantic similarity
Morphological Segment-ability (affix-specific)	segmentability of affix	affix
Morphological Informative-ness (word-specific)	word-specific informativeness	semantic transparency
Morphological Informative-ness (affix-specific)	affix-specific informativeness	affix

	Segmentability hierarchy	Additional assumption
Semantic Hierarchy	*un-* > {*dis-*, *in-*$_{\text{NEG}}$}> *in-*$_{\text{LOC}}$ > *-ly*	lexical meaning over productivity, transparency and type of base
Non-Semantic Hierarchy	*un-* > *-ly* > {*dis-*, *in-*$_{\text{NEG}}$}> *in-*$_{\text{LOC}}$	productivity, transparency and type of base over lexical meaning

Figure 4.6: Lexical segmentability hierarchies of affixes

In the last part of the chapter, I discussed implications for speech production models. It was shown that the acoustic realization of the homophonous prefixes negative and locative *in-* has important implications for the modeling the acoustic realization of complex words. Currently models of speech production (cf. Dell 1986; Johnson 1997; Levelt et al. 1999; Bybee 2002; Pierrehumbert 2001; 2002) are unspecified with regard to the processing of complex words. Finding differences in the acoustic realization of the homophonous *in*-prefixes would contribute further evidence for the presence of morphological structure in phonetic detail. Models of speech production would need to be revised, or specified, with regard to this aspect.

5 General method

The predictions developed in the previous chapter were investigated by conducting two studies: a corpus study and an experimental study. In each study the five affixes *un-*, locative *in-*, negative *in-*, *dis-* and *-ly* were investigated. Since both studies were conducted to answer the same research questions, they were designed to be comparable.

In both studies, multiple regression analysis was used to investigate which factors influence consonant duration with the five affixes under investigation. Multiple regression has the advantage of enabling us to look at the effect of one predictor in the presence of other, potentially intervening, predictors. To find out whether a word geminates, the influence of the number of consonants at the morphological boundary was investigated. In case of gemination, the double consonant (e.g. /nn/ in *unnatural*) is longer than the correspondent singleton (e.g. /n/ in *uneven*). In case of degemination, the double consonant is as long as the singleton. In addition to the number of consonants, I also tested the effects of other determinants of consonant duration and gemination. On the one hand, I tested the influence of phonetic, phonological and lexical factors which are assumed to affect duration, such as, for example, speech rate and stress. On the other hand, the investigated factors were chosen based on the predictions made in the last chapter. In other words, I tested the influence of factors which are predicted to affect consonant duration according to the different theories of the morpho-phonological interface discussed (cf. Table 4.3 in Chapter 4).

In the first part of this chapter, I will discuss general differences between corpus and experimental studies. This is important in order to understand why both types of studies were conducted. I will then describe the general methodology followed in both studies. After describing the composition of the data sets, I will explain the segmentation procedure of the sound files. Then, I will explain the main statistical models used in both studies. Finally, I will describe the variables included in the models. Even though large parts of the methodology are the same in both studies, there are several aspects in which the studies' methodologies differ from each other. The specific methodology for each study will be described

in chapters 6 and 7.[1]

5.1 Corpus studies vs. experimental studies on speech production

Corpus studies have the advantage of looking at natural conversational speech. As discussed by Tucker & Ernestus (2016), it is of high importance to investigate this kind of speech in order to theoretically model speech production. By reviewing previous studies Tucker & Ernestus (2016) demonstrate that conducting studies on carefully articulated speech does not suffice to shed light on speech production, and that there are profound differences between careful and conversational speech. The two types of speech deviate in word choice, sentence structure, tone and intonation, phonological assimilation and, crucially for this study, degree of speech reduction. Since degemination is realized by durational reduction of the double consonant, one can expect differences in gemination depending on the type of speech investigated. One would expect more reduction in natural conversational speech than in experimental careful speech. That the type of speech investigated is important for gemination is shown by the results in Oh & Redford (2012) where more reduction, i.e. a higher degree of degemination, was found in normal than in careful speech (see §2.4.2 for discussion). One might therefore conclude that natural conversational speech, as found in a corpus, is better suited to investigate gemination in English than experimental careful speech.

However, there are a number of drawbacks with investigating natural conversational speech. This type of speech is very difficult to elicit in controlled experiments, and one therefore has to resort to corpora. As laid out by Tucker & Ernestus (2016: 21), "corpora come at a high cost in terms of their creation", i.e. annotation and segmentation of the data is very time consuming. Furthermore, by their very nature corpus data entail various factors which are not controlled for, and which might confound results (such as contextual and pragmatic aspects of language). While a lot of these factors can be accounted for by using advanced statistical methods, corpus data will always be less controlled than experimental data, i.e. the potentially negative effect of confounding factors is always higher in corpus studies than in experimental studies (see also Kunter 2017: 144 for discussion). An additional potential problem with corpus data is the number of types and tokens available in a given corpus. While for some phenomena a corpus

---

[1]An earlier version of sections 5.3, 5.4 and 5.5 has been published in Ben Hedia & Plag (2017).

Draft of May 27, 2019, 18:04

might comprise a higher number of types and tokens than attainable in an experimental study, for other phenomena corpora may not feature higher numbers of types and tokens. This is the case with morphological gemination. As discussed in §3.3, morphological geminates are not frequent among prefixes. They do not occur very often in natural speech, and the corpus study in this work is restricted to a rather small set of types and tokens for some affixes.

The small size of the data set entails an additional problem which is related to the composition of the data set. Since the choice of adequate items for the corpus study is very limited, factors of interest are unevenly distributed among types in the study. Even though multiple regression models open up the opportunity to work with unevenly distributed data sets, some distributional aspects, such as the systematic co-occurrence of two or more factors, or the infrequency of types with specific attributes, cause problems which cannot be ignored. They have to be taken into account when conducting statistical models, and when interpreting the results. Some potential effects on gemination might not be testable in the corpus study. Therefore, an experimental study needs to be conducted to complement the results of the corpus study.

An experimental study has the advantage of offering the possibility to include a great variety of types and a high number of tokens. Furthermore, factors which possibly influence the duration of the boundary-adjacent consonant(s) can be controlled for in a carefully designed experiment. By choosing specific carrier sentences, as well as by a careful item selection, it is possible to control and manipulate factors, such as accentuation and word form frequency. However, experimental data are not as natural as corpus data and might therefore be argued to not represent natural speech processing.

As one can see, the drawbacks of one type of study form the advantages of the other. While the corpus data provides natural conversational speech but only a small amount of viable types and tokens, the experimental data is less natural but enables one to more systematically investigate the factors of interest. Following the approach taken by Arppe & Järvikivi (2007) and Kunter (2017), one can therefore state that the results of the corpus study and the ones from the experimental study complement each other to form a complete picture of the phenomenon under investigation, in this case the complete picture of the durational pattern of morphological geminates for the five affixes *un-*, locative *in-*, negative *in-*, *dis-* and *-ly*. By conducting a corpus study and an experimental study, I will furthermore be able to compare the results from spontaneous speech with the ones from experimental speech. This will give me the opportunity to detect differences in gemination depending on speech condition and affix. In other words, for each

affix I will be able to find out which factors influence gemination at different levels of speech processing. As stated by Arppe & Järvikivi (2007: 1), "each method adds to our understanding of the studied phenomenon, in a way which could not be achieved through any single method by itself."

5.2 Composition of the data sets

To investigate gemination different structures were included in the data sets. The term STRUCTURE refers to sets of words with a particular orthographic, phonological and morphological make-up. In this section, I will introduce the different structures included in the studies, and explain their relevance for the investigation. First, I will describe the structures investigated in the corpus study. Then, I will describe the structures investigated in the experimental study. Finally, I will give an overview of the composition of the data sets in both studies. The overview contains a comparison of the structures investigated in the corpus and the experimental study, as well as a summary of the type and token frequencies for both studies. A detailed description of how the items were selected in each study will be given in Chapters 6 and 7.

5.2.1 Corpus study

The tokens investigated in the corpus study were extracted from the Switchboard Corpus (Godfrey & Holliman 1997), which consists of about 2400 two-sided phone conversations among North American speakers of English. The data set was compiled of complex words featuring the five affixes *un-*, locative *in-*, negative *in-*, *dis-* and *-ly*. As already discussed in §3.1.2, a word counted as morphologically complex if it showed the affixational meaning and if its base is attested outside the derivative with a similar meaning. It did not matter whether the base occurs as a free morpheme (e.g. *natural* in *unnatural*) or as a bound morpheme (e.g. *-plicit* in *implicit* and *explicit*).

To investigate whether the five affixes geminate, it was essential to include two different structures. On the one hand, I included complex words which feature a phonological double consonant at the morphological boundary (e.g. *unnatural*). On the other, I included complex words which feature a phonological singleton at the morphological boundary (e.g. *uneven*). As discussed in §3.3, for the allomorph /ɪn/ morphological geminates are extremely rare. This is mirrored in the number of tokens of this category found in the Switchboard Corpus. It turned out that the corpus only contained 17 /ɪn/-prefixed tokens with a double

/n/. There are only five different types with these tokens (*innocuous, innovated, innovation, innovative, innovativeness*). Furthermore, out of these five types four share the same root. Because of the low frequency of double nasals with the allomorph /n/, I decided to focus on the allomorph /ɪm/ instead, for which enough types exist. As discussed in §2.4, the literature very often is not explicit about the degemination behavior of the different allomorphs of *in-*. If something is said, the authors state that all allomorphs behave in the same way, i.e. all allomorphs of *in-* are taken to undergo degemination (e.g. Borowsky 1986; Cruttenden & Gimson 2014). There is thus no obvious reason not to investigate the allomorph /ɪm/ as a representative of the morpheme *in-*. Investigating the allomorph /ɪm/ also has the advantage of giving us the possibility to directly link the results to the two previous studies on gemination which also analyzed /ɪm/ instead of /n/.

Table 5.1 gives an overview of the structures investigated in the corpus study. The columns show the two structures which are essential to investigate gemination, i.e. phonological doubles in complex words and singletons in complex words. Examples for the different affixes are given in each line.[2] For singletons in complex words two different phonological environments exist, either the consonant-adjacent segment is a vowel or the consonant-adjacent segment is a consonant. Phonological doubles in complex words are always followed by a vowel.

Note that, throughout this book, I will use the term ENVIRONMENT to refer to the particular combinations of phonological and morphological structure as found in the included types for each affix. In other words, the term *environment* refers to affix-specific combinations of sounds in a particular morphological structure. The notation for the pertinent environments can be seen below each example. It is composed of the underlying consonant(s) in question and the neighboring segment ('V' for vowel, 'C' for consonant). The '#' marks a morphological boundary. These notations will be used in both studies throughout the book.

For the prefix *in-*, the table shows an empty slot for singletons in complex words with a following vowel. The reason is that there is no attested sequence /ɪmV/, i.e. it is impossible to investigate /ɪm/-prefixed words with a following vowel. This is due to the allomorphy of the prefix. The prefix *in-* only takes the form /ɪm/ when it is followed by homorganic consonants, i.e. by the bilabials /m/, /b/ or /p/. Thus, for /ɪm/-prefixed words containing a single nasal only the sequences /ɪmb/ and /ɪmp/ exist.

---

[2]Note that both *in*-prefixes, i.e. locative and negative *in-*, were investigated in one data set. Therefore, they are not displayed separately at this point. The same holds for Table 5.2 and Table 5.3.

Table 5.1: Overview of the investigated structures in the corpus study

	Phonological double in complex word		Singleton in complex word	
{un-}	*unnatural* (n#nV)	*uneven* (n#V)	*untold* (n#C)
{in-}	*immortal* (m#mV)	NA	*impossible* (m#C)
{dis-}	*dissatisfy* (s#sV)	*disarm* (s#V)	*disgrace* (s#C)
{-ly}	*really* (Vl#l)	*truly* (V#l)	*probably* (C#l)

5.2.2 Experimental study

The tokens investigated in the experimental study were collected in two experiments conducted at the Cambridge University Phonetics Laboratory in October 2015 and October 2016. As in the corpus study, complex words featuring the five affixes *un-*, locative *in-*, negative *in-*, *dis-* and *-ly* were included in the experimental study. One part of these words featured a phonological singleton at the morphological boundary, and one part featured a phonological double at the morphological boundary. In addition to phonological doubles in complex words and singletons in complex words, two other structures were investigated in the experiment, orthographic doubles in simplex words and singletons in base words. The reason for not including these two structures in the corpus study is that not enough words with the pertinent structures were found in the Switchboard corpus. A statistical analysis of the tokens found in the corpus study was not reasonable.

Table 5.2 displays the four structures investigated in the experimental study. As in Table 5.1, the columns of the table display the investigated structures, and examples of the structures are given in each line for each affix. The notation for the pertinent environment is given below each example.

Table 5.2: Overview of the investigated structures in the experimental study

	Phonological double in complex word	Orthographic double in simplex word	Singleton in complex words		Singleton in base
{un-}	*unnatural* (n#nV)	NA	*uneven* (n#V)	*untold* (n#C)	*natural* (#n)
{in-}	*immortal* (m#mV) *innumerous* (n#nV)	NA	NA *inefficient* (n#V)	*impossible* (m#C) *intolerant* (n#C)	*mortal* (#mV) *numerous* (#nV)
{dis-}	*dissatisfy* (s#sV)	*dissertation* (sV)	*disarm* (s#V)	NA	*satisfy* (#sV)
{-ly}	*really* (Vl#l)	*belly* (Vl)	*truly* (V#l)	NA	*real* (Vl#)

I included orthographic doubles in simplex words, i.e. simplex words which feature a similar phonological make-up as the investigated affixed words, and which are spelled with an orthographic double, to investigate possible effects of orthography on gemination. That there is a relation between the orthography of a word and its phonology and phonetics is well established in the literature (see, for example, Smith & Baker 1976; Ehri 1993; Warner et al. 2004; 2006; Brewer 2008; Berg 2016). Brewer (2008), for instance, found that the duration of a segment is influenced by the number of graphemes it is represented by. The more graphemes are present, the longer the duration of the segment becomes. Similarly, Warner et al. (2004) and Warner et al. (2006) found an effect of orthography on duration in Dutch. In their data, phonological singletons which are represented by orthographic doubles (e.g. /t/ in *baatten*) are longer than phonological singletons which are represented by orthographic singletons (e.g. /t/ in *baten*). Warner et al.'s results suggest that gemination might not be a morpho-phonological phenomenon but an orthographic one. In other words, there is the possibility that the lengthening of double consonants is merely an orthographic effect and not a morpho-phonological effect. To test this possibility orthographic doubles in simplex words were included. In contrast to phonological doubles in complex words, they only feature one underlying consonant (e.g. *di*/s/ss/*atisfy* vs. *di*/s/*ertation*). If orthographic doubles in simplex words are longer than single-

tons in complex words, the number of underlying consonants is irrelevant for gemination. In that case, gemination is an orthographic phenomenon. If, on the other hand, orthographic doubles in simplex words are as long as singletons, and if they are simultaneously shorter than phonological doubles in complex words, gemination is amorpho-phonological phenomenon. In this case, only words with two underlying consonants geminate.

Simplex words with orthographic doubles that feature the same phonemic strings as the investigated affixes are extremely rare. The number of simplex words with an orthographic double starting in /ʌn/, /ɪn/ or /ɪm/ is so small that a statistical investigation is not reasonable. Only for words with the phonemic strings /dɪs/ and /li/ a sufficient number of types exists. Therefore, the effect of orthographic doubles was only investigated for /dɪs/ and /li/.

The fourth structure investigated in the experimental study are phonological singletons in base words. I included the base words of all complex words with an underlying double consonant, e.g. *natural* for *unnatural* and *real* for *really*. For prefixes, the base words were included to compare the duration of the double consonant in the complex word with the base-initial consonant in the base word. For suffixed words, the double consonant was compared to the base-final consonant, respectively. This comparison was conducted to ensure that the potential lengthening of the morphological geminate was not caused by an inherently long base-initial consonant (or base-final consonant in case of *-ly*, respectively).

As can be seen in Table 5.2, in addition to the allomorph /ɪm/ the experimental study also investigated /ɪn/ to test gemination with *in-*. This is different from the corpus study in which, due to a low token frequency of morphological geminates with /ɪn/, only /ɪm/ was investigated.

For the affixes *dis-* and *-ly*, the phonological environment was kept constant across all structures. In all words, the consonant(s) of interest is adjacent to a vowel. The phonological environment was kept constant to avoid effects of the affix-adjacent segment on the duration of the consonant(s) of interest. Due to the limited choice of types and tokens in the corpus study, controlling the phonological environment across words was only possible in the experimental study.

It would have also been possible to keep the following segment constant across all words for the prefix *un-* and the allomorph /ɪn/. However, for the allomorph /ɪm/ the phonological environment cannot be kept constant across structures. In case of a singleton at the morphological boundary, the allomorph is always followed by a stop consonant, i.e. /p/.[3] Doubles are always followed by a vowel.

---

[3]Because of low type frequency of items with a following /b/, and to keep the environment as constant as possible, no *in*-prefixed items with a following /b/ were included in the study.

Draft of May 27, 2019, 18:04

This evokes the problem of the interpretation of potential durational differences between doubles and singletons for /ɪm/, i.e. it is unclear whether potential differences are caused by the number of consonants or the deviating phonological environment. Teasing apart the two effects is impossible for /ɪm/.

However, teasing apart the effect of the number of consonants from the effect of the following segment is possible for *un-* and /ɪn/. As shown in Table 5.2, for *un-* and /ɪn/ the singleton environment with a following vowel (n#V) exists and was included in the study. I also included the environment n#C for /ɪn/ and *un-*-prefixed words in the data set. One can thus analyze durational differences in *un-* and /ɪn/-prefixed words across the three environments n#nV, n#V and n#C. This makes it possible to see which durational differences can be expected between nasals with a deviating number of underlying consonants (durational difference between n#nV and n#V), and which durational differences can be expected between nasals with deviating following segments (durational difference between n#V and n#C). The differences in duration between the environments for *un-* and /ɪn/ can then be used as a reference for the durational pattern of /ɪm/. In other words, one can compare the durational differences found for /ɪm/ with the ones found for *un-* and /ɪn/.[4] In turn, one can find out whether the potential difference between singletons and doubles for /ɪm/ should be interpreted as gemination, or whether it is merely caused by differences in their phonological environments.

5.2.3 Overview of the data sets

Table 5.3 gives an overview of the composition of the data sets in the two studies. On the left side of the table the environments investigated in the corpus study are displayed, on the right the ones of the experimental study. For each affix (and in case of *in-* its allomorphs), the included environments are indicated in the pertinent fields by the annotation introduced in Table 5.1. For example, the corpus study includes the following three environments for *un-*-prefixed words: phonological double nasals in complex words followed by a vowel (n#nV), singleton nasals in complex words followed by a vowel (n#V) and singleton nasals in complex words followed by a consonant (n#C). In the middle column of the table examples for each environment are given.

---

[4]Note that in all *un-* and *in-*-prefixed words with an n#C-environment, the following consonant is /t/ (e.g. *untold, intolerant*). The consonant /t/ was chosen because of its comparability to /p/, i.e. the consonant following /ɪm/-prefixed words with a singleton. In both cases a stop which shares its place of articulation with the preceding nasal follows the prefix.

For the analysis, I generated subsets for each affix (and in case of *in-* its allomorphs). This was necessary because the duration of a consonant heavily depends on its type. For example, fricatives are generally longer than nasals, and nasals are generally longer than laterals (see, for example, Umeda 1977). For the affixes investigated in this study this means that out of the affixational consonants investigated the fricative in *dis-*prefixed words will most likely be the longest, followed by the nasals in *un-* and *in-*prefixed words. The /l/ in *-ly*-suffixed words is expected to be the shortest. Even within types of consonants there are major differences in duration. Umeda (1977), for instance, found that bilabial nasals in word-medial position are almost twice as long as alveolar nasals in the same position (74 ms vs. 38 ms). Affixes which differ in their phonological make-up must therefore be investigated separately with regard to their duration, i.e. it is necessary to analyze the duration of the different consonants in different analyses.

Table 5.4 gives an overview of the type and token distribution in each study. For each subset, the number of investigated types and tokens is given in the table.[5] The compilation of the data sets will be discussed in detail in chapters 6 and 7.

5.3 Acoustic analyses

After all sound files were extracted from the corpus and the experimental recordings, the data was segmented and phonetically transcribed using the software Praat (Boersma & Weenink 2014). For each token, the segments of the affix in question, as well as the segments of the syllable immediately following or preceding the affix under investigation, were annotated. Figure 5.1 displays an example. As can be seen in the figure, doubles were segmented as one segment. This is because in almost all cases no boundary between the two adjacent consonants was distinguishable. Only for some words in the experimental data a pause was uttered between the affix and the base, i.e. the two consonants were pronounced as two independent segments. These tokens were noted down and will be discussed in Chapter 7.

There are two possible ways of segmenting speech data, manual segmentation and automatic segmentation. Automatic segmentation has the major advantage of demanding a lower workload and less time than manual segmentation. Furthermore, one can expect automatic segmentation to be very systematic. This is

---

[5]Note that locative and negative *in-* have the same phonemic form, and that they can therefore be analyzed in one data set.

Table 5.3: Overview of subsets and investigated environments in the two studies

	Corpus Study	Example	Experimental Study
	n#nV	*unnatural*	n#nV
	n#V	*uneven*	n#V
un-	n#C	*untold*	n#C
		natural	#nV
		innumerous	n#nV
		inefficient	n#V
in-		*intolerant*	n#C
		numerous	#nV
	m#mV	*immortal*	m#mV
im-	m#C	*impossible*	m#C
		mortal	#mV
	s#sV	*dissatisfy*	s#sV
	s#V	*disarm*	s#V
dis-	s#C	*disgrace*	
		satisfy	#sV
		dissertation	ssV
	Vl#l	*really*	Vl#l
	C#l	*probably*	
-ly	V#l	*truly*	V#l
		real	Vl#
		belly	Vll

Table 5.4: Type and token distribution across subsets in the two studies

	Corpus Study			Experimental Study	
	Types	Tokens		Types	Tokens
un-	101	158		89	2615
in-				83	1232
im-	83	156		64	1635
dis-	64	128		59	1114
-ly	150	154		103	1645
Total	398	596		398	8241
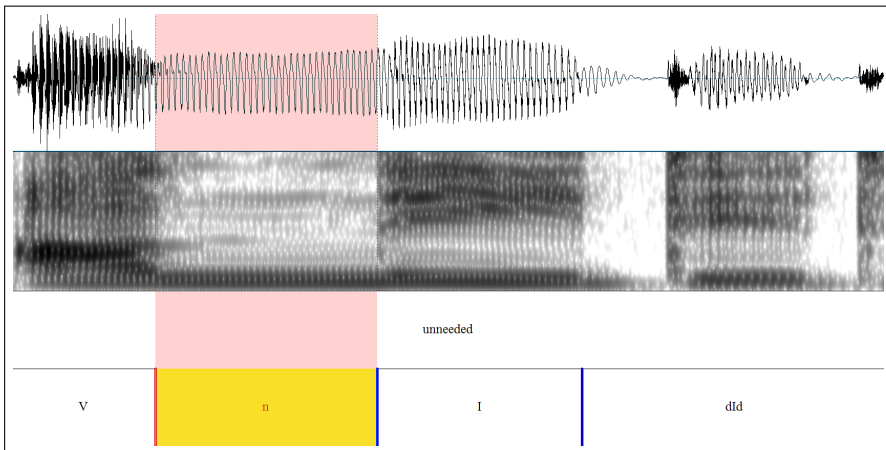
Figure 5.1: Segmentation example *unneeded*

because automatic segmentation relies on forced aligners which are based on algorithms. These algorithms ensure that every file is segmented according to the same criteria. This means, in contrast to manual segmentation, there is no risk of inter-annotator differences.

However, there are also major disadvantages with automatic segmentation. The forced aligners used in automatic segmentation rely on canonical pronunciations. This means that the forced aligner will always annotate all phonemes represented in the canonical pronunciation of a word, irrespective of whether they were produced by the speaker or not. Especially in conversational speech words are sometimes drastically reduced, i.e not all sounds of a word are produced. This

poses a serious problem for automatic segmentation. Segments which are not present are annotated by the system. An additional problem is related to the fact that forced aligners do not analyze the whole sound file but instead analyze the file in increments of a few milliseconds. The forced aligner software WebMAUS (Schiel 1999; Kisler et al. 2016), for instance, uses 10 ms increments. Especially when investigating fine phonetic detail these increments are problematic. For example, the duration of a word-medial /l/ is 40 ms on average (cf. Umeda 1977). Increments of 10 ms might distort the phonetic investigation of the duration of /l/ immensely by positioning a boundary 10 ms too early, or 10 ms too late, i.e. the automatic segmentation might show /l/ to be up to 50% shorter or longer than it actually is.

To test the accuracy of automatic segmentation, i.e. to test whether automatic segmentation can be used in this study, I carried out an automatic segmentation on a subset of the corpus and the experimental data. I used the software WebMAUS (Schiel 1999; Kisler et al. 2016) for the segmentation. WebMAUS takes speech files with orthographic transcriptions as an input and gives segmented, phonologically transcribed text grids as an output. The segmentation is based on a production system, which takes the canonical pronunciations of an utterance, as well as its sound wave, and then, on the basis of a Viterbi alignment procedure, computes the most probable pronunciation variant. Based on this pronunciation variant the speech wave is segmented in 10 ms increments.

The automatic segmentation of the subset showed that the automatic segmentation is inaccurate and therefore not suited to investigate fine phonetic detail. As expected, the system annotated sounds which were not present in the acoustic signal and misplaced boundaries, i.e. boundaries were set too early or too late in the speech signal. Especially for the corpus data the boundaries were very poorly placed. This might be due to the extreme reduction found in conversational speech. Having a valid, reliable and precise segmentation is extremely important for this investigation. Therefore, it was decided to not use automatic segmentation but instead segment all sound files manually. I decided to not first use automatic segmentation and then adapt the set boundaries since revising the automatic segmentation holds the risk of influencing the annotator unconsciously.

In contrast to automatic segmentation, manual segmentation is less prone to systematic mistakes and inaccuracies. This is because it does not rely on canonical representations and incremental analyses. It thus allows for more precise annotations than automatic segmentation. However, there are also disadvantages with manual segmentation. First, it is very time-consuming. Second, it is prone to inconsistent, unsystematic boundary setting, as well as to inter-annotator differ-

ences. While the possibilities to speed up manual segmentation are very limited, there are various possibilities to prevent inconsistencies and inter-annotator differences. To ensure the reliability and validity of the manual segmentation, I applied the following four strategies: 1. the development of strict segmentation criteria based on the specifics of each sound, 2. intensive training of the annotators, 3. segmenting a proportion of the data twice, and 4. testing the influence of the annotator on the segmentation statistically. In the following, I will discuss each strategy in detail.

1. The development of strict segmentation criteria

The criteria on which the segmentation was based were developed by consulting the relevant phonetic literature (cf. Ladefoged & Maddieson 1996; Johnson 1997; Ladefoged 2003; Machač & Skarnitzl 2009; Ladefoged & Johnson 2011) and were optimized during the segmentation process. The final criteria will be described in the following. First, I will describe the criteria for the segmentation of the nasals in *un*- and *in*-words. Then, I will give a description of the criteria for fricatives in *dis*-words. Finally, the criteria for laterals in *-ly*-words will be given.

The nasals in *un*- and *in*-prefixed words

Nasals have a regular waveform which has a lower amplitude than the waveform of vowels. Formants of nasals are quite low and faint in comparison to those of vowels. Boundaries between the preceding vowel and the nasal were thus set where the acoustic energy drops in the waveform, the spectrogram becomes visibly fainter and the higher formants visibly decrease (see Figure 5.1). In case of a following vowel, the boundary was marked at the point where the amplitude increases in the waveform and the formants become clearly visible (see Figure 5.1). Since approximants have, similar to vowels, a higher amplitude than nasals, as well as more acoustic energy, the identification of approximants following the nasal was similar to the identification of a following vowel. If a stop followed the nasal, the boundary was marked at the beginning of the occlusion, which was identified by the abrupt decrease of the waveform and the sudden diminishment of the formants (see Figure 5.2). In case of a following fricative, the boundary was set where the waveform became visibly irregular and the energy was concentrated in the upper part of the spectrogram with no distinct formants visible. All boundaries were set at the nearest zero crossing of the waveform.
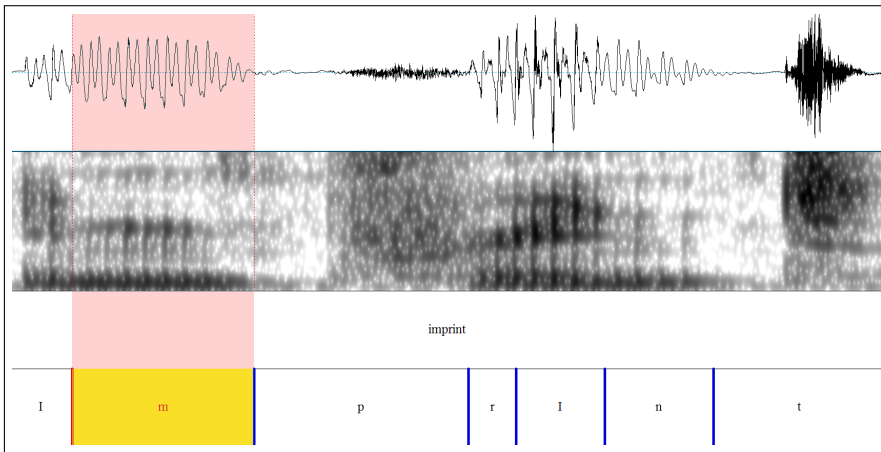
Figure 5.2: Segmentation example *imprint*

The fricative in *dis*-prefixed words

Fricatives are characterized by an irregular waveform, which is very easy to distinguish from the regular waveform of vowels. Furthermore, for fricatives, there is energy throughout the whole spectrogram and no separate formant bands are visible. Most energy is visible in the upper part of the spectrogram (see Figure 5.3). This is even more pronounced for voiceless fricatives, which are found in the majority of the *dis*-prefixed words. The boundary between the preceding vowel and the fricative was set where the waveform became irregular and the distinct formant structure vanished. The boundary between /s/ and the following vowel was set where the opposite was the case (see Figure 5.3). In case of a following approximant, the same criteria were applied. If a stop followed the fricative, the boundary was marked at the beginning of the occlusion. There were no fricatives immediately following the prefixal /s/ in the data sets.
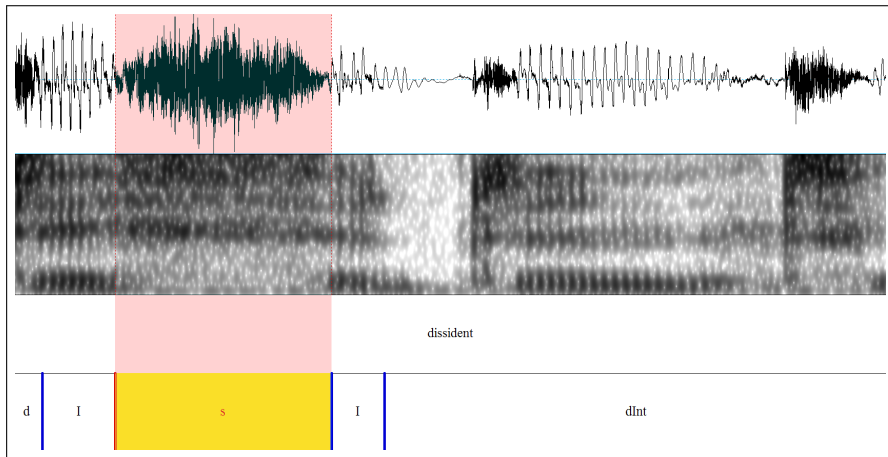
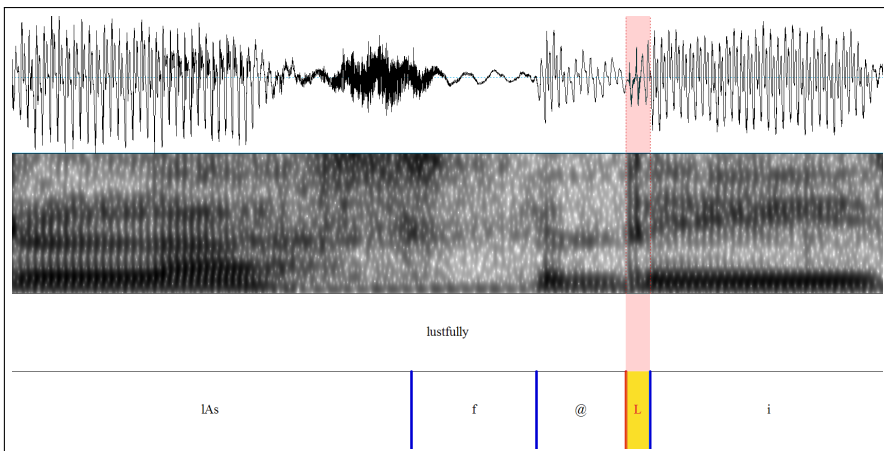Figure 5.3: Segmentation example *dissident*

The lateral in *-ly*-suffixed words

Out of the four consonants investigated in this study, laterals are the most difficult to segment. This is due to the fact that laterals are very similar to vowels regarding their acoustical properties. Thus, it is quite challenging to set a boundary between vowels and laterals (see also Machač & Skarnitzl 2009: chapter 7 for discussion). However, there are some aspects in which /l/ can be distinguished from vowels. There is less amplitude in the waveforms of laterals than in the one of vowels. Furthermore, their formant structure is, in contrast to the one of vowels, constant and, due to less energy in the speech signal, the formants of /l/ are in general fainter than the ones of vowels. This is especially the case for higher formants. Since in the suffix *-ly* the lateral is followed by a high vowel, which displays a high level of energy in the upper formants, it is possible to see the formant structure change between /l/ and the following vowel in my data sets. The boundary between /l/ and /i/ was thus set at the point at which the formant structure changed, i.e. the higher formants became more pronounced (see Figure 5.4). For intervocalic /l/ a visible decrease in the waveform, as well as the change in formant structure was used to mark the beginning of /l/ (see Figure 5.4).

Setting the boundaries between /l/ and a preceding consonant was generally not problematic since approximants can be distinguished quite easily from nasals, stops and fricatives. Approximants generally have a higher amplitude and more energy in the spectrogram than nasals. Their waveform is periodic, whereas the waveform of fricatives, as well as the waveform of the aspirational phase of stops,

Draft of May 27, 2019, 18:04
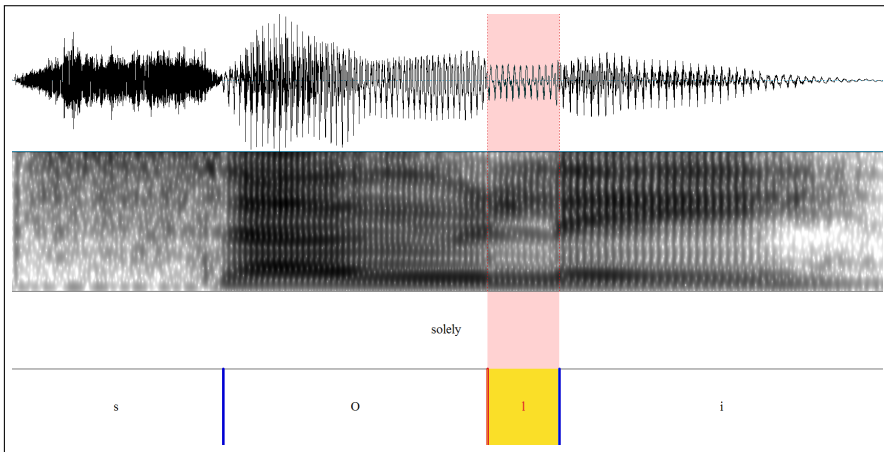
Figure 5.4: Segmentation example *solely*

is irregular.

Figure 5.5: Segmentation example *lustfully*

In some cases the waveform and the spectrogram for /l/ showed a completely different pattern. Figure 5.5 shows a case in which /l/ is marked by a dark, vertical bar which stretches throughout the whole spectrogram. In this case /l/ is realized as a tap, i.e. not as an approximant. The bar marks the tongue release from the alveolar ridge. Because of the high amount of energy in the spectrogram /l/ can be easily set apart from the neighboring vowels. This type of /l/ , i.e. a tap /l/, is

generally shorter than the one described above, i.e. the approximant /l/. To account for this difference in duration the type of /l/ pronounced was coded in the variable TYPEOFL. The variable was then incorporated in the statistical models as a covariate.

The criteria described above allowed for a valid and reliable segmentation of /l/. However, there were still items which were very difficult to segment. This was especially the case for items of the category l#, i.e. /l/ at the end of base words (e.g. *hateful, cool*). An example is displayed in Figure 5.6. There is no visible boundary between the preceding vowel and the lateral, i.e. the speaker did not pronounce them as two distinct sounds. There are various possible explanations, such as that the speaker might have deleted the word-final /l/, or that he might have vocalized it while deleting the preceding vowel. Either way, it is impossible to set a valid boundary between /l/ and the preceding vowel. Therefore, we marked the whole interval as /l/ (see Figure 5.6) and coded the pertinent tokens as featuring a vocalized /l/. Hence, the data set comprised three different types of /l/: the approximant /l/, the tap /l/ and the vocalized /l/. These three different types were coded in the variable TYPEOFL with the values `approximant`, `tap` and `vocalized`.

2. Intensive training of the annotators

The corpus data was segmented by five annotators. The experimental data was segmented by six. The reliability of the segmentation criteria was verified by a
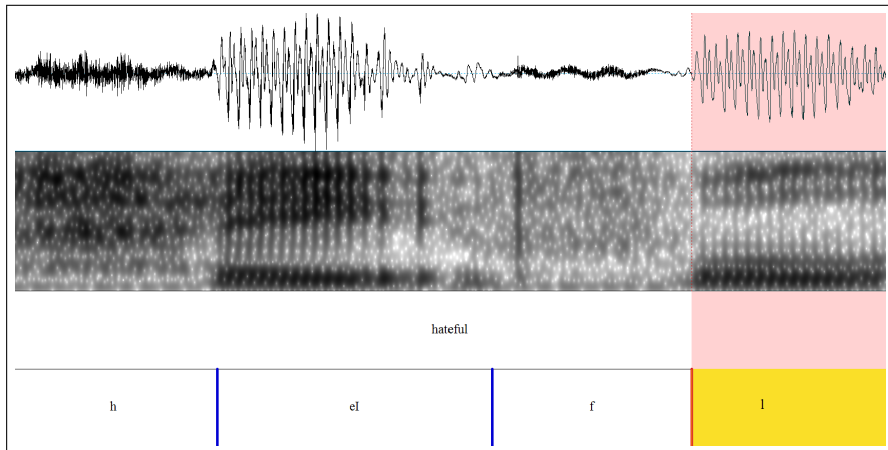

Figure 5.6: Segmentation example *hateful*

Draft of May 27, 2019, 18:04

set of trial segmentations. In these trials the annotators segmented the same 30 items. If boundaries differed by more than 10 milliseconds, the annotators discussed the discrepancy and refined the criteria in order to reduce inter-annotator variation. The trial segmentations were repeated two times until all boundaries were reliably placed with only small variations, i.e. variations within 10 ms. For the final measurement, each annotator worked on a disjunct set of items. If an annotator was not confident concerning the segmentation of an item, the item was discussed with all annotators. Items which could not be validly segmented were excluded. This was the case for several of the corpus items due to the poor quality of the sound files, as well as for various items including /l/.

3. Segmenting a proportion of the data twice

To further ensure reliability, 10 % of each annotator's items were segmented by a second annotator. These 10 % were then compared. Discrepancies between the segmentations were discussed and systematic mistakes were detected. To correct these mistakes, and to avoid them in future codings, the segmentation criteria were clarified and enhanced. The annotators revised all previous segmentations according to the enhanced criteria.

4. Testing the influence of the annotator on the segmentation statistically

After the segmentation process was completed, a script was used to measure and extract word duration, the duration of the consonants in question and the duration of the preceding and the following segments in milliseconds. For each subset, I compared the segmentations of each annotator by checking whether the durations varied significantly across annotators. For the comparison, I used linear models in which the dependent variable was the duration of the consonant in question (e.g. /n/ for *un*-prefixed words). I tested whether consonant duration varied significantly by annotator. The models revealed that the annotator did not have any effect on the duration of the consonant, i.e. the segmentation did not vary significantly across annotators.

5.4 Statistical analyses

In both studies similar statistical analyses were used to investigate morphological gemination. All statistical modeling was carried out using the software R (R Development Core Team 2014). The main analyses in all studies consisted of

the investigation of the duration of the consonant in question. In this section, I will give an introduction to the statistical models fitted to investigate duration. In addition to general information about the statistical procedures applied, I will discuss pertinent problems and explain how they were approached. The details of each model, as well as problems with specific data sets, will be described in the pertinent sections. Statistical procedures which were only relevant for specific subsets of the data will also be discussed in the pertinent sections.

This first durational analysis in both studies consisted of investigating the distributions of consonant duration across different environments. This investigation is of importance with regard to the question of whether gemination is a categorical or a gradient phenomenon (see §2.2 and §4.3.1 for discussion). If gemination is a categorical phenomenon, the data should show a bimodal distribution. If gemination is a gradient phenomenon, one would expect a gradient increase in duration from singletons to doubles.

To investigate the distribution of duration across environments, boxplots were generated, and differences in average duration between environments were tested for significance by using standard statistical tests. Boxplots have the advantage of displaying a lot of information about the data's distribution simultaneously (e.g. median, the distribution of quartiles). Therefore, these graphs are very well suited to compare the distribution of two or more categories with each other (see, for example, Benjamini 1988). I used boxplots to compare the durations of singletons with the durations of doubles. The plots show whether the distributions of singletons and doubles deviate, and whether the distribution of the whole data set (including singletons and doubles) is bimodal. If the distribution is bimodal, gemination can be assumed to be categorical. In case of a non-bimodal distribution more advanced statistics are needed to test whether the data at hand shows (gradient) gemination, or whether all morphological geminates in the data set degeminate. After the analysis of the raw data, more advanced statistical analyses were used to investigate gemination and the factors influencing duration more thoroughly.

In both studies multiple regression was used to test the effect of various factors on consonant duration. For each affix, i.e. each subset, at least one model was fitted (see Table 5.3 for an overview of the subsets). The dependent variable in all models was consonant duration, i.e. the models predicted the duration of the affixational consonant. Consonant duration was measured in absolute and in relative terms. Relative duration refers to the duration of the consonant relative to the duration of its preceding segment. Models with both absolute and relative duration were fitted for each affix. The independent variables were mainly deter-

mined by the predictions made in the previous chapter and varied slightly across models. They will be presented in further detail in the next section.

Multiple regression was used since it is an established and highly successful way to deal with the multitude of factors involved in predicting durational properties of morphemes (see, for example, Hay 2007; Hanique & Ernestus 2012; Smith et al. 2012; Plag et al. 2017). Using this type of model one can investigate one specific predictor while simultaneously accounting for other, potentially intervening, predictors. One can thus, for example, test whether the number of consonants influences the duration of a consonant, while simultaneously taking other factors, such as speech rate or stress pattern, into account. Another major advantage of multiple regression models is their capability to deal with unbalanced data sets. Especially for the corpus data this is of high importance.

However, multiple regression models also entail some statistical problems that need to be addressed. Two of them are especially relevant for the present analyses: collinearity and overfitting. Let us first discuss collinearity. A number of measurements I would like to include in the models are correlated, for example the number of segments in the word and word duration. A word which has more segments is most likely also longer in terms of duration. This can lead to serious problems in regression models (*multicollinearity*, for example, Baayen 2008: chapter 6). If one of the two variables has an influence on consonant duration, so will the other. This makes it difficult for the model to tease apart the individual explanatory power of each of the two variables.

There are several possible strategies to deal with collinearity. One strategy is to include only one of the correlating variables. This is a conservative and safe strategy, which may, however, decrease the power of the model. If collinearity only affects noise variables, i.e. variables which are known to affect the dependent variable but whose effect is not of primary interest for a study, another option is to keep the correlating variables in the model but not interpret their individual contribution to the model (cf. Wurm & Fisicaro 2014). A third strategy is to combine factors. The two variables NumberOfSegmentsInTheWord and WordDuration, for example, can be combined by calculating the variable SpeechRate, which is computed by dividing word duration by the number of segments. Another possibility of combining factors is by conducting a principal components analysis. In this type of analysis the dimensionality of the data is reduced by transforming the different variables into so-called principal components. The transformation results in linear combinations of the predictors, i.e. the principal components, that are uncorrelated with each other (see, for example, Baayen 2008: chapter 5.1; Venables & Ripley 2011: chapter 12). To address poten-

tial collinearity problems, I applied all of the strategies mentioned above.

The second potential problem with regression models is overfitting (see, for example, Draper & Smith 1998; Babyak 2004). The number of variables included in a model must be appropriate for the number of observations in the data set. If this is not the case, the model cannot be trusted. In other words, if too many terms are included in a model, the model will not be able to adequately approximate the effects of the included variables. Note that the notion *terms* not only refers to the number of predictor variables per se, but also to the number of variable levels and interactions included in the model. A common rule of thumb states that 10–15 observations per term are necessary to avoid overfitting (cf., for example, Draper & Smith 1998). Especially in the corpus data sets, which are of relatively small size, the number of variables included in the models had to be restricted, and the variables needed to be chosen carefully.

Let us now turn to the modeling strategy which was adopted in all models. Following established practices in the field (e.g. Baayen 2008), I first conducted an initial model incorporating all variables whose effect was to be tested. I then checked the residuals of the model, which need to be normally distributed. If visual inspection revealed that the residuals had a non-normal distribution, transformations and the exclusion of outliers were used to obtain the desired pattern. If a transformation of the dependent variable was necessary to alleviate problems of non-linearity, Box-Cox transformations were used to identify a suitable transformation parameter for a power transformation (see, for example, Box & Cox 1964; Venables & Ripley 2011).

After the residuals showed a satisfactory distribution, I checked for collinearity in the models by looking at the correlations between potentially correlated variables. In those cases where collinearity was a potential problem, I followed the strategies described above. In all models, I tested for relevant interactions. The strategy for testing interactions will be discussed in sections 6.3.1 and 7.3.1, after all variables included in the models are introduced.

The regression models were then simplified by stepwise excluding insignificant predictors. A predictor was considered significant if its p-value was lower than 0.05, and if the Akaike Information Criterion (AIC) of the model including the predictor was lower than when the predictor was not included. A lower AIC indicates that a model including the factor has a greater explanatory power than a model without the predictor variable. Linear models were generated using the `lme4 package` (Bates et al. 2014).

In addition to the stepwise exclusion of insignificant factors, i.e. finding the one model which explains the variation found in the data best, I also used multi-

model inferencing (see, for example, Barth & Kapatsinski 2014). Multi-model inferencing estimates the predictive value of each variable by looking at a multitude of possible models. Instead of just giving the significance of one variable in one specific model, it indicates the importance of a variable across a multitude of models. The importance of a variable is determined by the number of models in which the variable is significant and by the goodness of each model in which the variable is significant (measured in the AIC of the model). A variable which is significant in various models with a high AIC will have a high importance value. A variable which is significant in fewer models with a lower AIC will have a lower importance value. The multi-model inferencing was carried out using the `MuMin package` (Bartoń 2016), and was used as an additional clue to assess which variables influence affixational consonant duration.

In addition to the durational analyses, statistical analyses were applied to investigate decomposability. I investigated the relation of five different decomposability measures to find out whether these measures can be used as operationalizations of the same underlying concept. The investigated decomposability measures will be introduced in the next section. Furthermore, I compared the included affixes by means of the different decomposability measures to investigate their segmentability. This analysis was conducted to find out whether the segmentability hierarchies introduced in Chapter 3 are borne out by the data. The conducted decomposability analyses will be explained in detail in the pertinent sections of Chapters 6 and 7.

5.5 Coding of the variables

In both studies the data was annotated with regard to factors which potentially influence consonant duration and gemination. The annotation of the data resulted in the coding of various variables. These variables can be divided into two groups: variables of interest and noise variables. Variables of interest are those variables which are used to test the predictions made in Chapter 4. In other words, these variables serve to test the effects of the factors predicted to govern gemination according to the different theoretical approaches (see Table 4.3 for an overview of these factors). Noise variables, on the other hand, are those variables which are known to influence consonant duration but which are not directly linked to the predictions made by the theoretical approaches discussed.

There are seven variables of interest. The first one is ENVIRONMENT. This variable was coded to answer the question of whether an affix geminates. It holds information about the morphological and phonological environment of the in-

vestigated consonant(s) and is essential for all predictions.

Five variables of interest are closely related to the notion of boundary strength, and form possible operationalizations of decomposability. The five variables are SemanticTransparency, SemanticTransparencyRating, TypeOfBase, RelativeFrequency and LSAScore. As will be laid out below in further detail, they are relevant for the majority of the predictions made in Chapter 4.

The last variable of interest is Affix, which codes for the affix itself. On the one hand, this variable is needed to answer the question of which affix geminates. On the other, it is necessary for the comparison of affixes with regard to their segmentability. This comparison is of importance for the affix-specific predictions made by the decomposability and the morphological informativeness approach (see §4.5 for discussion).

For the most part, the same variables were included in all analyses. However, due to differences between corpus and experimental data, some variables were only used in one of the two studies. Furthermore, affix-specific features, such as the phonological make-up of an affix, called for the inclusion of additional variables in some of the models (e.g. Voicing for *dis-*). Inherent differences between affixes also led to some minor differences in the coding of some variables across subsets, e.g. in the variable Environment.

Below I will describe all variables which were initially considered in all models, including the ones only considered for the models of specific subsets. I will describe why the variable is of interest for the study and how it was coded. Differences in the coding between subsets will also be discussed. Furthermore, I will lay out problems with regard to the testing of some of the variables, and I will briefly discuss how these problems were dealt with. The specific modeling procedure with regard to the inclusion of certain variables will be discussed in the pertinent sections of this book.

First, I will describe the variables of interest. Then, I will turn to the noise variables, which can be categorized into three different types: phonetic factors, phonological factors and lexical factors. Finally, I will give an overview of which variables were included in which study.

5.5.1 Variables of interest

Environment. The variable Environment was coded to test whether a word geminates. It codes the phonological and the morphological environment of the consonant(s) investigated in a particular word. The variable is based on the four different structures included in the two studies, i.e. phonological doubles in complex words, orthographic doubles in simplex words, singletons in complex words and

singletons in bases. This means that every level coded in the variable represents one of the four structures. In §5.2, the different structures and their environments were already introduced.

Table 5.5 gives an overview of the levels of the variable Environment for the prefixes *un-*, *in-* (for both allomorphs) and *dis-*. For each level examples are given. For *un-* and /ɪn/ four environments exist. For the allomorph /ɪm/ three environments exist, and for *dis-* five environments were coded.[6]

Table 5.5: Levels of the variable Environment for *un-*, *in-* and *dis-*

| **un-** | |
Environment	Example
n#nV	*unnatural*
#nV	*natural*
n#C	*untold*
n#V	*uneven*

| **in-** | | **im-** | |
Environment	Example	Environment	Example
n#nV	*innumerous*	m#mV	*immortal*
#nV	*numerous*	#mV	*mortal*
n#C	*intolerant*	m#C	*impossible*
n#V	*inefficient*		

| **dis-** | |
Environment	Example
s#sV	*dissatisfied*
#sV	*satisfied*
s#C	*disgrace*
s#V	*disarm*
sV	*dissertation*

On the one hand the variable Environment codes the number of underlying segments found in each word, on the other the segment following the conso-

---

[6]In §5.2 the composition of the data sets is explained in detail, i.e. it is explained why for some affixes fewer environments were included in the investigation than for others.

nant of interest and the presence/absence of a morphological boundary is coded. With regard to the number of consonants, words featuring the environment n#nV, m#mV and s#sV feature two identical consonants. All other environments only feature one corresponding underlying segment. To test gemination, one can test whether underlying doubles, i.e. the consonants in n#nV-, m#mV- and s#sV-words, are longer than the corresponding singletons in the other investigated structures. For example, if the nasal in *unnatural* (n#nV) is longer than the nasal in *natural* (#nV), the nasal in *untold* (n#C) and the nasal in *uneven* (n#V), the word *unnatural* geminates.

The existence of different environments with an underlying singleton can be explained by referring to three of the four investigated structures, i.e. singletons in complex words, singletons in bases and orthographic doubles in simplex words. The different levels represent the three different structures, i.e. singletons in complex words are coded as n#C, n#V, m#C, s#C and s#V , singletons in bases are coded as #nV, #mV and #sV, and orthographic doubles are coded as sV. Note that the # marks a morphological boundary. While singletons in complex words were included to compare durations of doubles and singletons in comparable structures, singletons in base words were included to ensure that the potential lengthening of the double consonant is not due to an inherently long base-initial consonant. Orthographic doubles in simplex words were included to test the influence of orthography on gemination (see sections 5.2.1 and 5.2.2 for a thorough discussion of the investigated structures and their relevance for the investigation).

Let us now turn to the second important aspect coded in the variable, the segment following the consonant of interest. The following segment is only relevant for singletons in complex words. This is because in all other structures the consonant of interest is always followed by a vowel, i.e. we do not find variability. As already noted in §5.2, it is very important to account for the difference between a following vowel and a following consonant. The reason is that the following segment might affect the duration of the consonant of interest. This is evidenced by phonetic studies which show that the duration of consonants heavily depends on the neighboring segment (see, for example, Umeda 1977). For nasals, following vowels lead to shorter durations, following consonants increase it. For voiceless fricatives, a following vowel leads to a longer duration than a following consonant. For voiced fricatives, the following segment does not influence duration (Umeda 1977: 854). To code for possible influences of the following segment, two different levels for singletons in complex words were coded for each affix, singletons in complex words followed by a vowel (n#V, s#V) and singletons in complex words followed by a consonant (n#C, m#C, s#C). Note that for the allomorph /ɪm/,

no singletons in complex words followed by a vowel could be included. This is because in /ɪm/-prefixed words singletons are always followed by a consonant.

The coded environments for the suffix *-ly* deviate from the ones coded for the prefixes. As can be seen in Table 5.6, all in all six different levels were coded for *-ly*. Out of these six levels two feature an underlying double consonant (l#l, syllabic l#l), and four an underlying singleton (l#, #l, syllabic l# , l). The four different singleton levels correspond to the three singleton structures investigated, i.e. singletons in complex words (#l), singletons in bases (l#, syllabic l#) and orthographic doubles in simplex words (l).

Table 5.6: Levels of the variable Environment for *-ly*

-ly Environment	Example
l#l	*really*
syllabic l#l	*ment(a)lly*
l#	*real*
syllabic l#	*ment(a)l*
#l	*probably, truly*
l	*belly*

The difference between the two environments with an underlying double, as well as the difference between the two environments for base words, is syllabicity. In base words and words featuring a underlying double, the lateral sometimes is syllabic. This occurs quite often when the base ends in the suffix *-al* (e.g. in *educationally/educational* or *mentally/mental*). The schwa-preceding /l/ is deleted, and /l/ becomes syllabic. The literature often claims that syllabic consonants have longer durations than non-syllabic consonants (see, for example, Jones 1959: 67; Clark & Yallop 1995: 135; Price 1981: 329). This claim is, however, only partly supported by empirical research. For instance, while Toft (2013) has shown that syllabic /l/ is longer than non-syllabic /l/, a study by Barry (2000) has found the opposite. To consider possible effects of syllabicity, I coded this factor in the variable Environment. If in a pertinent word (e.g. *mentally, mental*) the vowel preceding /l/ was deleted, i.e. the vowel was neither detected in the waveform, nor in the spectrogram, the word was coded as syllabic. Hence, two levels for doubles in complex words and two levels for doubles in base words emerged (l#l, syllabic l#l, l#, syllabic #l.)

As for the prefixed words, the consonant-adjacent segment might influence the

duration of the consonant of interest in *-ly*-words. For laterals, a preceding consonant leads to shortening (Umeda 1977: 851). Therefore, one must code for the type of preceding segment in *-ly*-suffixed words. To prevent the variable ENVIRONMENT from featuring too many levels, I coded the type of preceding segment in a separate variable (PRECEDINGSEGMENT) for *-ly*. The variable had two levels: `consonant` and `vowel`.

Semantic Transparency. The factor semantic transparency is important for testing the predictions made by the prosodic word approach, the morphological segmentability approach and the morphological informativeness approach. The variable represents one way of operationalizing decomposability. It has been used extensively in psycholinguistic research to investigate the question of whether words are processed as wholes or whether they are decomposed into their constituent morphemes (see, for example, Marslen-Wilson (2009) for an overview). These studies have shown that transparent words are more easily decomposed than non-transparent words. Thus, there is some evience that semantic transparency might be a well-suited measure of decomposability.

To test the pertinent predictions, and to investigate the suitability of semantic transparency as a measure of decomposability, I created the variable SEMANTICTRANSPARENCY. In the variable I coded whether the meaning of a derivative was transparent or opaque. I checked the meaning of each derivative, as well as the meaning of its base, in the online version of the *Oxford English Dictionary* (OED 2013). If the meaning of the derivative is fully compositional, i.e. it can straightforwardly be computed by combining the meaning of the affix with the meaning of the base, it was categorized as `transparent`. Examples of transparent words are *unnatural* and *impossible*. Words that were not fully compositional were categorized as `opaque` (e.g. *impression* and *imposed*).

Semantic Transparency Rating. SEMANTICTRANSPARENCYRATING is a second variable used to measure decomposability. It is thus primarily relevant for the segmentability approach and the question of how to operationalize decomposability. The variable is based on ratings in which all complex words included in the studies were rated for their decomposability in terms of semantic transparency. In an online experiment using LimeSurvey (LimeSurvey Project Team & Carsten Schmitz 2015) participants were asked how easy it is to decompose a given word into two meaningful parts on a scale from 1 ("very easy to decompose") to 4 ("very difficult to decompose"). Furthermore, participants could indicate if they did not know a specific word. In addition to complex words, the

study also included simplex words featuring the same phonemic strings as the affixed words. Including these simplex words made it possible to assess the validity of the rating by checking whether these words were rated as very difficult to decompose. Furthermore, inter- and intra-rater reliability were tested using different statistical procedures, such as calculating intra-rater correlations (Bartko 1966) and Cronbach's α (Cronbach 1951) .

I conducted two separate ratings, one for the corpus study and one for the experimental study. Since the corpus study investigates American English, the rating of the corpus data was done by native speakers of American English. For the experimental study, the participants of the production experiment, who were native speakers of British English, rated the items. In the corpus study, the medians of the ratings for each type were coded in the variable SemanticTransparencyRating. Since in the experimental study each recorded token was rated by the speakers themselves, it was possible to test the effect of the rating of each token directly. The design of the ratings, as well as their outcomes, i.e. their distributions and the results of the reliability tests, will be discussed further in the pertinent chapters.

Type of Base. The type of base of a derivative is of importance with regard to the predictions made by Stratal OT, the prosodic word approach and, because of its relation to segmentability, both psycholinguistic approaches. It is the third measure of decomposability used in this study. The factor is structural in nature and concerns the distinction between bound roots and words as bases. Derivatives with words as bases (e.g. *unnatural*) can be assumed to be more decomposable than words that have a bound root as their base (e.g. *implicit*). This distinction was coded for each derivative in the variable TypeOfBase. The variable has two levels: bound root and word.

Relative Frequency. Relative frequency is the fourth measure of decomposability used in this study. It is of relevance for the segmentability approach. Relative frequency is defined as the ratio of the frequency of a derived word to the frequency of its base (Hay 2003). The more frequent a derivative is in comparison to its base, the less decomposable is the complex word, and the higher is its relative frequency. I computed the variable RelativeFrequency by dividing a word's lemma frequency by its base lemma frequency. Since the variety of English deviates between studies, i.e. American English in the corpus study and British English in the experimental study, the frequencies for the two studies were extracted from two different databases. For the corpus data, frequencies

were extracted from the DVD version of COCA (Davies 2008–2014). For the experimental data, frequencies were extracted from the British National Corpus (Davies 2007). To allow for the calculation of relative frequency for all complex words in the data sets, the base frequency of derivatives with bound roots was set to 1, i.e. the lowest possible frequency. A base frequency of 1 automatically leads to a high relative frequency, which mirrors a very low degree of decomposability, which in turn mirrors the low decomposability of words with bound roots. I log-transformed the variable RELATIVEFREQUENCY before it entered the models.

Semantic Similarity (LSA). Semantic similarity is the fifth decomposability measure in this study. It is calculated with the help of latent semantic analysis, which compares the contexts of two words and calculates a similarity score (LSA score) on that basis. The more similar the contexts in which two words occur, the higher the pertinent LSA score (see Landauer et al. 1998). It can be assumed that if a derivative is very decomposable, its meaning will be similar to the meaning of its base. Therefore, it can be assumed that more decomposable words have a higher LSA score than less decomposable words. The higher score is assumed to mirror the semantic similarity between the derivative and its base, and in turn the derivative's decomposability. LSA scores have, for example, been used as measures of semantic transparency for compounds (see, for example, Wang et al. 2014; Gagné et al. 2016).

The variable LSAScore was coded by calculating the LSA score for the derivatives on a web-interface (University of Colorado Boulder 2017). As mentioned above, the data sets comprise derivatives with bound roots. As the LSA score can only be computed for words with a word as a base, the variable LSAScore was only coded for a subset of the data, i.e. derivatives with a word as a base. Furthermore, one should note that this measure of decomposability is rather exploratory, i.e. while the other possible measures of decomposability are well established and have been used in earlier studies, LSA scores have up until now only been used rarely in this way. Because the decomposability analyses of the corpus data revealed that the variable LSAScore did not correlate with the other decomposability measures, and because the variable did not affect gemination in the corpus study, the variable LSAScore was only used in the corpus study, i.e. it was not coded in the experimental study.

Affix. The variable AFFIX was coded with five levels: un, inLoc, inNeg, dis and ly. The variable is of interest in two ways. Firstly, as discussed in previous chap-

ters, morphological gemination is often assumed to depend on the affix involved. To test this assumption, one must compare the gemination behavior of the different affixes. While for the most part affixes have to be investigated using separate analyses, a few models were fitted in which affixes were compared directly. For these analyses it was necessary to code for the affix.

Secondly, the variable Affix is of interest with regard to the segmentability hierarchies introduced in §3.2. To validate the hierarchies one must compare the different decomposability measures across affixes, i.e. one needs to code for the affix. The segmentability hierarchies are, if validated by the data, relevant for the affix-specific predictions made by psycholinguistic approaches.

5.5.2 Noise variables: Phonetic factors

Consonant-specific factors. Some phonetic factors are only relevant for specific consonants, i.e. for specific affixes. Hence, they are only coded for specific subsets. The variable Voicing is one of them. Voiceless /s/ is longer than voiced /z/ (cf. Umeda 1977). This must be accounted for in the *dis*-data set. Therefore, the variable Voicing with the two levels `voiced` and `voiceless` was coded. The coding relied on the canonical pronunciation variant of the words found in the Longman pronunciation dictionary (Wells 2008). Voicing was only relevant for the corpus data since no *dis*-prefixed words with voiced fricatives were included in the experimental study.

Another variable only used for one of the five affixes is TypeOfL. It codes whether /l/ is pronounced as an approximant, a tap or as a vocalized /l/. The coding was based on the segmentation of -*ly*-suffixed words, as explained in §5.3. The three levels are `approximant`, `tap` and `vocalized`. Since there were no occurrences of vocalized /l/ in the corpus data, and only very few cases of taps, the variable TypeOfL was only used in the experimental data.

Duration of the Preceding Segment. There are two reasons for including the duration of the preceding segment in the analyses. The first reason is that it can be used as an extremely local measure of speech rate (as, for example, in Ernestus et al. 2006). The second reason is that, as discussed in Chapter 2, gemination may manifest itself on the vowel preceding the geminated segment (relative duration, for example, Ridouane 2010; Miller 1987; Oh & Redford 2012). In order to test whether there are effects of relative duration it is necessary to include PrecedingSegmentDuration in the models. The inclusion of this variable has the additional advantage that it helps to tease apart degemination effects and other kinds of reduction effects. In her study of the phonetics of *un*- Hay (2007)

finds, for example, that with declining decomposability, not only the nasal but the whole prefix becomes shorter. Including PRECEDINGSEGMENTDURATION as a noise variable controls for this effect.

Speech Rate. Speech rate can be defined as the number of linguistic units which are produced by a speaker in a given amount of time. The more units are uttered by a speaker in a certain amount of time, the shorter these units become. Different measures of speech rate are conceivable and the choice is largely determined by the kind of data at hand. One way of measuring speech rate is calculating the number of syllables per second (see, for example, Pluymaekers et al. 2005; Plag et al. 2017). To compute this ratio, relatively long strings of uninterrupted speech produced by one speaker are required. I used a similar measure in the experimental study. The variable GLOBALSPEECHRATE codes the number of words uttered per second. It is calculated by dividing sentence duration by number of words in the sentence. Since the sentence structure in the experiment was controlled for, i.e. except for the investigated word the sentences in the experiment are identical, and the investigated words do not vary much with regard to their length, this measure is comparable throughout the data set.

Due to a large amount of turn taking found in the Switchboard Corpus, neither the number of syllables per second, nor the number of words per second was feasible for the corpus data. Therefore, a second, more local measure of speech rate was calculated: the number of segments per second. This measure can be computed on the bases of the word alone, meaning no long strings of uninterrupted speech are necessary. I computed the values for the variable LOCALSPEECHRATE for each item by dividing the number of segments included in the word by the total word duration in seconds. This variable was coded for both data sets. It is expected that the higher the speech rate, the shorter the duration of the consonant(s) in question will be.

Word Length. There are various, closely related, measures of word length which potentially influence segment duration. The first one is word duration. The longer the duration of a word, the longer the duration of each segment. However, this measure is problematic since it is closely related to the variable LOCAL-SPEECHRATE, which is calculated by means of word duration. Using word duration as a separate variable would lead to serious statistical problems (collinearity). Furthermore, it is unnecessary to include word duration separately since it is already included in the variable LOCALSPEECHRATE.

Apart from word duration, there are two other measures of length which could

influence consonant duration – the number of syllables and the number of segments in the word. Early studies on Swedish and Dutch vowels have found that the more syllables a word consists of, the shorter a given vowel becomes (Lindblom 1963; Nooteboom 1972). Plag et al. (2017) have shown the same effect for word-final /s/ and /z/. To take these facts into account, I coded three types of phonological word length: the number of syllables in a word as noted in the lexical database CELEX (Baayen et al. 1995), the actual number of syllables in the word as coded by the annotators, and the number of segments in the word as coded by the annotators. All three of these measurements are highly correlated. Furthermore, they are correlated with the variable LocalSpeechRate, which is computed by means of number of segments in the word. This means that much of the variation brought in by word length, i.e. number of syllables or segments in the word, is already accounted for by including LocalSpeechRate. Therefore, none of the word length measures was included as a separate variable. Measures of word length are, however, implicitly integrated in the model using the variable LocalSpeechRate.

5.5.3 Noise variables: Phonological factors

Accentuation. Previous research has revealed that words which bear sentence accent show less reduction and a longer duration than words which are not accented (e.g. Sluijter & van Heuven 1996; Sugahara & Turk 2009; Bergmann 2014). The effect manifests itself in the duration of the individual segments of the word. Applied to this study, this translates into the prediction that in accented words the consonant in question will be pronounced with a longer duration than in unaccented words.

The variable Accentuation was only coded in the experimental data, in which items were produced in either accented or unaccented position (see Chapter 7 for details on both conditions). Items which were in accented position were coded as `accented`, items in unaccented position as `unaccented`. The reason for not including Accentuation in the corpus study is related to the type of speech investigated, and the difficulty to reliably code for accentuation in this type of speech. Generally, it is often impossible to hear a clear pitch accent in conversational speech. Annotation of accent is even more demanding in the recordings at hand, which are of rather poor quality. Because of this difficulty to code for accentuation in the corpus data, it was decided to first only code a subset of the data for pitch accent, and then decide, based on that subset, whether the rest of the data should also be coded. The annotation was done by two independent raters. Since the measure did not prove to be significant, i.e. it did not influence

consonant duration in the subset, it was decided to not code the rest of the data for accentuation.

Position. Words uttered at the end of an utterance or phrase have been shown to be pronounced with a longer duration than words in mid-positions (see, for example, Berkovits 1993; Hay 2007; Oller 1973). Some research found the lengthening effect being restricted to the final syllable of a word. For example, utterance-final position of *un*-prefixed words did not have a lengthening effect on prefixal /n/ (Hay 2007). But there is also evidence that segments occurring in the first syllable of a word participate in phrase- or utterance-final lengthening processes (Oller 1973). Ends of utterances and phrases are often marked by a pause. To account for possible effects of phrase-final lengthening, one can thus code whether a pause is present after an item. I included the variable PostPause with the two levels (pause) and (noPause) to account for possible effects of phrase-final lengthening.

In addition to effects of phrase-final position on duration, studies have also found that phrase-intial position affects duration. Items following a phrasal boundary are pronounced with longer durations. This might be due to initial-strengthening (see, for example, Cho & Keating 2001; Byrd et al. 2006; Cho et al. 2007). Again, the presence of a pause may be used as a marker for a phrasal boundary. As shown in Umeda (1977), segments after a pause are pronounced with a longer duration. I included the variable PrePause to account for possible effects of initial-strengthening. The variable coded whether a pause was present. It has two levels pause and noPause.

Preliminary inspection of the variable PostPause revealed that this variable did not have any effect on the duration of the corpus data. With regard to the variable PrePause, the corpus data only featured few items with a preceding pause. Including the variable was therefore not reasonable. Hence, both variables were only included in the experimental models.

Stress. Stressed syllables tend to have a longer duration than unstressed syllables (see, for example, Fry 1955; 1958; Lieberman 1960; Beckman 1986; Eriksson & Heldner 2016, see also Laver 1994 for an overview). For this study, this is relevant in two ways. First, a stressed affix is expected to feature a longer consonant than an unstressed affix. In other words, segments in stressed affixes might be longer than segments in unstressed affixes. It is therefore desirable to code for affix-stress. Second, the stress status of the affix-adjacent syllable, i.e. the first syllable of the base in case of prefixes and the penultimate syllable of *-ly*-affixed

words, might influence the duration of the investigated consonant(s).

Affix-adjacent stress might be relevant for the present investigation in various ways. First, it might influence the duration of morphological geminates. In case of a morphological geminate, the double consonant is represented by two underlying consonants (see §2.3 for discussion of the phonological representation of geminates). One of them belongs to the affix, the other belongs to the affix-adjacent syllable. The latter might participate in the stress-caused lengthening of the affix-adjacent syllable. The expectation that adjacent-syllable stress influences the duration of geminates is also supported by findings on lexical geminates (see, for example, Dmitrieva 2017 for discussion).

The second way in which affix-adjacent stress might be relevant is with regard to the structure singletons in bases (e.g. *natural*). In these words the consonant of interest is part of the base-initial syllable and might therefore be lengthened if it is stressed.

A third important aspect is that there might be an independent effect of affix-adjacent stress on the duration of the affixational consonant. Umeda (1977), for instance, found that nasals before unstressed vowels are shorter. For the data at hand, this could mean that affixational consonants which are adjacent to stressed syllables might be longer than affixational consonants which are adjacent to unstressed syllables. A possible explanation for this effect is that the lengthening of the adjacent stressed syllable spills over to the adjacent syllable.

To sum up, the stress status of the affix itself, as well as the stress status of its adjacent syllable, might influence consonant duration. One therefore needs to code for stress. While coding for affix-adjacent stress is not problematic, coding for affix-stress is quite challenging (see discussion in §3.1.1). While the suffix *-ly* is never stressed, the stress status of prefixes is difficult to determine and not well researched. While it seems uncontroversial that prefixes bear (secondary) stress when followed by an unstressed syllable, it is often unclear whether they are stressed or unstressed when followed by a stressed syllable. In pronunciation dictionaries, such as Wells (2008), the prefix in those cases is sometimes stressed, sometimes unstressed and sometimes variably stressed. However, as shown by Hanote et al. (2010) for the prefix *un-*, the stress assignment in Wells (2008) does not follow any systematic pattern. Furthermore, in conversational speech (as found in the corpus data), additional contextual factors might influence the stress status of the prefixes (cf. Videau & Hanote 2015). The matter is further complicated by the difficulty to determine the relative prominence relation between the prefix and a following stressed syllable, i.e. coding prefix stress is quite challenging.

Because of the difficulty to code prefix-stress (unsystematic annotation in dictionaries, potential contextual influences, difficulty of determining prefixal stress based on acoustic properties) I did not explicitly code the stress status of the prefix. Instead, I coded for the stress status of the affix-adjacent syllable. The lexical stress status of the base-initial syllable (or the base-final syllable in case of *-ly*) is uncontroversial. I used the Longman Pronunciation Dictionary (Wells 2008) for the coding. As explained above, only when the base-initial syllable is stressed, a prefix can be unstressed. If the base-initial syllable is unstressed, the prefix must be stressed. Therefore, one can at least partially account for prefixal stress by coding for the stress status of the affix-adjacent syllable of a prefixed word. Coding for base-initial and base-final stress is also relevant in view of the possible independent effect of the affix-adjacent syllable.

Stress was thus coded with regard to the affix-adjacent syllable. For prefixes, the variable BaseInitialStress was coded with the two levels `stressed` and `unstressed`. For the suffix *-ly*, the variable BaseFinalStress was coded with the two levels `stressed` and `unstressed`.

5.5.4 Noise variables: Lexical factors

Word Form Frequency. Frequency has been shown to affect the duration of a word. More frequent words tend to have shorter durations (see, for example, Aylett & Turk 2004; Gahl 2008). Frequency was therefore included as a covariate. I collected two different types of frequency, word form frequency and word lemma frequency. The frequencies for the corpus study were extracted from COCA (Davies 2008–2014). The frequencies for the experimental study were taken from the BNC (Davies 2007). Preliminary inspection of the data revealed that the two frequency measurements highly correlate. This means that it essentially does not make a difference whether one tests the influence of word form frequency or lemma frequency on duration. Hence, only one was included in the models: WordFormFrequency. I log-transformed this variable before it entered the models.

5.5.5 Overview of variables

Table 5.7 summarizes which variables were included in the two studies. All variables are listed in the middle. They are divided into variables of interest and noise variables. On the left side of the table one can see which variables were included in the corpus study, on the right side which variables were included in the experimental study.

Table 5.7: Overview of variables in the two studies

	Corpus Study	Variable	Experimental Study
	yes	ENVIRONMENT	yes
	yes	SEMANTICTRANSPARENCY	yes
	yes	SEMANTICTRANSPARENCYRATING	yes
Variables of interest	yes	TYPEOFBASE	yes
	yes	logRELATIVEFREQUENCY	yes
	yes	LSASCORE	no
	yes	AFFIX	yes
	yes	PRECEDINGSEGMENTDURATION	yes
	yes	LOCALSPEECHRATE	yes
	no	GLOBALSPEECHRATE	yes
	yes	VOICING	no
	no	TYPEOFL	yes
Noise variables	no	ACCENTUATION	yes
	no	PREPAUSE	yes
	no	POSTPAUSE	yes
	yes	BASEFINALSTRESS	yes
	yes	BASEINITIALSTRESS	yes
	yes	logWORDFORMFREQUENCY	yes

6 Corpus Study

The corpus study was conducted to investigate gemination in natural, conversational speech. The data was extracted from the Switchboard Corpus (Godfrey & Holliman 1997), which consists of over 3 million words and comprises North American conversational speech. The methodology of the corpus study followed the general methodology described in the previous chapter. However, as the previous chapter was restricted to the description of the general methodology followed in both gemination studies presented in this book, i.e. the corpus and the experimental study, it is necessary to describe the methodology specific to the corpus study in further detail. This will be done in the first part of this chapter. First, I will describe how the corpus data was collected. Then, I will turn to the decomposability rating which was conducted to code the variable SemanticTransparencyRating. After having discussed the validity and reliability of the rating, I will briefly describe the annotation of the variables in the data set. I will then turn to the two different types of analyses conducted. On the one hand, the data was analyzed with regard to decomposability, on the other it was analyzed with regard to consonant duration, i.e. gemination. First, I will discuss the decomposability analyses and lay out their results. Then, I will turn to the durational analyses. Again, I will describe the conducted analyses and discuss results. At the end of the chapter, I will summarize the results and briefly discuss them with regard to the predictions made in Chapter 4. An extensive discussion of the results and their theoretical implications will be conducted in Chapter 8 alongside with the results from the experimental study.[1]

6.1 Methodology

6.1.1 Sampling

In the corpus study four subsets of complex words were investigated. One subset contains *un*-prefixed words, one *in*-prefixed words, one *dis*-prefixed words and

---

[1]Earlier versions of sections 6.1.1, 6.1.3, 6.3.2, 6.3.3, 6.3.4, and 6.3.5 of this chapter have been published in Ben Hedia & Plag (2017).

one *-ly*-suffixed words. The complex words either feature a phonological dou-
ble consonant at the morphological boundary (e.g. *unnatural*) or a phonological
singleton (e.g. *uneven*). All double consonants in the data set are surrounded by
vowels. Singletons in prefixed words are either followed by a vowel or a conso-
nant. Singletons in suffixed words are preceded by either a vowel or a consonant.

  To compile the four data sets, I first extracted all words with one of the desired
phonological structures from the corpus using the speech corpus management
system LaBB-CAT (Fromont & Hay 2012; Fromont 2003-2015). I checked all ex-
tracted words for their morphological status by using the criteria described in
§3.1.2. As expected, the number of prefixed words with morphological geminates
was quite low. For the prefix *in-*, it turned out that the corpus only contained a
few /ɪn/-prefixed tokens with two underlying /n/s at the morphological bound-
ary (17 tokens of 5 types). Therefore, I decided to not include the allomorph /ɪn/
in the study, but instead to focus on /ɪm/ for which enough tokens were found.
This means that to investigate *in-*, its allomorph /ɪm/ was used (see also sections
3.3 and 5.2.1 for discussion).

  Similarly to *in-* (in its allomorphic form /ɪn/), the number of tokens with a dou-
ble consonant was quite low for *un-* and *dis-*. For *un-* the corpus contained 22 per-
tinent tokens, for *dis-* it contained 24 pertinent tokens. In contrast to *in-*, neither
un- nor *dis-* feature allomorphic variants. In other words, there are no additional
tokens with morphological geminates which can be investigated. Therefore, I in-
cluded all of the available *un-* and *dis-*prefixed tokens with a phonological double,
i.e. 22 for *un-* and 24 for *dis-*.

  For *in-* (in its allomorphic form /ɪm/) and *-ly*, the number of tokens with a
phonological double was much higher than for *un-* and *dis-*. However, a closer
look revealed that a lot of the available tokens were of the same type. Including a
lot of tokens of the same type was not desirable for two reasons. First, one aim of
this study is to investigate word-specific factors. It is therefore crucial to include
a high number of different types. Second, including a high number of tokens of
one type, while only including a low number of tokens of another, might cause
serious statistical problems. If the type-token ratio in a data set is very unbal-
anced, statistical models will be highly influenced by specific types, i.e. by those
types which occur frequently in the data set. Including a random effect for type
in the model would not provide a solution. This is because the variation found
in types which only occur a few times in the data set would be accounted for
by the random effect exclusively, i.e. the variation of infrequent types would be
explained by inherent differences between types. This would make it impossible
to find any other effects. It was thus decided to avoid large numbers of tokens of

the same type by restricting the number of tokens with a double consonant for *in-* and *-ly* to 90.

The 90 *un-* and the 90 *-ly*-affixed tokens with a phonological double were semi-randomly selected from the list of all available tokens. The semi-random selection included selecting as many different types as possible, and selecting only one token of each type from one pertinent speaker. Most speakers in the Switchboard Corpus provided us with only one pertinent token, which would have made it impossible to incorporate speaker-specific effects in the statistical analysis. I therefore decided to include only one token of a specific type from one given speaker.

Note that, while selecting only one token of one type per speaker was possible for *in-* and *-ly*, the small number of items with double consonants in the *un-* and *dis-*subsets prohibited this selection procedure for *un-* and *dis-*. In other words, while for *in-* and *-ly*, all tokens of the same type come from different speakers, and all speakers provided only one token per type, for *un-* and *dis-*, two speakers contributed two tokens for one type.

For all affixes, the selection of words with a singleton at the morphological boundary was executed in a similar way to the selection of *in-* and *-ly*-affixed words with a double consonant, i.e. the items were semi-randomly selected. To keep the data sets as large as possible (to ensure sufficient statistical power), but to also keep the differences in size between the subsets relatively small, the number of singleton tokens in the study was restricted. For the prefixes, 70 tokens were sampled for each of the singleton environments. For *-ly*, 90 tokens were sampled for each environment. Hence, I included 70 *in-*prefixed tokens with a single nasal (e.g. *impossible*), 90 *-ly*-suffixed tokens with a single lateral (e.g. *truly, probably*), and 140 *un-* and *dis-*prefixed tokens with a singleton each. For *un-* and *dis-*, 70 of the singletons were followed by a vowel (e.g. *unable, disarm*) and 70 were followed by a consonant (e.g. *unfit, disgrace*).

After compiling the list of word tokens to be included, I extracted the sound files containing the pertinent tokens from the corpus. All tokens were segmented according to the criteria described in §5.3, and afterwards coded for their environment. Some of the tokens sampled had to be removed after closer inspection of the sound files, for example because the quality of the recording was insufficient to provide valid segmentation, or because they were pronounced in very unnaturally-sounding ways. The final data sets were of comparable size and contained 158 *un-*prefixed words, 156 *in-*prefixed words (with the allomorph /ɪm/), 128 *dis-*prefixed tokens and 154 *ly*-suffixed words. The type and token distribution for each environment for each affix is displayed in Table 6.1.

Table 6.1: Distribution of types and tokens in corpus study

un-

Environment	Example	Number of Types	Number of Tokens
n#nV	*unnatural*	6	23
n#C	*untold*	53	68
n#V	*uneven*	42	67
Total		101	158

in-

Environment	Example	Number of Types	Number of Tokens
m#mV	*immortal*	16	89
m#C	*impossible*	67	67
Total		83	156

dis-

Environment	Example	Number of Types	Number of Tokens
s#sV	*dissatisfied*	9	24
s#C	*disgrace*	21	45
s#V	*disarm*	34	59
Total		64	128

-ly

Environment	Example	Number of Types	Number of Tokens
non-syllabic l#l	*really*	29	33
syllabic l#l	*ment(a)lly*	48	48
#l	*possibly*	73	73
Total		150	154

Draft of May 27, 2019, 18:04

6.1.2 The decomposability rating

All types included in the corpus study were rated for their decomposability in terms of semantic transparency. The rating was carried out online using the software LimeSurvey (LimeSurvey Project Team & Carsten Schmitz 2015). All raters were native speakers of American English. The median rating of each item was coded in the variable SEMANTICTRANSPARENCYRATING. Before the median was computed, the ratings were tested for their reliability and their validity. In this section, I will describe the design of the rating and the analyses conducted to ensure reliability and validity.

The rating was designed similar to the rating conducted in Hay (2001), in which participants were asked to decide which of two derivatives was more complex. As in Hay's task, participants read an explanation about the make-up of complex words before rating the words' decomposability. The participants were informed that some words consist of more than one meaningful unit and were given examples. They were then asked to rate on a scale from 1 to 4 how easy it is to decompose the words into two meaningful units. The participants were explicitly told which units to look for, e.g. *in* and the rest of the word for the rating of *in*-prefixed words, and *un* and the rest of the word for *un*-prefixed words. The participants were given the opportunity to indicate if they did not know a word. Some biographical information about the participants, such as their language background, their profession, and their knowledge about linguistics, was recorded in order to test whether any of these factors influenced the rating.[2]

Four ratings were conducted, two ratings included *un-* and *in-*prefixed words, one *dis-* prefixed words and one *-ly*-affixed words. In addition to the complex words included in the corpus study, the ratings also included simplex words which featured the same orthographic strings as the affixed words, e.g. *uncle* and *family*. They were included to test whether the raters understood the task correctly. Simplex words should be rated as being very difficult to decompose.

All in all ,133 native speakers of American English between the ages of 14 and 62 rated 450 items. After the ratings were conducted, I analyzed the data in order to test the validity and the reliability of the rating. As a first step, I checked the distribution of ratings for each participant. Participants who did not vary in their rating, i.e. who rated all the items the same, and participants who rated the simplex items as very decomposable were excluded. It can be assumed that they either did not understand the task correctly, or did not fulfill the task

---

[2]The questionnaire the participants of the rating studies filled out, and the instructions they were given can be found in appendix A.

thoroughly. Their ratings were not valid. To further check the ratings' reliability, I computed three consistency estimates of inter-rater reliability, the intraclass correlation coefficient (intraclass correlation coefficient, Bartko 1966), item-total correlations and Cronbach's α (Cronbach 1951). All three measures can be used to test whether the subjects rated the items with internal consistency (see also Stemler & Tsai 2008: 38ff. for discussion).

The intraclass correlation coefficient (intraclass correlation coefficient) assesses the agreement rate of multiple responses across different stimuli, i.e. it assesses the agreement in the ratings of all raters across all items. Two different ICCs were computed, intraclass correlation coefficient 2 and intraclass correlation co-efficient 3. While intraclass correlation coefficient 2 takes differences in the average ratings of the participants into account, intraclass correlation coefficient 3 does not. For example, if rater 1 rated item A with 1 and item B with 2, while rater 2 rated item A with 2 and item B with 3, the intraclass correlation coefficient 2 would show perfect agreement between the two raters (an intraclass correlation coefficient 2 of 1). Taking the difference in the intercept into account, rater 1 and rater 2 agreed perfectly by rating item B with one value more than item A. The absolute agreement of the ratings would, however, not be perfect because the two raters did not rate both items with the same values. This degree of absolute agreement is shown by intraclass correlation coefficient 3, which would be below 1 in the example.

The second measure of consistency are item-total correlations. This measure computes the correlation between the ratings of one participant and the ratings of all other participants across all items. This means that one value for each participant is computed. This measure can be used to detect participants which differed to a high degree from the other participants in their rating. After inspecting the distribution of item-total correlations across speakers, I excluded all participants whose item-total correlation was clearly lower than the other participants' correlations. This resulted in excluding all participants who featured an item-total correlation below 0.6.

The third measure of inter-rater reliability is Cronbach's α, which is a reliability coefficient often used to measure internal consistency. Cronbach's α displays the average inter-correlation among raters across items. Its value is between 0 and 1, with 1 indicating excellent internal consistency and 0 unacceptable internal consistency. I computed all measures using the R package `psych` (Revelle 2017).

After excluding invalid ratings, based on the inspection of the ratings' distributions and the analysis of inter-rater reliability, the ratings proved to be valid and reliable. This is indicated by the consistency estimates which are summarized in

Table 6.2 for all ratings. For the item-total correlations, the minimum, the maximum, the mean and the median correlations for each rating are displayed. The table also shows the final number of raters for each rating, i.e. the number of raters after the exclusion of invalid raters. All in all, 23 raters were excluded.

Table 6.2: Overview consistency estimates for all ratings in corpus study

	number of raters	Cronbach's α	ICC1	intraclass correlation coefficient 2	item-total correlation			
					Min	Max	Mean	Median
Rating 1: *un*- and *in*-	26	0.99	0.75	0.77	0.69	0.95	0.88	0.91
Rating 2: *un*- and *in*-	32	0.99	0. 69	0.72	0.69	0.94	0.85	0.87
Rating 3: *dis*-	32	0.98	0.62	0.66	0.63	0.91	0.83	0.81
Rating 4: *-ly*	20	0.98	0.72	0.72	0 .62	1.0	0.93	0.99

After the reliability of the ratings was ensured, I checked for possible effects of age, sex and linguistic background (including knowledge of Latin and linguistics) on the rating by conducting regression models. The dependent variable of these models was the rating. Fixed factors included age, sex, linguistic background and knowledge of Latin. The participant was included as a random effect. The analyses showed that neither age, nor sex, nor the linguistic background or knowledge of Latin had a significant effect on the rating.

I computed the median rating for each type and coded it in the variable Seman-ticTransparencyRating. I chose to code the median in the variable, i.e. not the mean, to avoid the influence of a few extreme ratings.

6.1.3 Annotation

After extracting and phonetically annotating the words included in the corpus study, they were annotated with regard to factors possibly influencing consonant duration. These factors were described in detail in §5.5. Because of differences between corpus and experimental data, not all of the discussed variables were included in both studies. For the corpus study, the five variables Glob-alSpeechRate, TypeOfL, Accentuation, PrePause and PostPause were not coded (see variable descriptions in §5.5 for a discussion on why these variables were not included). All other variables described in §5.5 were included in the

study. Overviews of all variables initially included in the models predicting absolute consonant duration are given in Tables B.1 – B.4 in appendix B.

6.2 Decomposability

Analyzing decomposability had two aims. The first aim was to test the suitability of the decomposability measures used in this study. This was done by investigating the relation between the different decomposability measures. The second aim was to test the validity of the segmentability hierarchies proposed in §3.2. This is important with regard to the theoretical predictions proposed in Chapter 4. To test whether the hierarchies are validated by the data, the segmentability of the five affixes was compared.

6.2.1 The relation between decomposability measures

As discussed thoroughly in §4.3.1, we find different operationalizations of decomposability in the literature. It is yet unclear whether all of them are well-suited measures of the concept. In this study five variables are used as potential measures of decomposability: SemanticTransparency, SemanticTransparencyRating, TypeOfBase, logRe-lativeFrequency and LSAScore. To test whether all of these variables tap into the same phenomenon, i.e. decomposability, it is necessary to investigate their relation to each other. If they are all measures of the same underlying property, they are expected to be highly correlated. To test the relation between the variables, hierarchical cluster analyses were conducted.

Hierarchical cluster analyses are usually used to find patterns and groups with similar characteristics in a data set (see, for example, Baayen 2008: chapter 5.1.5, Zumel & Mount 2014: chapter 8.1). They can also successfully be applied to investigate the similarity between different variables, and are thus well suited to provide insights into the relation between the five decomposability variables (see Baayen 2008: 200f.). The type of cluster analysis applied here first computes Spearman's rank correlations between all included variables, then squares them, and then puts them into a correlation matrix. This matrix hence includes pairwise comparisons of all decomposability variables. The use of squared figures is motivated by the need to avoid negative values. Since correlations can only be calculated for numerical variables, categorical variables have to be recoded into numerical variables in order to be included into the analysis. Therefore, the categorical variable SemanticTransparency was recoded into the numerical variable numSemanticTransparency, and the categorical variable TypeOfBase was

recoded into the numerical variable NUMTYPEOFBASE. The variable NUMSEMAN-TICTRANSPARENCY featured the two levels 0 (= opaque) and 1 (= transparent), the variable NUMTYPEOFBASE the levels 0 (= bound root) and 1 (= word as base). After the correlation matrix was created, a dendrogram in which the correlations are displayed was generated. Variables which cluster together in the dendrogram are more similar, i.e. have a higher correlation, than variables which are more distant in the dendrogram.

I conducted five cluster analyses, one for each subset and one for the whole data set. I decided to not only conduct one analysis looking at the whole data set, but to also investigate the relation between the variables in each of the four subsets, to check for possible differences between subsets. The affixes might deviate in their segmentability (as will be discussed in further detail in §6.2.2), and these differences between affixes might be mirrored in the relation between the different decomposability variables in the subsets, i.e. the correlations might deviate between subsets.

All analyses were based on types, not on tokens. Types were chosen to avoid that the decomposability features of one very frequent type in the data set affect the results for the whole data set. In other words, to avoid that the correlations between the variables were influenced by the relation between them in one specific type, each type was only included once in the analyses. All in all, 396 types were analyzed. However, for the correlations with the variable LSASCORE, not all types could be considered. This is due to the fact that, as explained in §5.5.1, a similarity score can be computed only for derivatives with words as bases, and thus only derivatives with words as bases were coded for LSASCORE. It follows that the correlations between the variable LSASCORE and all other variables only considered those items for which the score was available (303 types). For all other correlations all 396 types were taken into consideration. All cluster analyses were generated in R using the `Hmisc package` (Harrell Jr 2017).

The first cluster analysis conducted investigated all types of the corpus study, i.e. derivatives with all five affixes. The correlation matrix created in the analysis is shown in Table 6.3. The table shows that the highest correlations are between the variables NUMSEMANTICTRANSPARENCY, SEMANTICTRANSPARENCY-RATING and NUMTYPEOFBASE. The correlations between logRELATIVEFREQUENCY and LSASCORE and all other variables are much lower.

Figure 6.1 displays the relation between the variables in a dendrogram. On the y-axis the squared Spearman correlation score between the variables is displayed. The figure shows three splits which structure the variables into four clusters. The lower the splits in the figure, the higher are the correlations between

Table 6.3: Correlation matrix for decomposability measures in corpus study

	numSemantic-Transparency	logRelative-Frequency	SemanticTrans-parencyRating	numType-OfBase	LSAScore
numSem.Transp.	1.00				
logRel.Freq.	0.09	1.00			
Sem.Transp.Rating	0.70	0.10	1.00		
numTypeOfBase	0.65	0.08	0.70	1.00	
LSAScore	0.11	0.03	0.08	0.05	1.00

the variables of the pertinent clusters. The first split separates the variable LSA-Score from all other variables. LSAScore thus forms its own cluster, which indicates the dissimilarity of this variable to the other four decomposability variables. The second split separates the variable logRelativeFrequency from numSemanticTransparency, SemanticTransparencyRating and numTypeOfBase. This shows that logRelativeFrequency is also rather dissimilar from the other three decomposability variables. That logRelativeFrequency and LSAScore are different from the other three decomposability variables is also shown by the high position of the first two splits in the figure.
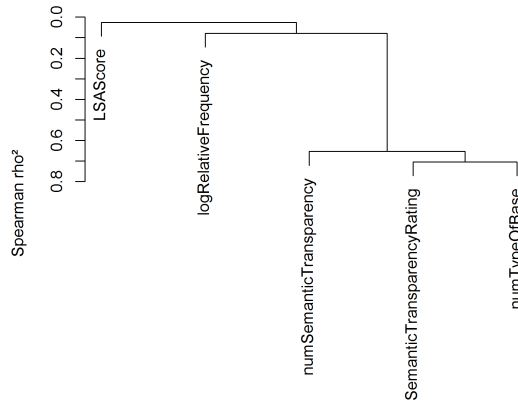
Figure 6.1: Dendrogram of the five decomposability measures for all types in the corpus study

The third split is the one splitting NUMSEMANTICTRANSPARENCY from SEMAN-TICTRANSPARENCYRATING and NUMTYPEOFBASE. This split is displayed in the lower part of the figure, which indicates that even though SEMANTICTRANS-PARENCYRATING and NUMTYPEOFBASE are more similar to each other than they are to NUMSEMANTICTRANSPARENCY, all three variables correlate to a high degree and are very similar.

Let us now turn to the analyses of the individual affixes. It turned out that fitting a cluster analysis was only reasonable for the prefixes *in-* and *dis-*, i.e. it was not reasonable to conduct cluster analyses for *un-* and *-ly*. This is due to the distribution of the decomposability variables in the *un-* and the *-ly*-data sets. Three of the five decomposability variables do not show enough variability to be investigated. All of the *un-* and *ly*-affixed words are semantically transparent, and only a few feature a bound root. Furthermore, all *un-*prefixed, and the majority of *-ly*-suffixed words were rated as very decomposable. Because of this lack of variability in the variables SEMANTICTRANSPARENCY, SEMANTICTRANSPARENCY-RATING and TYPEOFBASE, only the correlation between logRELATIVEFREQUENCY and LSASCORE was investigated for *un-* and *-ly*. The squared Spearman correlation score is below 0.05 for all correlations tested. This means there is no indication that the two variables logRELATIVEFREQUENCY and LSASCORE can be used as the operationalization of the same concept.

For *in-* and *dis-*, all variables showed enough variation to conduct cluster analyses. The results of the analyses for the two prefixes are displayed in the dendrograms in Figure 6.2. For both prefixes, the variables NUMSEMANTICTRANS-PARENCY, SEMANTICTRANSPARENCY-RATING and NUMTYPEOFBASE cluster together in the lower part of the figure. This means that the correlations between the three variables are quite high. LSASCORE and logRELATIVEFREQUENCY, on the other hand, do not correlate to a high degree with any other variable. This resembles the results of the first cluster analysis.

To sum up, the cluster analyses have revealed that the three variables SEMANTIC-TRANSPARENCY, SEMANTICTRANSPARENCYRATING and TYPEOFBASE are highly correlated. The two variables logRELATIVEFREQUENCY and LSASCORE, in contrast, barely correlate with the other decomposability variables. These outcomes can be interpreted in the following way: while the three variables SEMANTICTRANSPAR-ENCY, SEMANTICTRANSPARENCYRATING and TYPEOFBASE can certainly be used as measures of the same concept, this cannot be stated for the two variables logRELATIVEFREQUENCY and LSASCORE. In other words, while possible effects of SEMANTICTRANSPARENCY, SEMANTICTRANSPARENCYRATING and TYPEOFBASE on duration are probably caused by the same underlying property, any effects of
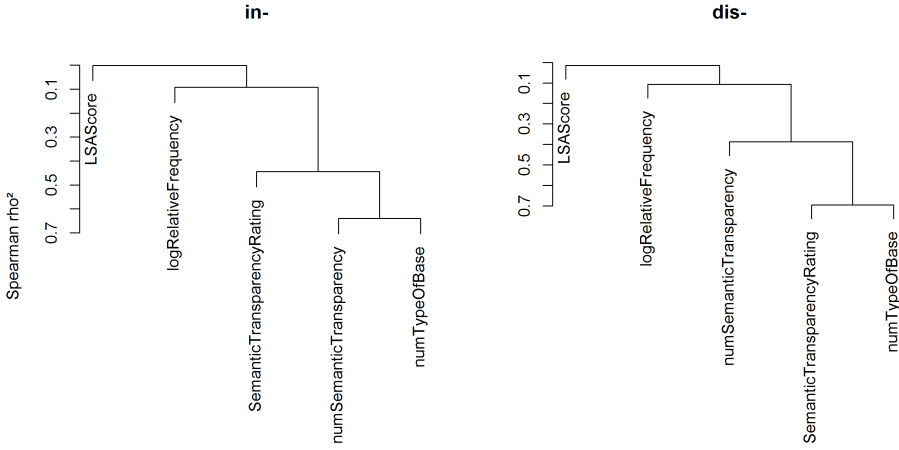
Figure 6.2: Dendrogram of the five decomposability measures for *in-* and *dis*-prefixed words in the corpus study

logRELATIVEFREQUENCY and LSASCORE can be regarded as independent.

The result that only three of the five investigated variables are measuring the same underlying property, i.e. that two measures are independent, raises the question of which variable, or which set of variables, is the right measure of decomposability. The answer to this question is not trivial and depends on one's definition of the concept DECOMPOSABILITY. As thoroughly discussed in §4.3.1, decomposability is not explicitly defined in the literature and different operationalizations are used. In other words, decomposability is not one uniform theoretical concept and cannot be treated as such. Different operationalizations call for different definitions of the concept. The analyses conducted in this section indicate that the decomposability measures used in this study form operationalizations of three different types of decomposability. The first type is operationalized by SEMANTICTRANSPARENCY, TYPEOFBASE and SEMANTICTRANSPARENCYRATING, the second is operationalized by logRELATIVEFREQUENCY, and the third is operationalized by LSASCORE. The answer to the question of which variable measures decomposability is thus that all five variables are measures of decomposability but that they do not measure the same type of decomposability.

6.2.2 The segmentability of the affixes: A comparison

Two of the theoretical predictions about gemination, i.e. the affix-specific morphological segmentability prediction and the affix-specific morphological informativeness prediction (cf. §4.3.1), are based on the lexical segmentability hierarchies proposed in §3.2. To ensure the validity of these predictions, it is necessary to ensure the validity of the proposed segmentability hierarchies (see Figure 6.3 for a repetition of the hierarchies). The Semantic Segmentability Hierarchy is mainly based on a qualitative analysis of the affixes' semantics (see discussion on semantics of the affixes in §3.2). Testing its validity by a quantitative analysis, such as the one applied here, is therefore quite challenging. The validity of the Non-Semantic Segmentability Hierarchy can, however, well be tested by checking whether the hierarchy is mirrored in the distribution of the decomposability measures in this data set.

	Segmentability hierarchy	Additional assumption
Semantic Hierarchy	*un-* > {*dis-*, *in-*$_{\text{NEG}}$}> *in-*$_{\text{LOC}}$ > *-ly*	lexical meaning over productivity, transparency and type of base
Non-Semantic Hierarchy	*un-* > *-ly* > {*dis-*, *in-*$_{\text{NEG}}$}> *in-*$_{\text{LOC}}$	productivity, transparency and type of base over lexical meaning

Figure 6.3: Lexical segmentability hierarchies of affixes

I tested the validity of the Non-Semantic Segmentability Hierarchy by comparing the segmentability of the five investigated affixes as found in the data. I looked at the distributions of the decomposability variables[3] for each affix in the data set, and used standard test statistics, such as the χ-square-test and the

---

[3]Note that I will continue to use the term *decomposability measures* for the five variables SEMANTICTRANSPARENCY, SEMANTICTRANSPARENCYRATING, TYPEOFBASE, logRELATIVEFREQUENCY and LSASCORE, even though the results of the cluster analyses have revealed that these variables do not form measures of the same underlying property. I will continue to use the term *decomposability variables* because of two reasons. First, to avoid confusion. The term is used in the literature as well as §5 of this book to refer to these variables, and it might cause confusion to change the terminology at this point. Second, even though the variables might not measure

Kruskal-Wallis test, to see whether differences between affixes were statistically significant. If the Non-Semantic Segmentability Hierarchy is valid, the comparison should reveal the same segmentability hierarchy as the one proposed.

First, I looked at the distribution of the variable SEMANTICTRANSPARENCY for each affix. The distribution is shown in Table 6.4. The table shows how many types of each affix were classified as semantically opaque, and how many were classified as semantically transparent. The percentage of opaque and transparent types per affix is given in parentheses next to the total number of types. The more transparent types an affix features, the more segmentable it is. The affixes are ordered from the least to the most segmentable.

Table 6.4: Semantic Transparency by affix

SEMANTIC TRANSPARENCY	$in\text{-}_{\text{Loc}}$		$dis\text{-}$		$in\text{-}_{\text{Neg}}$		$un\text{-}$		$\text{-}ly$	
opaque	42	(78%)	28	(45%)	7	(24%)	0	(0%)	0	(0%)
transparent	12	(22%)	34	(55%)	22	(76%)	101	(100%)	150	(100%)

The affixes *un-* and *-ly* only feature transparent items, i.e. they are the semantically most transparent affixes out of the five. They are followed by negative *in-* and *dis-*. Locative *in-* has the most opaque items and is thus the least segmentable affix in terms of semantic transparency. A χ-square-test ($\chi\text{-}squared$ = 204.48, df = 4, p < 0.001) and pairwise comparisons of proportions (see, for example, Crawley 2012: chapter 6.5) revealed that the contrasts between all affixes are significant (except for the one between *un-* and *-ly*).

The distribution of the variable SEMANTICTRANSPARENCYRATING reveals a similar picture. Table 6.5 shows the distribution of the median ratings (of types) for each affix. Again, the total number of types and the percentage of types are given. The two affixes *un-* and *-ly* were rated to be the easiest to segment. Locative *in-* and *dis-* were rated to be the most difficult to segment.

To test whether the differences between affixes were significant, pairwise Kruskal-Wallis tests were applied. Only a few differences proved to be significant. The ratings for locative *in-* differ significantly from the ratings for all other affixes. Furthermore, the ratings for *dis-* differ significantly from the ratings for *un-* and *-ly*. One can thus say that there is a significant difference between the affixes

---

the same underlying property, it can still be argued that they all form operationalizations of decomposability. Importantly, decomposability has to be defined in different terms depending on the decomposability measure used.

Draft of May 27, 2019, 18:04

Table 6.5: Semantic Transparency Rating by affix

Semantic Transparency Rating	$in\text{-}_{\text{Loc}}$		dis-		$in\text{-}_{\text{Neg}}$		-ly		un-	
1 - most decomposable	2	4%)	40	(65%)	20	(69%)	145	(97%)	101	(100%)
2	13	(24%)	4	(6%)	1	(3%)	4	(2%)	0	(0%)
3	25	(46%)	14	(23%)	4	(14%)	1	(1%)	0	(0%)
4 - least decomposable	14	(26%)	4	(6%)	4	(14%)	0	(0%)	0	(0%)

which were rated as the most difficult to segment ($in\text{-}_{\text{Loc}}$ and *dis-*) and the affixes which were rated as being the most easy to segment (*un-* and *-ly*). However, due to the small size of the tested data sets combined with the rather high number of levels (i.e. 4 levels for each affix) and the differences in number of observations between affixes, one must be cautions to not over-interpret the significance, or insignificance, of the contrasts. All in all, the rating clearly shows that locative *in-* is the least segmentable affix, and that *un-* and *-ly* are the most segmentable affixes.

Table 6.6 gives the distribution for the variable Type Of Base for all affixes. It shows that locative *in-*, unlike all other affixes, has a strong preference for bound roots. Almost all of the *un-* and *-ly*-affixed words feature a word as a base. For negative *in-* and *dis-*, there are quite a few words which feature a bound root. As indicated by pairwise tests of proportions, the differences between all affixes, except for the one between *un-* and *-ly* and the one between negative *in-* and *dis-*, are significant. On can thus state that in terms of type of base, locative *in-* is the least segmentable affix, *un-* and *-ly* are the most segmentable affixes, and *dis-* and negative *in-* pattern in between.

Table 6.6: Type of base by affix

Type Of Base	$in\text{-}_{\text{Loc}}$		dis-		$in\text{-}_{\text{Neg}}$		un-		-ly	
bound root	44	(81%)	15	(24%)	7	(24%)	1	(1%)	1	(1%)
word	10	(19%)	47	(76%)	22	(76%)	100	(99%)	149	(99%)

Turning to the gradient measures of decomposability, Figure 6.4 displays the distribution of the variable logRelativeFrequency for the five affixes using boxplots. The logarithmized relative frequency is displayed on the x-axis of the plot.

Each of the five boxes in the plot represents 50% of the types of one affix. The black dot in each box marks the median relative frequency for each affix. For example, the graph shows that 50% of the *-ly*-words have a logarithmized relative frequency between −2.191 and −0.058. It also shows that half of the *-ly*-words have a relative frequency above −1.415, and half have a relative frequency below −1.415. All significant differences are indicated in the plot.
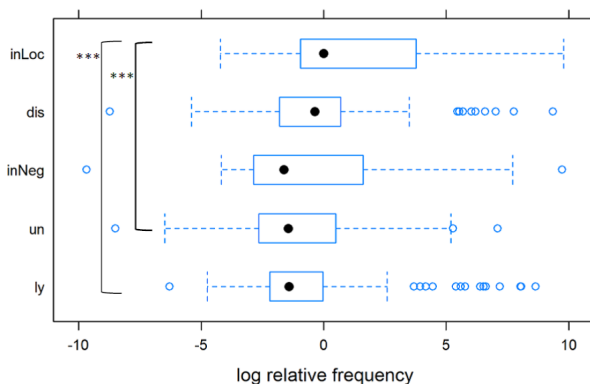
Figure 6.4: Comparison of the relative frequency of the five affixes

The plot suggests that locative *in-* is the least segmentable affix having the highest relative frequency. It is followed by *dis-*. The median relative frequencies of *in-*, *un-* and *-ly* do not differ to a high degree, but the plot suggests that, with regard to their decomposability, *-ly* and *un-*affixed words are more uniform than words with negative *in-*. In other words, while the majority of the *-ly-* and *un-*words have a rather low relative frequency, we find some variation with negative *in-*. An ANOVA (F = 6.716, p < 0.001) indicates that there is a significant difference between the relative frequency of the five affixes. Pairwise comparisons using Tukey contrasts reveal, however, that only two of the differences are significant. The affix with the highest relative frequency, locative *in-*, is significantly different from the two affixes with the lowest relative frequency, *un-* ($t\text{-}value$ = 4.813, p < 0.001) and *-ly* ($t\text{-}value$ = 4.489, p < 0.001).

One reason for the insignificance of the contrasts between the affixes might be that there are not enough observations for each affix to reach significance. Furthermore, the differences in relative frequency between the affixes are quite small. This is partly due to the gradient nature of the variable logRELATIVEFRE-QUENCY. In combination, the size of the data set and the gradient nature of the

variable might lead to a lack of statistical power. To alleviate this problem, I re-coded the variable logRELATIVEFREQUENCY into a categorical one. This especially makes sense considering that the variable has a natural threshold. For all types with a relative frequency below 0, the derivative is less frequent than its base. For all items having a relative frequency higher than 0, the opposite is the case. Therefore, items with a relative frequency below 0 were recoded as being *more decomposable*, and items with a relative frequency higher than 0 were recoded as *less decomposable* (see Hay 2001; Collie 2008 for a similar coding of relative frequency). Table 6.7 displays the distribution across affixes. Pairwise compar-isons revealed that *-ly* has significantly more more decomposable words than negative *in-*, locative *in-* and *dis-*. Furthermore, locative *in-* has significantly less more decomposable words than *un-* and negative *in-*.

To sum up the comparisons of the affixes' relative frequency, *un-* and *-ly* are the most segmentable affixes, locative *in-* is the least segmentable affix. The af-fixes *dis-* and negative *in-* seem to be in the middle of the scale. However, in both analyses (gradient and categorical relative frequency), only few contrasts proved to be significant.

Table 6.7: Categorical relative frequency by affix

CAT. RELATIVE FREQUENCY	*in*-$_\mathrm{LOC}$		*dis-*		*in*-$_\mathrm{NEG}$		*un-*		*-ly*	
less decomposable	38	(70%)	30	(48%)	9	(31%)	35	(35%)	35	(23%)
more decomposable	16	(30%)	32	(52%)	20	(69%)	66	(65%)	115	(77%)

Turning to the last measure of decomposability, the comparison of the affixes' LSA scores reveals that there are only minor differences between the affixes. Fig-ure 6.5 displays these differences using boxplots. On the x-axis the LSA score is displayed. Each box represents the distribution of the LSA scores for one affix in increasing order. Locative *in-* has the lowest score indicating that words fea-turing this affix are the least semantically similar to their base words. In terms of semantic similarity, they can thus be said to be the least decomposable words in the data set. Words with the affix *-ly* have the highest score, and are thus the most similar to their base word. They are the most decomposable. However, the figure shows that there is a big overlap in the distribution of the LSA scores across affixes, i.e. the differences between affixes are rather small. Statistical anal-yses confirm this impression. While an ANOVA ($F = 8.512, p < 0.001$) showed a significant effect of the affix on the LSA score, the pair-wise comparison of
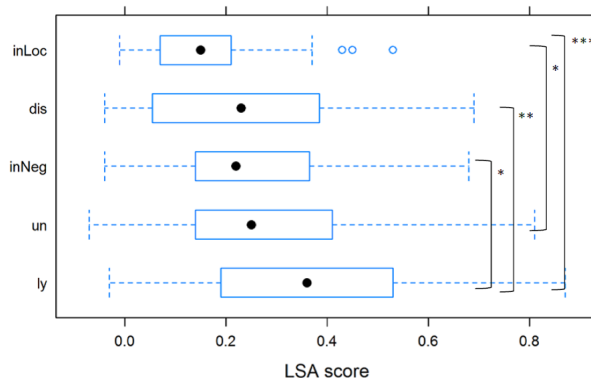
Figure 6.5: Comparison of LSA score of the five affixes

the means using Tukey contrasts shows that only four of the contrasts are significant. The affix *-ly* has significantly higher LSA scores than the affixes *un-* ($t\text{-}value$ = 0.027, p = 0.01), locative *in-* ($t\text{-}value$ = 5.197, p < 0.001) and *dis-* ($t\text{-}value$ = 3.443, p = 0.006). Furthermore, the affix *un-* has significantly higher scores than the affix locative *in-* ($t\text{-}value$ = 2.772, p = 0.044).

When interpreting the results for the variable LSAScore, it should be remembered that for some affixes not all items were taken into consideration in the comparisons. While for most of the *-ly-* and *un*-affixed words the LSA score could be computed (118 types for *-ly*, 89 types for *un-*), for the other affixes fewer types were considered in this comparison (between 23 and 40). Therefore, it is not surprising that we only find significant contrasts with *-ly* and *un-*, but not with the other affixes.

One can summarize that despite the fact that for a lot of types no LSA score could be computed, and the comparison of affixes in terms of their LSA score is thus based on a rather small number of observations, the distribution of the variable LSAScore reveals the same picture as the other decomposability variables. Locative *in-* is the least segmentable affix, followed by *dis-* and negative *in-*. The affixes *un-* and *-ly* are the most segmentable affixes.

After having looked at each decomposability measure individually, let us now take a look at the whole picture. Table 6.8 summarizes the results of the comparisons by showing a segmentability hierarchy for each decomposability measure. The segmentability hierarchies rank the affixes from the most segmentable to the least segmentable. The table shows very similar hierarchies for all measures.

Locative *in-* is the least segmentable affix, followed by *dis-* and negative *in-*. The affixes *un-* and *-ly* are the most segmentable affixes out of the five.

Table 6.8: Segmentability hierarchies for each decomposability measure

Decomposability measure								
SEMANTICTRANSPARENCY	{*un-*	,	*-ly*}	>	*in-*$_{\text{Neg}}$	>	*dis-*	>	*in-*$_{\text{Loc}}$
SEMANTICTRANSPARENCYRATING	*un-*	>	*-ly*	>	*in-*$_{\text{Neg}}$	>	*dis-*	>	*in-*$_{\text{Loc}}$
TYPEOFBASE	{*un-*	,	*-ly*}	>	{*in-*$_{\text{Neg}}$	,	*dis-*}	>	*in-*$_{\text{Loc}}$
logRELATIVEFREQUENCY	*-ly*	>	*un-*	>	*in-*$_{\text{Neg}}$	>	*dis-*	>	*in-*$_{\text{Loc}}$
CAT.RELATIVEFREQUENCY	*-ly*	>	*un-*	>	*in-*$_{\text{Neg}}$	>	*dis-*	>	*in-*$_{\text{Loc}}$
LSASCORE	*-ly*	>	*un-*	>	*in-*$_{\text{Neg}}$	>	*dis-*	>	*in-*$_{\text{Loc}}$

The differences between the hierarchies mostly concern the ranking of *un-* and *-ly*. In terms of LSASCORE and logRELATIVEFREQUENCY, *-ly* is more segmentable than *un-*, in terms of SEMANTICTRANSPARENCYRATING, *un-* is more segmentable than *-ly*, and in terms of SEMANTICTRANSPARENCY and TYPEOFBASE, there is no difference between the two. While all hierarchies display a very similar picture, it is to note that they differ with regard to how significant differences between affixes are. While for SEMANTICTRANSPARENCY and TYPEOFBASE all differences were significant, this was not the case for the other measures of decomposability.

The results of the segmentability comparison fit in with previous research on the segmentability of prefixes. Zirkel (2010), for example, investigated the parsability of 15 prefixes in terms of four different measures (productivity, type parsing ratio, token parsing ratio and average boundary strength). The comparison of the prefixes revealed that out of the 15 investigated prefixes, *un-* is the third most segmentable, and *dis-* is the 10th most segmentable. In other words, as in the present study, the prefix *un-* is very segmentable and the prefix *dis-* is far less segmentable. The prefix *in-* was not investigated in Zirkel (2010).

Overall, the segmentability hierarchies displayed in Table 6.8 match the Non-Semantic Segmentability Hierarchy. Thus, the Non-Semantic Segmentability Hierarchy is borne out by the data. This supports its validity and thus shows that the predictions made by the affix-specific Morphological Segmentability Approach are valid.

6.2.3 Summary

In the first part of this section the relation between the five measures of decomposability SemanticTransparency, SemanticTransparencyRating, TypeOfBase, logRelativeFrequency and LSAScore was investigated. The cluster analyses revealed that the three variables SemanticTransparency, SemanticTransparencyRating and TypeOfBase are highly correlated. The two gradient measures logRelativeFrequency and LSAScore did not correlate with any other decomposability variable. One can thus summarize that, while SemanticTransparency, SemanticTransparencyRating and TypeOfBase tap into the same phenomenon, and are therefore well suited to be used as operationalizations of the same concept, this cannot be said for logRelativeFrequency and LSAScore. It was argued that the results can be interpreted as evidence for three different types of decomposability. The first type is operationalized by SemanticTransparency, SemanticTransparencyRating and TypeOfBase, the second by logRelativeFrequency, and the third by LSAScore. Potential effects of the decomposability variables must be interpreted with these different types of decomposability in mind.

In the second part of the section, the distributions of the decomposability variables across affixes were compared. All comparisons revealed the same picture. Locative *in-* is the least segmentable affix, *un-* and *-ly* are the most segmentable. The prefixes negative *in-* and *dis-* pattern in between. Even though not all contrasts proved to be significant, and the categorical measures seemed to provide a clearer picture than the gradient measures, one can say that the decline in segmentability from *un-* and *-ly* to locative *in-* is supported by all measures. This order resembles the Non-Semantic Segmentability Hierarchy (cf. Figure 6.3). The corpus data thus empirically verifies one of the proposed segmentability hierarchies. In turn, the gemination prediction of the affix-specific Segmentability Approach, which is based on the Non-Semantic Segmentability Hierarchy, is valid.

6.3 Duration

6.3.1 Analyses

The first durational analysis consists of investigating the distribution of consonant duration in each subset to test whether gemination is a categorical or a gradient phenomenon (cf. 'Nature of gemination: Predictions' in §4.3.1). To see whether the distributions significantly differ between environments, I generated boxplots for each environment of each subset and applied standard test statis-

tics. If doubles show a significantly higher mean than singletons, and the box-plots indicate that the distributions of doubles and singles differ significantly, the data shows a bimodal distribution. In that case, one can assume gemination to be categorical. If there is no significant difference in the distribution of the environments, two explanations are possible. The first one is that the morphological geminates in the given data set degeminate, i.e. there is no difference in the duration of doubles and singles in the data set. The other possibility is that gemination is a gradient phenomenon which is not traceable by merely looking at distributions of durations and comparing averages. To check which of the two explanations holds, further statistical models are needed, i.e. linear regression models.

I fitted at least two linear regression models for each subset, one predicting absolute consonant duration in milliseconds (AbsoluteConsonantDuration) and one predicting relative consonant duration (RelativeConsonantDuration)▉ Relative consonant duration refers to the duration of the consonant relative to the duration of the preceding vowel. In addition to the models for each subset, one model directly comparing the three prefixes with nasals, i.e. *un-*, locative *in-* and negative *in-*, was fitted. While this model has the advantage of directly comparing the three prefixes with each other, it also faces several problems, such as the systematic difference in the distribution of variables between the prefixes. As will be discussed in the pertinent section, these problems limit the usefulness of the model to a great degree, e.g. some variables cannot be investigated in the model.

In the relative duration models, the independent variable PrecedingSegmentDuration was not included. This was because relative duration is computed by means of preceding segment duration. In other words, the variable PrecedingSegmentDuration is part of the dependent variable and therefore not a suitable predictor variable.

With regard to the decomposability measures, only in the *in-* and *dis-*models all five decomposability variables were included. In the *un-* and *-ly*-models, only the effects of logRelativeFrequency and LSAScore were tested. The reason is the distribution of the variables in the different subsets (see §6.2.2 for discussion). Since the variable LSAScore was not coded for all items, i.e. the inclusion of the variable results in a loss of a lot of data points, the effect of this variable was tested individually in all subsets.

In the models predicting consonant duration with *in-* and *dis-*, collinearity problems had to be addressed. As discussed in the previous section, the three decomposability variables SemanticTransparency, SemanticTransparency-

Rating and TypeOfBase highly correlate, and it was thus problematic to test all of them simultaneously in the model. Therefore, the effect of these variables was tested by including them individually in the model, and by conducting principal component analyses (cf. §5.4).

The use of mixed effects models was precluded by the data's unnestedness. Almost every item is produced by a different speaker and many items occur only once in the corpus, so that it did not make sense to use speaker and item as random effects (see also discussion in §6.1.1). All models were fitted according to the modeling strategy described in §5.4.

All models were tested for two types of interactions. First, I tested for interactions which are predicted to affect gemination according to the theoretical approaches discussed in Chapter 4. These are the interactions between the variable Environment and the decomposability variables, and the interaction between Environment and Affix. Then, I tested for interactions which, based on previous empirical work and theoretical considerations, can be assumed to affect affixational consonant duration. In other words, I tested for interactions between variables for which one can assume that their relationship leads to a situation in which the simultaneous influence of the variables on affixational consonant duration is not additive. One of those interactions is, for example, the interaction between Environment and BaseInitialStress. For prefixed words, it can be assumed that base-initial stress affects the duration of a double consonant differently than the duration of a singleton. The reason is that, as discussed in §2.3, part of the double consonant belongs to the base-initial syllable, while the singleton only belongs to the prefix. Some interactions were not testable because some level combinations were not attested. An example of such an untestable interaction is the interaction between Environment and BaseInitialStress in the *un*-data set. There are no *un*-prefixed words with a double consonant and an unstressed base-initial syllable. All interactions tested in the corpus study are listed in appendix C.

After fitting the regression models for each subset, I used multi-model inferencing to verify the results, i.e. to ensure that multi-model inferencing predicts the same variables to be important for predicting consonant duration as the final regression model (see §5.4 for a discussion of multi-model inferencing).

The plots of the regression models were generated with the `visreg` package (Breheny & Burchett 2015). For a plot showing the effect of a variable, all other variables are held constant at the median (for numeric variables) or at the most common category (for factors). For the models predicting absolute consonant duration, the plots always show the response variable AbsoluteNasalDuration
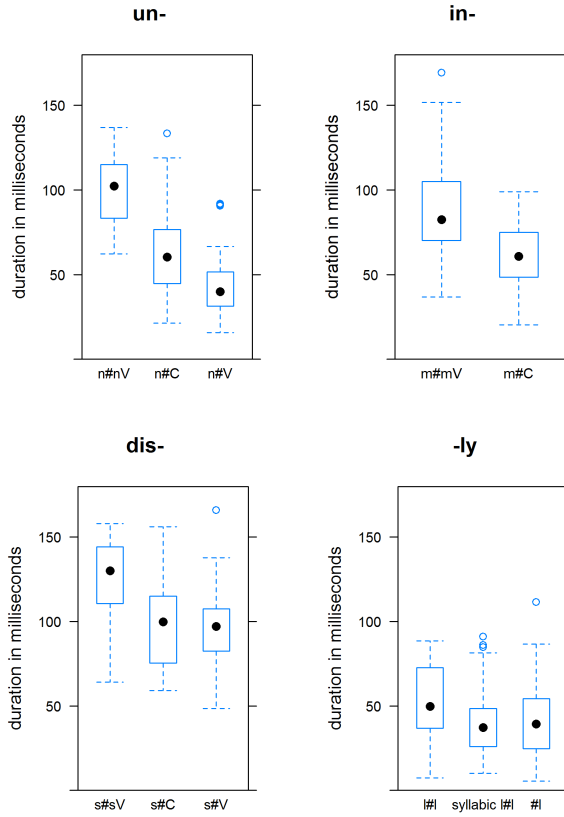
Figure 6.6: Distribution of consonant duration in the four data sets

in milliseconds, i.e. in cases where the dependent variable had to be transformed in the modeling process, the plots show the back-transformed variable.

6.3.2 Overview

Figure 6.6 depicts the distribution of consonant duration for each environment in each subset using boxplots. The distribution in the *un*-data set is shown in the upper left panel, the one for *in-* in the upper right panel, the one for *dis-* in the lower left panel, and the one for *-ly* in the lower right panel. The y-axis of each plot displays the duration of the consonant in milliseconds. The leftmost box in each plot represents the distribution for items with a double consonant, and the boxes on the right show the distributions for items with a single consonant.

For *-ly*, the box in the middle of the panel shows the distribution of the syllabic doubles.

For the prefixes, the figure shows a clear difference in duration between double and single consonants. Doubles are longer than singletons. The figure also shows that there is a clear binary distribution in the *un-*, *in-* and *dis*-data sets. For *in-* and *dis-*, there is hardly any overlap in the interquartile range of doubles and singletons. For *un-*, there is no overlap at all. For the prefixes, the durational distribution thus suggests a clear categorical difference between doubles and singletons. In other words, the plots suggest that the prefixes geminate, and that gemination is a categorical phenomenon. The *-ly*-data set does not show a bimodal distribution, i.e. there is a big overlap in the distribution of doubles and singletons, and doubles are not significantly longer than singletons in this data set. One might thus suspect degemination for *-ly*. Further statistical analyses are, however, necessary to confirm this impression.

Table 6.9 shows a summary of the distribution of consonant duration for each environment in the four subsets. Overall the durations of the consonants in the data set are in the same range as those found in other studies. For example, Umeda (1977: Tables II and X) finds in her North American English data that intervocalic word-internal singleton /n/ is between 34 and 38 ms long (depending on stress). Double /n/s across a word boundary have a duration of 100 ms. For singleton intervocalic word-medial /m/, Umeda finds mean durations between 70 and 74 ms, and for singleton /s/, mean durations range from 90 to 120 ms in that position. Singleton /l/ is between 40 and 47 ms long in Umeda's study. This indicates that the data and the durational measurements are valid.

Let us now turn to the distribution of duration for single and double consonants in the data, i.e. to the question of gemination. As already shown in Figure 6.6, for the three prefixes the mean and median duration of the double consonants is much higher than the mean and median duration of the single consonants. The differences in mean range from 26 ms (m#mV to m#C) to 55 ms (n#nV to n#V). For *-ly*, there is a difference in mean duration of 8 ms between double /l/ (l#l) and single /l/ (#l). The double is thus only slightly longer than the singleton.

To get a first idea about the nature of these differences between environments, some univariate analyses were carried out. The pair-wise comparison of the means for *un-* using Tukey contrasts yields significant contrasts for all three pairs, i.e. the differences between all environments are significant (see Table 6.10). This suggests gemination for *un-*. For *in-*, a Welch t-test shows a significant difference between the two environments m#mV and m#C ($t(152.98) = 7.1122$, $p < 0.001$). Thus, as in the *un-*data set, in the *in-*data set doubles are significantly longer than

Table 6.9: Duration of consonant(s) in milliseconds per environment for all affixes

un-

Environment	Example	Mean	Median	Standard Deviation
n#nV	*unnatural*	100	102	21
n#C	*untold*	64	60	24
n#V	*uneven*	45	40	18
Overall		60	54	28

in-

Environment	Example	Mean	Median	Standard Deviation
m#mV	*immortal*	87	81	27
m#C	*impossible*	61	61	19
Overall		76	74	27

dis-

Environment	Example	Mean	Median	Standard Deviation
s#sV	*dissatisfied*	127	130	35
s#C	*disgrace*	100	100	29
s#V	*disarm*	95	96	22
Overall		103	102	30

-ly

Environment	Example	Mean	Median	Standard Deviation
l#l	*really*	50	50	23
syllabic l#l	*ment(a)lly*	41	37	21
#l	*possibly*	42	39	21
Overall		43	41	22

singletons. For *dis-*, the comparison of the means also yields significant contrasts for the difference between doubles and singletons (see Table 6.11). However, the difference between the two singleton levels is not significant. None of the differences between the *-ly*-environments proved to be significant. Thus, while for all the prefixes, doubles are significantly longer than singletons, this is not the case for the suffix *-ly*.

Table 6.10: Multiple comparison of means of nasal duration for *un*-prefixed words (Tukey contrasts)

	Estimate	Std. Error	*t*-value	Pr(>|t|)
n#C - n#nV	-35	5	-6.903	< 0.001
n#V - n#nV	-56	5	-10.798	< 0.001
n#C - n#V	-20	4	-5.487	< 0.001

Table 6.11: Multiple comparison of means of consonant duration for *dis*-prefixed words (Tukey contrasts)

	Estimate	Std. Error	*t*-value	Pr(>|t|)
s#C - s#sV	- 28	7	-6.903	0.001
s#V - s#sV	-32	7	-10.798	< 0.001
s#C - s#V	5	5	-0.858	0.666

All in all, the first analyses of durations have shown that the singleton consonant durations in the data are similar to the ones found in previous studies and the data is thus valid. Furthermore, the investigation of distributions suggests that the investigated prefixes geminate and that gemination is categorical. It also suggests degemination for *-ly*. However, since it is well known that the duration of segments in natural speech is subject to a variety of different influences, more advanced statistical analyses are necessary to investigate the matter. In the next subsections I will present such analyses for each subset.

6.3.3 The prefix *un*-

6.3.3.1 Absolute duration

The linear model predicting absolute duration with *un*- was fitted according to the procedure described in sections 5.4 and 6.3.1. The residuals of the initial model showed a non-normal distribution. Therefore, the dependent variable ABSOLUT-CONSONANTDURATION was transformed by the Box-Cox-transformation parameter 0.303, and outliers were removed. The removal of outliers resulted in the loss of 2 observations, i.e. 1.3% of the observations. After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix C for a list of all tested interactions). None of the interactions was significant, and only two significant predictors remained in the final model, ENVIRONMENT and LOCALSPEECHRATE. The model explains 57% of the variance found in the data. Table 6.12 documents the estimates for each predictor and their p-values in the final model.

Table 6.12: Summary of linear model for variables predicting the Box-Cox-transformed duration of [n] in *un*-prefixed words

	Estimate	Std. Error	t-value	p-value
Intercept	0.580	0.015	38.502	< 0.001
ENVIRONMENT-n#C	-0.050	0.010	-5.072	< 0.001
ENVIRONMENT-n#V	-0.097	0.010	-9.770	< 0.001
LOCALSPEECHRATE	-0.008	0.001	-6.814	< 0.001

Adjusted R-squared: 0.562

Figure 6.7 depicts the effect of LOCALSPEECHRATE. The y-axis of the graph displays the duration of the nasal in milliseconds, the horizontal axis represents the local speech rate. The line represents the estimated effect of the variable. The shaded areas in the graphs represent the 95% confidence intervals. The plot shows, the higher the speech rate, i.e. the more segments are pronounced in a given amount of time, the shorter becomes the nasal. This is an expected effect.

Let us now turn to our variable of interest, ENVIRONMENT. Its effect is shown in Figure 6.8. The blue lines in the figure represent the estimated consonant duration for each of the three investigated environments. The graph shows that
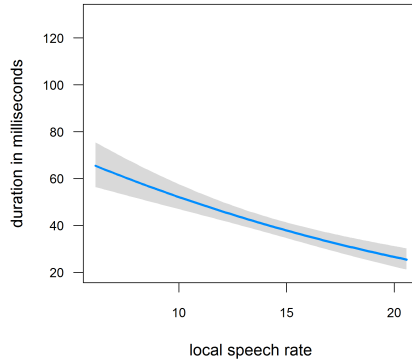
Figure 6.7: Effect of local speech rate on consonant duration in *un*-data set

words containing a double nasal have a significantly longer duration than words with one nasal, no matter whether the single nasal is followed by a non-nasal consonant or by a vowel. In the case of a following vowel, the single /n/ is shortest.
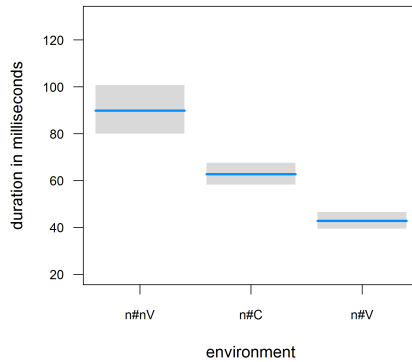
Figure 6.8: Effect of environment on consonant duration in *un*-data set

The predicted mean duration for double nasals is 90 ms. For words with the n#C environment the nasal is predicted to be 63 ms long, and for words having the n#V

environment it is predicted to be 43 ms long. If we compare the two environments with a following vowel (and thus hold the type of following segment constant), the model predicts double nasals to be even a bit longer than twice the duration of the average single nasal in this environment (90 ms as against 43 ms). This result clearly speaks in favor of gemination with *un-*.

When a consonant follows the single nasal at the morpheme boundary, we also find a highly significant contrast between the two environments, but the difference is smaller. We do not find twice the duration for the double nasal, but only a difference of 27 ms, i.e. an increase in duration of 43% from single to double nasal.

The question may be raised whether this increase in phonetic duration can be interpreted as gemination in spite of the fact that the duration is not doubled. The literature on gemination has shown, however, that the durational differences between geminates and their corresponding singletons may vary substantially (see §2.1 for discussion). For /n/ in English, differences between 34% and 109% were found. For word boundary geminates, Delattre (1969) finds an increase from singleton to geminate /n/ of 50%. All investigated environments in Delattre's study were, however, vowel-initial. For normal speech, Oh & Redford (2012: 86, Figure 2) arrive at an estimated 82 ms for word-internal singletons, 110 ms for *un*-geminates and geminates across word boundaries. This is an increase of 60%. For careful speech, they find estimated durations of 110 ms for singletons, 225 ms for *un*-geminates, and 230 ms for geminates across word boundaries. This is an increase of 104% to 109%.[4] Note that again, only vocalic environments were tested.

The comparison to previous findings on gemination with English /n/ shows that there is good reason to interpret even the smaller of the two contrasts in the data (i.e. the one between n#C vs. n#nV) as evidence for gemination. While some studies have found bigger singleton-geminate ratios, some found smaller ratios. Differences in experimental set-up, especially speech condition, and environment might be the cause of the different ratios found. In this study, the smaller singleton-geminate ratio between pre-consonantal singletons and doubles (as against the ratio of pre-vocalic singletons and doubles) can be attributed to the type of following segment (C vs. V). As discussed in §5.5.1, following consonants generally lead to longer durations for nasals.

After fitting the linear model, multi-model inferencing was used to detect which of the variables included in the initial model are the most predictive vari-

---

[4]Oh & Redford (2012) do not give the estimated means in their article. The figures given here are read off from the partial effects plot given in Figure 2 of their article.

ables across a multitude of models. The variable LSASCORE was not included in the multi-model inferencing analysis because of the low number of items which were coded for LSASCORE. The analysis revealed that the two most important variables are those which also ended up in the final linear model, i.e. ENVIRON-MENT (importance value: 1) and LOCALSPEECHRATE (importance value: 1). The importance values of the other variables are much lower (PRECEDINGSEGMENT-DURATION: 0.4, BASEINITIALSTRESS: 0.29, logWORDFORMFREQUENCY: 0.27, logRE-LATIVEFREQUENCY: 0.33). This indicates that these variables are far less predictive of nasal duration than ENVIRONMENT and LOCALSPEECHRATE.

6.3.3.2 Relative duration

The model predicting relative consonant duration with *un-* was fitted similarly to the model predicting absolute consonant duration and the same interactions were tested. Again the dependent variable was transformed using Box-Cox-transformation to achieve a normal distribution of residuals. The transformation parameter was 0.141. In contrast to the absolute duration model, no outliers had to be excluded. Table 6.13 summarizes the final model. The model explains about 33% of the variance and features only one variable, ENVIRONMENT.

Table 6.13: Summary of linear model for variables predicting the Box-Cox-transformed relative duration of [n] in *un*-prefixed words

	Estimate	Std. Error	t-value	p-value
Intercept	1.029	0.013	80.326	< 0.001
ENVIRONMENT-n#C	-0.072	0.015	-4.858	< 0.001
ENVIRONMENT-n#V	-0.127	0.015	-8.527	< 0.001

Adjusted R-squared: 0.327

The model reveals that, as in absolute duration, double /n/ is significantly longer than both singleton /n/s in relative duration. It thus confirms the results of the absolute duration model, i.e. *un-* geminates. In contrast to the absolute duration model, the variable LOCALSPEECHRATE is not significant in this model. This is not surprising, as the dependent variable RELATIVECONSONANTDURATION does not measure duration per se, but the relation of consonant duration and preceding vowel duration, i.e. a ratio. While speech rate is known to affect absolute

duration, an effect on relative duration is not expected.

Multi-model inferencing confirms the final model. The variable ENVIRONMENT is the most predictive variable (importance value: 1). All other variables have very low importance values (logWORDFORMFREQUENCY: 0.56, BASEINITIALSTRESS: 0.46, LOCALSPEECHRATE: 0.26, logRELATIVEFREQUENCY: 0.26).

6.3.3.3 Summary

Two models were fitted to predict consonant duration with *un-*. The model predicting absolute consonant duration explains more variance than the one predicting relative consonant duration, i.e. the absolute duration model is the better model. In the absolute duration model, the noise variable LOCALSPEECHRATE had the expected effect. In the relative duration model, no noise variables remained in the final model. Both models found a significant effect of the variable ENVIRONMENT. Phonological doubles (n#nV) are significantly longer than phonological singletons, irrespective of whether the following segment is a consonant (n#C) or a vowel (n#V). The durational difference between doubles and singletons is more pronounced when the singleton is followed by a vowel, i.e. a following consonant lengthens the prefixal /n/. The results clearly show that the prefix *un-* geminates. With regard to the other variables of interest, only LSASCORE and logRELATIVEFREQUENCY could be tested. Neither had a significant effect.

6.3.4 The prefix *in-*

6.3.4.1 Absolute duration

The initial *in*-model predicting absolute duration showed a non-normal distribution of residuals. Therefore, the dependent variable was transformed (Box-Cox-transformation parameter: 0.465). After the transformation, the model showed a satisfactory distribution of residuals. The model was then fitted according to the strategy described in §5.4 and interactions were tested (see appendix C for a list of all tested interactions). To avoid collinearity, the effects of the decomposability variables were tested individually.

The final model explains about 50% of the variance in the data and includes four variables with a significant effect on consonant duration, ENVIRONMENT, LOCALSPEECHRATE, BASEINITIALSTRESS and AFFIX. None of the tested interactions proved to be significant. An overview of the model coefficients is given in Table 6.14.

Figure 6.9 displays the effects of the two noise variables in the model. The

Table 6.14: Summary of linear model for variables predicting the Box-Cox-transformed duration of [m] in *in*-prefixed words

	Estimate	Std. Error	t-value	p-value
Intercept	0.368	0.015	24.931	< 0.001
Environment-m#C	-0.048	0.007	-6.662	< 0.001
LocalSpeechRate	-0.003	0.001	-4.335	< 0.001
BaseInitialStress-unstressed	-0.038	0.008	-4.826	< 0.001
Affix-inNeg	0.016	0.007	2.121	0.036

Adjusted R-squared: 0.504

left panel shows the effect of LocalSpeechRate. This effect is as expected: the higher the speech rate, the shorter the nasal. The right panel shows the effect of BaseInitialStress. With an estimated mean duration of 74 ms, the consonant is 21 ms shorter before an unstressed base-initial syllable than before a stressed base-initial syllable (95 ms). This result is expected, too. As mentioned in §5.5, Umeda (1977) also found that nasals before unstressed vowels are shorter than nasals before stressed vowels.
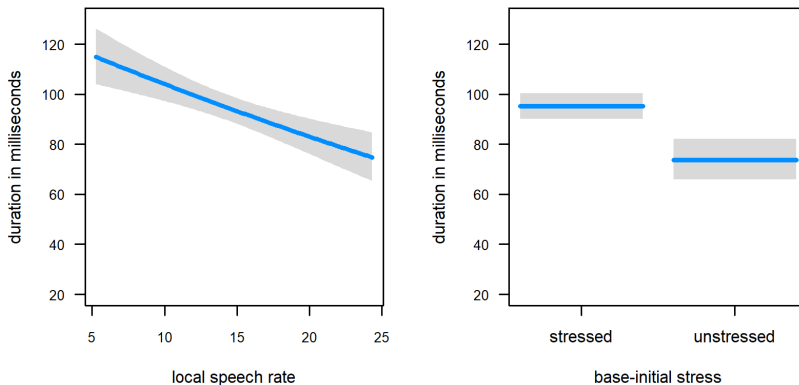
Figure 6.9: Effects of local speech rate and base initial-stress on consonant duration in *in*-data set

Let us turn to the variables of interest. The left panel of Figure 6.10 displays the effect of Environment. Double consonants are significantly longer than sin-

gletons. The estimated mean duration for double consonants is 95 ms, while it is 68 ms for single consonants, a difference of 27 ms. This difference is significant, and shows that *in-* geminates.
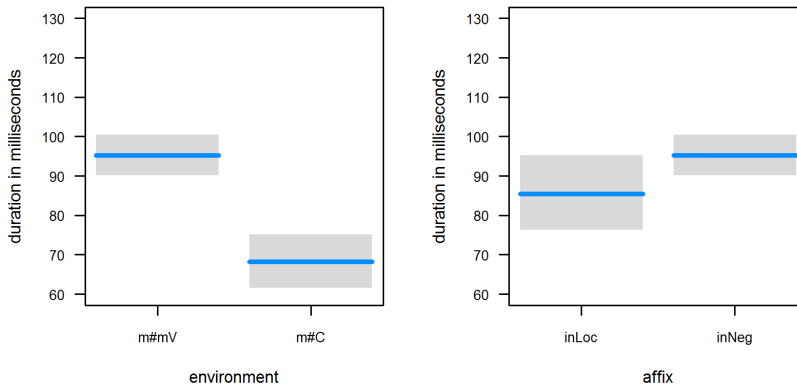
Figure 6.10: Effects of environment and affix on consonant duration in *in*-data set

However, one could venture the idea that the difference is not due to a difference between one consonant and two, but due to a difference in the following segment, i.e. consonant versus vowel. This idea is, however, unsupported, since following vowels lead to shorter durations of the nasal. This can also be seen in the *un*-data set, in which a single nasal preceding a consonant (63 ms) is longer than a single nasal preceding a vowel (43 ms). In other words, the double nasals (which are by their very nature followed by a vowel) are likely to be shortened, not lengthened, due to their vocalic environment. In other words, the double nasals show longer duration in spite of being in an environment that would trigger shorter duration. The significant difference between m#mV and m#C is thus a sure sign of gemination.

The effect of Affix is displayed in the right panel of Figure 6.10. The nasal in negative *in-* is significantly longer (by 10 ms) than the one in locative *in-*. Hence, there is a difference in the duration of the nasal depending on which of the two affixes is used. There was no interaction of Environment and Affix, which means that the two prefixes do not differ significantly in their gemination behavior.

None of the investigated decomposability measures proved to be significant. However, it is possible that, while the measures do not affect duration when

tested individually, a combined measure of decomposability has a significant effect on consonant duration. As evidenced by the decomposability analyses in §6.2, at least some of the decomposability variables highly correlate and can be assumed to measure the same underlying property. It is thus possible to test the effect of a combined measure of this underlying property.

To test the effect of a combined decomposability measure, an additional model with combined decomposability measures (as opposed to individual decomposability measures) was fitted. The combined measures were created by means of a principal component analysis (cf. §5.4 on principal component analyses). The principal component analysis was fitted with the variables logRELATIVE-FREQUENCY, SEMANTICTRANSPARENCY, SEMANTICTRANSPARENCYRATING, TYPEOF-BASE and AFFIX. The variable AFFIX was included because of the differences in segmentability between locative and negative *in-* (see §6.2.2 for discussion). LSASCORE was excluded because of the low number of observations coded for this variable. Categorical variables were recoded as numerical before entering the analysis, and all variables were scaled. Table 6.15 summarizes the analysis by showing the composition of each principal component, i.e. the loading of each variable for each principal component, and by displaying the proportion of variance covered by each component.

Table 6.15: Summary of principal components

Composition of principal components					
	PC1	PC2	PC3	PC4	PC5
scaledAFFIX	0.078	-0.817	-0.244	0.449	0.255
scaledRELATIVEFREQUENCY	-0.428	0.450	-0.530	0.574	0.054
scaledSEMANTICTRANSPARENCYRATING	-0.521	-0.233	0.450	0.269	-0.631
scaledTYPEOFBASE	-0.487	-0.275	-0.529	-0.623	-0.140
scaledSEMANTICTRANSPARENCY	-0.550	-0.002	0.420	-0.087	0.717
Variance explained by principal components					
	PC1	PC2	PC3	PC4	PC5
Proportion of Variance	0.551	0.264	0.086	0.064	0.035

The analysis revealed that the first two components can account for most of the variance expressed by the five variables (81%). An inspection of the rotation matrix shows that the first component is dominated by all four decomposability measures, and that the second is mainly dominated by the variable AFFIX. One can thus conclude that the first component represents a combined measure of

decomposability, and the second represents the variable Affix. Both components were included as predictor variables in the linear model.

The linear model with the principal components was fitted similarly to the model with the individual decomposability measures. After simplification, the model showed very similar effects as the model with the individual decomposability measures (see appendix D for model summary). The effects of Environment, LocalSpeechRate and BaseInitialStress are identical. Instead of the variable Affix, this model shows an effect of PC2. The effect is shown in Figure 6.11. The higher the value of PC2, the longer the duration of the nasal. As explained above, PC2 is dominated by the variable Affix. A higher PC2-value indicates negative *in-*, a lower PC2-value indicates locative *in-*. One can thus interpret the effect of PC2 as being an effect of Affix. Negative *in-* is longer than locative *in-*. The principal component model hence shows the same effects as the model with the individual decomposability measures.
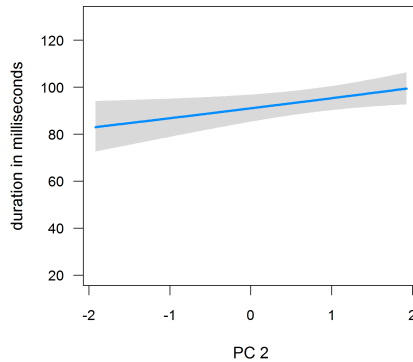
Figure 6.11: Effect of PC2 on consonant duration in *in*-data set

After the final models were fitted, multi-model inferencing was used to detect the most important predictor variables. Due to the collinearity problems with the decomposability variables, these variables were not included separately in the analysis. Instead, the combined decomposability measure from the principal component analysis, i.e. principal component 1 (PC1), was included. The analysis revealed that LocalSpeechRate, Environment and BaseInitialStress are the most predictive variables. They all have an importance value of 1. Affix has an importance value of 0.79, i.e. it is the fourth most important variable. With importance values of 0.29 (PC1), 0.28 (logWordFormFrequency) and 0.27

(PRECEDINGSEGMENTDURATION), the other variables are far less predictive of consonant duration. Multi-model-inferencing thus confirms the results of the final models.

6.3.4.2 Relative duration

In the model predicting relative consonant duration with *in-*, the dependent variable was Box-Cox-transformed by the parameter −0.101 to achieve a normal distribution of the residuals. No outliers were excluded. The model was fitted similarly to the absolute duration model, i.e. the same interactions were tested and the decomposability measures were tested in the same way. On the one hand the decomposability variables were tested individually. On the other, decomposability was tested by means of principal components in an additional model.

The simplification of both models, i.e. the one with the individual measures and the one with the principal components, resulted in same final model. The model features the three predictor variables ENVIRONMENT, LOCALSPEECHRATE and BASEINITIALSTRESS. Neither the individual decomposability measures nor the principal components proved to be significant. There are no interactions. The summary of the final model is given in Table 6.16. The model explains about 42% of the variance in the data.

Table 6.16: Summary of linear model for variables predicting the Box-Cox-transformed relative duration of [m] in *in*-prefixed words

	Estimate	Std. Error	t-value	p-value
Intercept	0.987	0.015	65.285	<0.001
ENVIRONMENT-m#C	0.053	0.007	7.308	<0.001
LOCALSPEECHRATE	-0.003	0.001	-3.416	0.001
BASEINITIALSTRESS-unstressed	0.066	0.008	7.876	<0.001

Adjusted R-squared: 0.422

As in the model predicting absolute duration, this model reveals that double consonants are longer than singletons, i.e. we also find gemination in relative duration. Also similar to the absolute duration model, LOCALSPEECHRATE affects consonant duration. The higher the speech rate, the shorter the nasal relative to the vowel. This effect of speech rate indicates that speech rate has a bigger

effect on the consonant of the prefix *in-* than on its vowel, i.e. in faster speech the consonant is more reduced than the vowel. A possible explanation might be that the prefixal vowel in *in-* is too short to be reduced to the same degree as the prefixal nasal, i.e. there is simply less material to be reduced.

The model also reveals a significant effect of BASEINITIALSTRESS. In relative duration, the consonant is longer before unstressed syllables. This is the opposite of what is found for absolute duration, where the consonant is shorter in that condition. This difference between relative and absolute duration can be explained by the role of the preceding vowel for relative duration. Longer preceding vowels lead to shorter relative durations. One can thus assume that the shorter relative duration of the consonant in words with an unstressed base-initial syllable is caused by a longer preceding vowel in those words. Longer vowels before unstressed syllables are expected. This is because, in the investigated words, unstressed base-initial syllables indicate a stressed prefix, which in turn might influence the duration of the prefixal vowel. It is particularly the duration of the vowel which is lengthened in a stressed syllable, i.e. one might expect the vowel of a stressed prefix (which is followed by an unstressed syllable) to be lengthened. A longer vowel in turn leads to shorter relative duration. This explains that in relative duration the consonant is shorter before unstressed base-initial syllables than before stressed base-initial syllables. The absolute duration of the consonant, in contrast to its relative duration, is not affected by the duration of the preceding segment, i.e. by its stress status. Instead, the consonant participates in the stress-caused lengthening of its following syllable, i.e. the consonant is longer before stressed base-initial syllables.

In contrast to the absolute model, in the relative duration model AFFIX does not have a significant effect. This indicates that negative and locative *in-* only differ in absolute duration, i.e. the ratio of consonant and vowel duration does not differ between the two prefixes. As with absolute duration, no effect of the decomposability variables was found in the relative duration models.

Multi-model inferencing revealed that the three variables which are significant in the final model are the most predictive variables across a multitude of models. BASEINITIALSTRESS has an importance value of 1, ENVIRONMENT has an importance value of 0.99, and LOCALSPEECHRATE has an importance value of 0.99. The other tested variables are much less predictive of relative duration (importance value of logWORDFORMFREQUENCY: 0.54, importance value of AFFIX: 0.48, importance value of PC1: 0.37).

6.3.4.3 Summary

The two linear models predicting consonant duration with *in-* clearly show that locative and negative *in-* geminate. In both models, phonological doubles are longer than phonological singletons. The absolute duration model furthermore revealed that the nasal in negative *in-* is significantly longer than the one in locative *in-*, irrespective of whether it is a double consonant or a singleton. This effect of Affix was, however, not found in relative duration. With regard to the decomposability measure, none had an effect on nasal duration, neither in absolute, nor in relative duration. In both models, the two noise variables LocalSpeechRate and BaseInitialStress showed expected effects. None of the tested interactions was significant. The model predicting absolute consonant duration explains more of the variance in the data than the model predicting relative duration.

6.3.5 The prefixes *un-* and *in-*

The model predicting consonant duration with *un-* and *in-* was fitted to directly compare the durational behavior of the nasals in the three prefixes *un-*, negative *in-* and locative *in-*. The model has, however, the disadvantage that a number of interesting variables cannot be properly tested. The reason is that the *un-* data set and the *in-*data set differ in important respects. First, the prefix *un-* and the allomorph of *in-* that is being invested here end in two different consonants, i.e. /n/ vs. /m/. Therefore, durational differences between *un-* and *in-* are not directly comparable. I used scaling of the durational variables to alleviate this problem. This, however, means that durational differences are not straightforward in their interpretation. Second, the phonological environments of singleton *un-* and singleton *in-* are not the same, since /ɪm/ is necessarily always followed by a base-initial consonant, while *un-* is followed by both consonants and vowels. Only the double nasal in both prefixes is always followed by a vowel. Third, variables of decomposability cannot be tested in an interesting way. This is because *un-*, as described in §6.2.2, does not vary in most of the decomposability measures, and because relative frequency measures are not well comparable across *un-* and *in-*. The prefix *in-* has very many bound roots, which is problematic with regard to computing relative frequency measures that are comparable to the relative frequency measures of affixes with hardly any or no bound roots, i.e. in this case *un-*. With these limitations in mind, a regression model was fitted to the lumped data set.

Since the environments for the prefixal nasal are not the same across prefixes, I created a new variable in which I coded whether the word has one or two under-

lying nasals (NUMBEROFCONSONANTS), and an additional variable encoding whe-
ther a vowel or a consonant followed the nasal (FOLLOWINGSEGMENT). I included
the following predictors in the model: NUMBEROFCONSONANTS, FOLLOWINGSEG-
MENT, LOCALSPEECHRATE, BASEINITIALSTRESS, AFFIX, PRECEDINGSEGMENTDURA-
TION and logWORDFORMFREQUENCY. The model was then simplified and interac-
tions were tested. Crucially, all interactions between the variable AFFIX and all
other variables were tested (see appendix C for a list of all tested interactions).
The final model explains 49% of the variance and features five variables, NUMBER-
OFCONSONANTS, FOLLOWINGSEGMENT, LOCALSPEECHRATE, BASEINITIALSTRESS and
AFFIX. The model is documented in Table 6.17.

Table 6.17: Summary of linear model for variables predicting the nor-
malized duration of the nasal in *un-* and *in*-prefixed words

	Estimate	Std. Error	t-value	p-value
Intercept	2.484	0.223	11.127	< 0.001
NUMBEROFCONSONANTS-double	-1.454	0.144	-10.065	< 0.001
FOLLOWINGSEGMENT-vowel	-0.537	0.130	-4.136	< 0.001
LOCALSPEECHRATE	-0.088	0.012	-7.016	< 0.001
BASEINITIALSTRESS-unstressed	-0.347	0.103	-3.365	0.001
AFFIX-inLoc	-0.469	0.133	-3.521	0.001
AFFIX-un	0.343	0.123	2.794	0.006

Adjusted R-squared: 0.49

As in the *un-* and *in*-models, this model shows that doubles are longer than sin-
gletons. Interestingly, there is also a main effect of AFFIX. Figure 6.12 shows that
negative *in-* is significantly longer than locative *in-*, and that *un-* is significantly
longer than negative *in-*. We thus find a decline in duration from *un-*, to negative
in-, to locative *in-*. This decline is fully in line with the decline in segmentability
of the prefixes found in §6.2.2. Crucially, there is no significant interaction be-
tween AFFIX and NUMBEROFCONSONANTS, which means that all three prefixes
geminate.

In addition to the effects of the variables of interest, there is the expected effect
of LOCALSPEECHRATE (nasals become shorter with increasing speech rate) and
the expected effect of the FOLLOWINGSEGMENT (nasals are shorter before vowels).
There is also an effect of BASEINITIALSTRESS such that the nasal is shorter before
unstressed syllables.
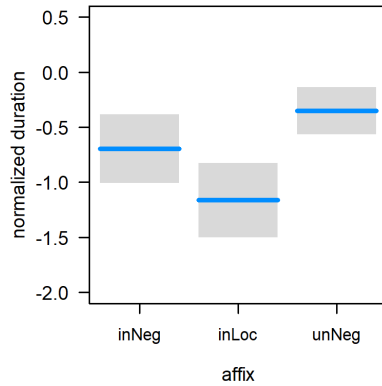
Figure 6.12: Effects of affix on consonant duration for words prefixed with *un-* and *in-*

6.3.6 The prefix *dis-*

6.3.6.1 Absolute duration

The residuals of the initial model predicting absolute consonant duration with *dis-* were not distributed normally. Therefore, the dependent variable was Box-Cox-transformed (parameter: 0.222). After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. No outliers were removed. The model was then simplified according to the strategy described in §5 and interactions were tested (see appendix C for a list of all tested interactions).

As in the *in-*model, testing the effect of the decomposability measures simultaneously was not possible due to collinearity problems. Therefore, their effect was tested individually, as well as by using principal components. In other words, first models were fitted in which the effects of the decomposability measures were tested individually, and then an additional model with a combined decomposability measure was fitted. Let us first discuss the model which initially included the individual measures.

An overview of the final model, i.e. the model after simplification, is given in Table 6.18. Insignificant estimates are printed in light gray. The model yields an adjusted R-squared of 0.333, i.e. it explains about 33% of the variance found in the data. The model includes four variables, ENVIRONMENT, LOCALSPEECHRATE,

Voicing and BaseInitialStress. There is a significant interaction between Environment and BaseInitialStress. None of the decomposability variables is significant.

Table 6.18: Summary of linear model for variables predicting the Box-Cox-transformed duration of [s] in *dis*-prefixed words

	Estimate	Std. Error	t-value	p-value
Intercept	0.644	0.015	43.659	< 0.001
Environment-s#C	-0.045	0.009	-5.176	< 0.001
Environment-s#V	-0.026	0.010	-2.506	0.014
LocalSpeechRate	-0.005	0.001	-4.595	< 0.001
Voicing-voiceless	0.057	0.011	5.367	< 0.001
BaseInitialStress-unstressed	-0.036	0.019	-1.851	0.066
Environment-s#C: BaseInitialStress-unstressed	0.048	0.023	2.068	0.041
Environment-s#V: BaseInitialStress-unstressed	0.017	0.021	0.782	0.436

Adjusted R-squared: 0.333

Figure 6.13 displays the effects of LocalSpeechRate and Voicing. The effect of LocalSpeechRate can be seen in the left panel. As expected, with increasing speech rate the fricative in *dis*-prefixed words becomes shorter. The right panel of the figure shows the effect of Voicing. Voiced fricatives are significantly shorter than voiceless fricatives. For doubles this difference is predicted to be 47 ms, for singletons voiced fricatives are predicted to be 36 ms shorter than voiceless fricatives. With regard to Voicing, it is important to note that the distribution of voiced items in the data set is unbalanced. All voiced *dis*-prefixed words are semantically opaque, followed by a vowel and have a stressed base-initial syllable. This skewed distribution might have influenced the final model, i.e. it might have skewed the effects of the affected variables. This problem will be discussed in further detail below.

The interaction between Environment and BaseInitialStress is displayed in Figure 6.14. The left panel of the figure displays the estimated effect of Environment on words with a stressed base-initial syllable. The right panel shows the estimated effect of Environment on words with an unstressed base-initial syl-
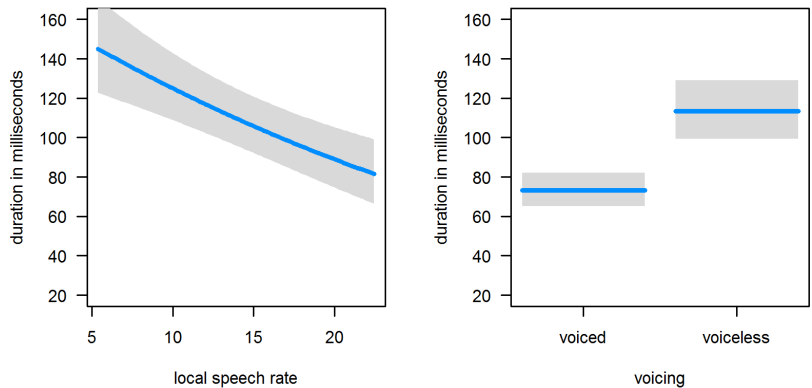
Figure 6.13: Effect of local speech rate and voicing on consonant duration in *dis*-data set
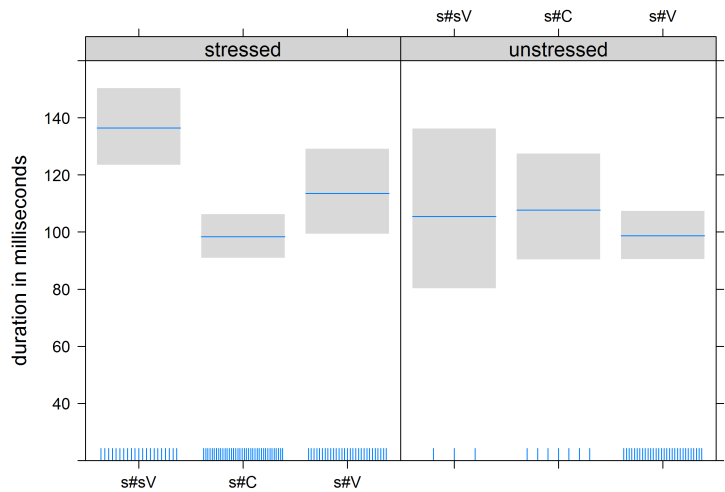
Figure 6.14: Effect of environment by base-initial stress on consonant duration in *dis*-data set

lable. The figure shows that while there is a significant effect of Environment for words with a stressed base-initial syllable, there is no significant effect for words with an unstressed base-initial syllable. For words with a stressed base-initial syllable, doubles are estimated to be 38 ms longer than singletons which are followed by a consonant, and 23 ms longer than singletons which are followed by a vowel. The difference in duration between the two singleton levels is marginally significant (p-value: 0.066). The fricative before a consonant is predicted to be 15ms shorter than the fricative before a vowel. There is no difference between the three environments for words with an unstressed base-initial syllable, i.e. doubles are of the same duration as singletons, and the two singleton levels are also of the same duration.

One could interpret the interaction between Environment and BaseInitial-Stress as evidence that the prefix *dis-* only geminates when followed by a stressed syllable, and that only in this condition the following segment, i.e. consonant or vowel, affects the duration of the prefixal consonant. However, there are two problems with the pertinent interaction, and the interpretation of the effects is therefore not that straightforward. Both problems are related to the distribution of variables in the data set.

The first problem is the number of types and tokens for each category. The blue lines on the bottom of the plot (in Figure 6.14), the rugs, indicate the number of tokens for each category. It is striking that there are only three tokens with an unstressed base-initial syllable and a double consonant at the morphological boundary. These three tokens are all of the same type, i.e. *dissolution*. The interaction found in the model is hence caused by only three observations (of one type). In other words, there are only three tokens with a double consonant which do not geminate. All three tokens are of the type *dissolution*.

One might argue that the word *dissolution* behaves differently because the word is simplex. There are two arguments for this assumption. First, *dis-* does not carry any meaning in the word *dissolution*. Second, in the morphologically related derivative *dissolve* the prefixal fricative is voiced. As discussed in §3.1.1, there are claims that voiced /s/ is only found in simplex *dis*-words. That the word *dissolve* features a voiced fricative might thus indicate that it is lexicalized, which in turn might also mean that the related word *dissolution* is lexicalized and no longer complex.

The second problem with the interaction is related to the distribution of the variable Voicing. As already mentioned above, all voiced *dis*-prefixed words are semantically opaque, followed by a vowel and have a stressed base-initial syllable. The significant effect of Voicing in the final model is thus solely based on

opaque, vowel-adjacent items with stressed base-initial syllables. Even though the effect is based on only a few items with a particular combination of features, the model accounts for Voicing in predicting consonant duration for all items of all levels. In other words, the model allocates the effect of Voicing to all levels irrespective of whether voiced items are actually present in the pertinent level. This might distort the effects of the affected variables, i.e. SemanticTransparency, Environment and BaseInitialStress. With regard to the final model, the interaction between Environment and BaseInitialStress might be affected by this problem. Only two of the six pertinent categories feature voiced items, i.e. only words with base-initial stress and a double consonant and words with base-initial stress and a singleton followed by a vowel can be voiced. All other categories only feature voiceless items.

To test whether the effects of Environment and BaseInitialSyllable are affected by the distribution of voiced items in the data set, and to also test the effect of SemanticTransparency without the possible harming influence of the variable Voicing, an additional model was fitted to the *dis*-data set. In this model only voiceless *dis*-prefixed words were included. The model included 104 observations, i.e. 24 voiced *dis*-prefixed words were excluded.

Table 6.19: Summary of linear model for variables predicting the Box-Cox-transformed duration of [s] in *dis*-prefixed words with voiceless /s/

	Estimate	Std. Error	t-value	p-value
Intercept	0.716	0.017	41.928	< 0.001
Environment-s#C	-0.050	0.009	-5.607	< 0.001
Environment-s#V	-0.052	0.011	-4.918	< 0.001
SemanticTransparency-transparent	-0.004	0.008	-0.523	0.602
SpeechRate	-0.005	0.001	-5.088	< 0.001
BaseInitialStress-unstressed	-0.046	0.017	-2.748	0.007
SemanticTransparency-transparent: BaseInitialStress - unstressed	0.052	0.020	2.646	0.010

Adjusted R-squared: 0.375

The model predicting consonant duration with voiceless *dis*-prefixed words was fitted similarly to the model with all items. Table 6.19 shows the final model for the voiceless fricatives. It explains about 38% of the variance found in the data. As in the model with all *dis*-prefixed words, LocalSpeechRate has the expected

effect on duration, and we find an effect of SMALL-CAPS:BASEINITIALSTRESS and ENVIRON-MENT. However, in contrast to the complete model, in this model BASEINITIAL-STRESS and ENVIRONMENT are not interacting. Figure 6.15 shows the effect of ENVIRONMENT. One clearly sees that the double is predicted to be longer than both singleton consonants, irrespective of base-initial stress. Furthermore, the durational difference between the two singleton levels is not significant in this model. One can hence assume that the marginally significant difference between the two singleton levels in the complete model was caused by the uneven distribution of voiced items in the complete *dis*-data set.
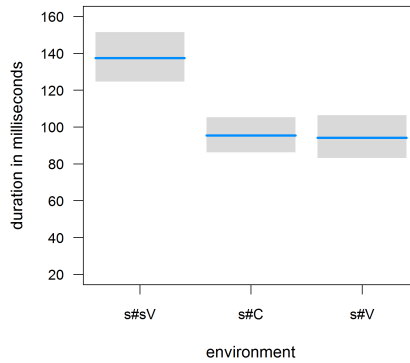
Figure 6.15: Effects of environment on consonant duration in voiceless *dis*-data set

The voiceless model shows an interaction between the two variables SEMANTIC-TRANSPARENCY and BASEINITIALSTRESS. This interaction can be seen in Figure 6.16. The left panel shows the effect of SEMANTICTRANSPARENCY for items with a stressed base-initial syllable, the right panel shows the effect for items with an unstressed base-initial syllable. When the base-initial syllable is stressed, there is no difference between opaque and transparent items. When the base-initial syllable is unstressed, the fricative in opaque items is predicted to be 42 ms shorter than the fricative in transparent items. However, as in the complete model one must be cautious to interpret this interaction. There are only four opaque tokens with an unstressed base-initial syllable in the data set. These four tokens cause the interaction, i.e. in these four tokens the fricative is significantly shorter than in all other tokens. Crucially, three of the four tokens are of the type *dissolution*. This means that, even though we find two different interactions in the two *dis*-models
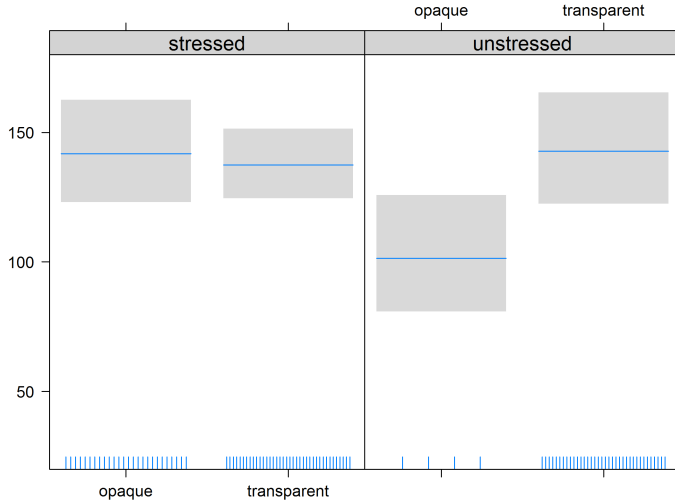
Figure 6.16: Effects of semantic transparency by base-initial stress on consonant duration in voiceless *dis*-data set

(BASEINITIALSTRESS and ENVIRONMENT in the complete model vs. BASEINITIAL-STRESS and SEMANTICTRANSPARENCY in the voiceless model), the same tokens which cause the interaction in the complete model cause the interaction in the model with only voiceless items. It remains unclear whether the short fricative in the pertinent words is due to the unstressed base-initial syllable in combination with semantic opacity, or whether the pertinent words behave differently because the stress status of the base-initial syllable is crucial for gemination with *dis*-, i.e. only words with a stressed base-initial syllable geminate, or whether the shorter fricative duration is caused by type-specific effects.

After fitting a model with the individual decomposability measures, I fitted a model with combined decomposability measures. The combined measures were created by means of a principal component analysis (cf. §5.4 on principal component analyses). The principal component analysis included the four variables logRELATIVEFREQUENCY, SEMANTICTRANSPARENCYRATING, TYPEOFBASE and SEMANTICTRANSPARENCY. As with *in*-, the variable LSASCORE was not included in the principal component analysis as it would have led to an extreme reduction of the data set. Categorical variables were recoded as numerical before they entered the analysis, and all variables were scaled.

Table 6.20 shows a summary of the principal components. The first princi-

pal component accounts for most of the variance and is composed more or less equally of all measures. The second component is dominated by TYPEOFBASE and logRELATIVEFREQUENCY, the third component mostly represents SEMANTIC-TRANSPARENCY, TYPEOFBASE and logRELATIVEFREQUENCY, and the fourth is mostly composed of SEMANTICTRANSPARENCY and SEMANTICTRANSPARENCYRATING. The second and the third component explain much less variance than the first, and the last principal component explains barely any variance. The first three principal components were included in the model.

After model simplification none of the principal components remained in the model. The final model resembles the final model of the complete data set. This means that the four variables ENVIRONMENT, LOCALSPEECHRATE, VOICING and BASEINITIALSTRESS proved to be significant, and that ENVIRONMENT and BASEINITIALSTRESS form an interaction in the model. Decomposability did not affect fricative duration.

Table 6.20: Summary of principal components

Composition of principal components				
	PC1	PC2	PC3	PC4
scaledRELATIVEFREQUENCY	0.449	0.653	0.609	0.036
scaledSEMANTICTRANSPARENCYRATING	0.563	-0.094	-0.269	-0.776
scaledTYPEOFBASE	0.441	-0.735	0.447	0.255
scaledSEMANTICTRANSPARENCY	0.535	0.158	-0.598	0.576
Variance explained by principal components				
	PC1	PC2	PC3	PC4
Proportion of Variance	0.679	0.174	0.105	0.042

Two multi-model inferencing analyses were conducted for *dis-*, one for the complete data set, one for the voiceless data set. As with *in-*, the principal components were used to test the predictive value of the decomposability variables. In other words, to avoid collinearity problems, the principal components were used in the analyses instead of the individual decomposability measures.

The models revealed that LOCALSPEECHRATE and ENVIRONMENT are the most important variables in both models (importance value: 1). In the complete model, VOICING was also of high importance (importance value: 1). For both data sets, BASEINITIALSTRESS proved to be the next most important variable. However,

its importance value is much lower than the ones of Environment, Local-SpeechRate and Voicing (0.47 in complete model, 0.32 in voiceless model). All other variables were much less important in both models. The importance values for the complete model are 0.37 for logWordFormFrequency, 0.26 for PrecedingSegmentDuration, 0.29 for PC2, 0.25 for PC1, and 0.25 for PC3; the importance values for the voiceless model are 0.30 for PC3, 0.28 for logWordFormFrequency, 0.28 for PC1, 0.27 for PC2 and 0.30 for PrecedingSegmentDuration.

6.3.6.2 Relative duration

In the model predicting relative duration with *dis*-, the residuals showed a normal distribution and therefore no transformation of the dependent variable was necessary. The model was fitted similarly to the absolute duration model, i.e. decomposability measures were tested individually, as well as in terms of principal components, and the same interactions were tested. Let us first discuss the model with the individual decomposability measures.

After model simplification, the following four variables remained in the final model: Environment, SemanticTransparency, Voicing and BaseInitialStress. With an adjusted R-squared of 0.262 the model explains less of the variance than the absolute duration model. The model summary is printed in Table 6.21.

Table 6.21: Summary of linear model for variables predicting the relative duration of [s] in *dis*-prefixed words

	Estimate	Std. Error	t-value	p-value
Intercept	1.937	0.231	8.384	< 0.001
Environment-s#C	-0.838	0.203	-4.134	< 0.001
Environment-s#V	-0.643	0.208	-3.089	0.002
SemanticTransparency-transparent	-0.390	0.194	-2.015	0.046
Voicing-voiceless	1.187	0.281	4.225	< 0.001
BaseInitialStress-unstressed	-0.370	0.185	-1.998	0.048

Adjusted R-squared: 0.262

The variable Voicing shows the same effect as in the absolute duration model. Voiced fricatives are shorter than voiceless fricatives. The effects of Environment and BaseInitialStress are also similar to the ones found in the absolute

duration model. Doubles are longer than singletons, and consonants are shorter before unstressed base-initial syllables than before stressed base-initial syllables. In contrast to the absolute duration model, the relative duration model does not show an interaction between ENVIRONMENT and BASEINITIALSTRESS, or between SEMANTICTRANSPARENCY and BASEINITIALSTRESS. This means that, in contrast to the absolute duration model, all *dis*-prefixed types geminate in terms of their relative duration. Furthermore, all prefixal fricatives are affected by an unstressed base-initial syllable, i.e. not just opaque words or words with a double consonant.

The effect of BASEINITIALSTRESS on relative duration with *dis*- deviates from the effect of BASEINITIALSTRESS on relative duration with *in*-. For *in*-, we find shorter relative durations before stressed syllables than before unstressed syllables. This effect was explained by reference to the preceding vowel duration in *in*-. There are different possibilities for the deviating results between *in*- and *dis*-. It might for example be that inherent durational differences between the consonants, i.e. /m/ vs. /s/, led to different results. Another possibility is that preceding vowel durations differ between *in*- and *dis*-.

The final relative duration model shows a significant effect of SEMANTICTRANS-PARENCY. Transparent items are predicted to have shorter fricatives than opaque items. This is an unexpected effect, which is not found in absolute duration, and one must be cautious to interpret the effect. As discussed above, there is a close and complex relation between the variables VOICING and SEMANTICTRANSPARENCY. This relation might have caused the surprising result. To see whether this is the case, an additional model with only voiceless items was fitted. The model is displayed in Table 6.22. With an R-squared of 0.204, the model explains 20% of the data. This is less than any other model fitted to the *dis*-data set. Only one variable remained significant in the model, ENVIRONMENT. The fricative in s#sV-structures is longer than the fricative in s#C- and s#V-structures. In other words, the model again shows that *dis*- geminates. Crucially, SEMANTICTRANSPARENCY does not reach significance in the model. One can therefore conclude that the effect found in the complete model is probably caused by the distribution of voicing in the data set. In other words, the effect is not stable and therefore not trustworthy.

After fitting the model with the individual decomposability measures, the model with the principal components was fitted. After simplification, the principal component model showed similar effects as the model with the individual decomposability measures (see Table D.1 in appendix D for model summary). The effects of ENVIRONMENT and VOICING are identical. However, instead of the variables BASEINITIALSTRESS and SEMANTICTRANSPARENCY, this model includes the variable

Table 6.22: Summary of linear model for variables predicting the relative duration of [s] in *dis*-prefixed words with voiceless /s/

	Estimate	Std. Error	t-value	p-value
Intercept	2.878	0.172	16.692	< 0.001
Environment-s#C	-0.823	0.207	-3.970	< 0.001
Environment-s#V	-1.123	0.212	-5.295	< 0.001

Adjusted R-squared: 0.204

PC1. The effect of PC1 is shown in Figure 6.17.
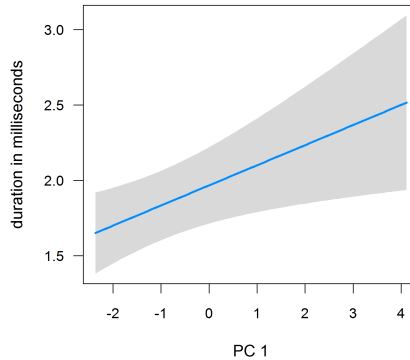
Figure 6.17: Effect of PC1 on relative consonant duration in complete *dis*-data set

The higher the value of PC1, the longer the relative duration of the fricative. As explained above, PC1 is composed of all decomposability measures. The higher the PC1-value, the less decomposable a word is. One can thus interpret the effect of PC1 in the following way: the less decomposable a derivative, the longer the relative duration of /s/. This is unexpected with regard to the decomposability predictions. It is, however, somewhat expected with regard to the results from the complete relative duration model, in which semantically transparent words are predicted to have shorter fricatives than semantically opaque words. As discussed above, this effect might be caused by the distribution of certain proper-

ties among types and tokens in the subset, and one must therefore be cautious with the interpretation of the effect. All in all, one can state that the principal component model shows similar effects as the model with the individual decomposability measures, and that the interpretation of the decomposability effect is yet unclear.

As for absolute duration, two multi-model inferencing analyses were conducted to identify the most important variables for predicting relative duration with *dis-*, one for the complete data set and one for the voiceless data set. As in the absolute duration models, ENVIRONMENT is clearly the most important variable in both models (importance value: 1). In the complete model, VOICING is also of very high importance (importance value: 0.99). For both data sets, the first principal component (PC1) reaches a rather high importance value (complete data set: 0.85, voiceless data set: 0.71). The importance values of all other variables are much lower. BASEINITIALSTRESS has an importance value of 0.56 in the complete data set and one of 0.41 in the voiceless data set, PC2 has an importance value of 0.47 in the complete data set and one of 0.40 in the voiceless data set, PC3 has an importance value of 0.27 in the complete data set and one of 0.28 in the voiceless data set, logWORDFORMFREQUENCY has an importance value of 0.27 in the complete data set and one of 0.25 in the voiceless data set, and LOCALSPEECHRATE has an importance value of 0.26 in the complete data set and one of 0.24 in the voiceless data set.

6.3.6.3 Summary

The models predicting absolute duration with *dis-* and the models predicting relative duration with *dis-* reveal similar effects. In general, all models feature the same set of variables, and there are only small differences between the models. The models predicting relative duration feature fewer significant variables, and only in absolute duration models interactions were found to be significant. The absolute duration models feature a higher R-squared than the relative duration models. They thus explain more of the variance in the data than the relative duration models. This finding is similar to what was found for *un-* and *in-*.

With regard to the noise variables, LOCALSPEECHRATE, VOICING and BASEINITIALSTRESS had a significant effect on consonant duration in the models. In absolute duration, LOCALSPEECHRATE influenced fricative duration in the expected direction, i.e. the higher the speech rate, the shorter the fricative. LOCALSPEECHRATE did not affect the relative duration of the fricative. An effect of VOICING was found in the two models containing all *dis-*prefixed items, i.e. voiced and voiceless items. Voiced items are shorter than voiceless items in absolute and relative

duration. The effect of BASEINITIALSTRESS is less clear. While in relative duration, an unstressed base-initial syllable leads to the shortening of the fricative in all *dis*-prefixed words, in the absolute duration models the variable forms interactions. In the model with only voiceless items BASEINITIALSTRESS interacts with SEMANTICTRANSPARENCY, in the complete model it interacts with ENVIRONMENT.

Both interactions with the variable BASEINITIALSTRESS are caused by only few tokens in the data set, three in the complete model and four in the voiceless model. Three of those tokens are of the type *dissolution*, which is the only type in the data set which is opaque, features a double consonant and has an unstressed base-initial syllable. In the voiceless model, we find that in opaque items with an unstressed base-initial syllable, i.e. in the three tokens of the type *dissolution* and in the one token of the type *discount*, consonants are shorter than in all other items. In the complete model, we find that doubles in words with an unstressed base-initial syllable, i.e. doubles in the type *dissolution*, are shorter than in all other double consonant items. They are as long as singletons. This means that while all other double consonants geminate, the double consonant in these words is not significantly longer than a singleton.

The two absolute duration models thus predict particularly short fricative durations for the same tokens, in particular tokens of the type *dissolution*. The models differ, however, in how they explain the shortness of the consonant. While in the complete model the short double consonant is explained by the following unstressed syllable, in the voiceless model it is explained by the combination of an unstressed base-initial syllable with semantic opacity. It remains unclear whether the short duration of /s/ in *dissolution* is caused by its unstressed base-initial syllable, i.e. whether all *dis*-prefixed words degeminate when followed by an unstressed syllable, or whether it is caused by the combination of an unstressed base-initial syllable and semantic opacity, or whether this is a type-specific effect.

All in all, the analyses suggest that, in general, *dis-* geminates. Except for the complete absolute duration model, in which one type with a double consonant degeminates, i.e. *dissolution*, all models show a robust effect of the variable ENVIRONMENT. In both absolute and relative duration, doubles are significantly longer than singletons. There is no significant difference in duration between the two singleton levels. Furthermore, multi-model inferencing supports the claim that *dis-* geminates by showing the importance of the variable ENVIRONMENT for all models. One can thus summarize that the corpus study shows that in the majority of cases, the prefix *dis-* geminates. The only *dis*-prefixed word in the data set which might degeminate is the word *dissolution*.

With regard to the other variables of interest, i.e. the decomposability measures, only SemanticTransparency showed a significant effect in absolute duration. The effect was only significant in one model and formed an interaction with BaseInitialStress. As discussed above, the interaction is caused by only a few items and therefore difficult to interpret. In the relative duration models, effects of SemanticTranspaency and PC1 were found. Due to the distribution of certain properties in the data set, these effects are, however, also difficult to interpret.

6.3.7 The suffix *-ly*

6.3.7.1 Absolute duration

After the exclusion of two outliers, the *-ly*-model showed a satisfactory distribution of residuals, and an initial model was fitted. With regard to decomposability, only the two decomposability measures LSAScore and logRelativeFrequency were tested in the model. As described in §6.2.2, in the *-ly*-data set the variables SemanticTransparency, TypeOfBase and SemanticTransparencyRating did not show any, or barely any, variation. Therefore, their effect on consonant duration in *-ly* could not be tested. The variable PrecedingSegmentDuration was also discarded. This was because of its high correlation with the variable PrecedingSegment, which would have caused collinearity problems in the model. In preliminary analyses, the variable PrecedingSegment proved to be the better predictor. Therefore, this variable was used in the model and not PrecedingSegmentDuration.

The model for the *ly*-data set was simplified according to the same procedure as the previous models. A list of all tested interactions in the model can be found in appendix C. The final model is summarized in Table 6.23.

The model explains about 24% of the variance and features four variables of which three are significant. The three noise variables LocalSpeechRate, PrecedingSegment and logWordFormFrequency are significant. The variable Environment does not show a significant effect but remained in the model as it is the crucial variable with regard to gemination.

Figure 6.18 displays the effects of LocalSpeechRate and PrecedingSegment. The left panel of the figure shows that with increasing speech rate /l/ becomes shorter, and the right panel shows that after a vowel /l/ is longer than after a consonant. These effects are expected.

Figure 6.19 shows the effects of logWordFormFrequency and Environment. The left panel of the figure shows that with increasing frequency /l/ becomes

Table 6.23: Summary of linear model for variables predicting the duration of [l] in *-ly*-suffixed words

	Estimate	Std. Error	t-value	p-value
Intercept	77.323	8.438	9.163	<0.001
ENVIRONMENT-syllabic l#l	7.387	5.641	1.309	0.192
ENVIRONMENT-#l	3.965	4.761	0.833	0.406
LOCALSPEECHRATE	-2.335	0.438	-5.332	<0.001
PRECEDINGSEGMENT - vowel	15.694	4.429	3.544	0.001
logWORDFORMFREQUENCY	-1.631	0.777	-2.100	0.038

Adjusted R-squared: 0.236
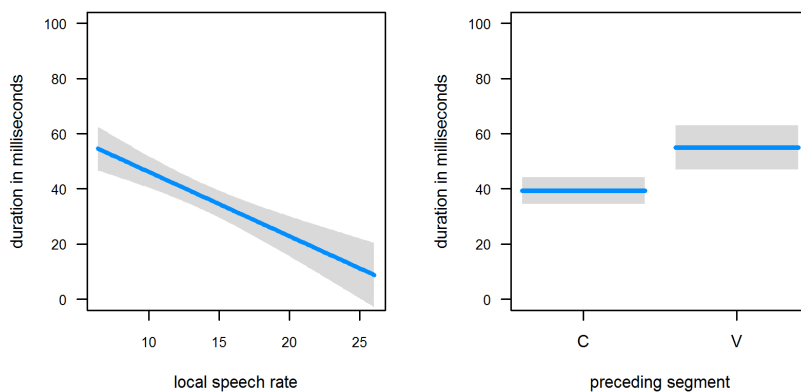
Figure 6.18: Effects of local speech rate and preceding segment on consonant duration in *-ly*-data set

shorter. This is expected. The right panel of the figure shows the insignificant effect of ENVIRONMENT on consonant duration in -*ly*. The figure clearly shows that there is no durational difference between double /l/ (l#l), syllabic double /l/ (syllabic l#l) and singleton /l/ (#l). The suffix -*ly* thus clearly degeminates in absolute duration.
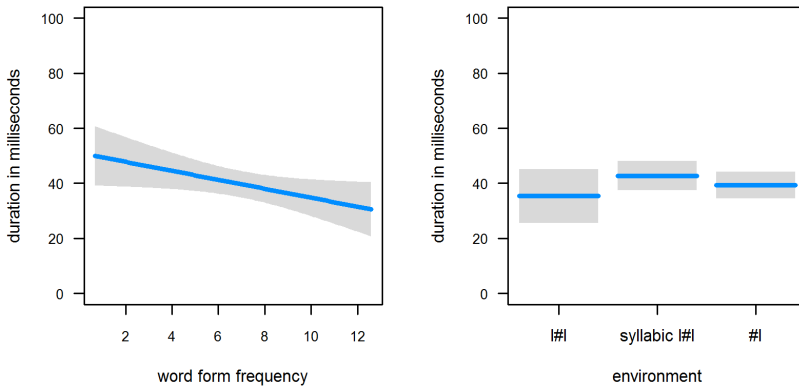
Figure 6.19: Effects of word form frequency and environment on consonant duration in -*ly*-data set

To see which are the most predictive variables for consonant duration in the -*ly*-data set, I conducted a multi-model inferencing analysis. As for the other data sets, the analyses did not include the variable LSASCORE because of the low number of observations which were coded for LSASCORE. The analysis revealed that LOCALSPEECHRATE is the most important variable (importance value: 1). With an importance value of 0.99 the variable PRECEDINGSEGMENT is the second most important value. It is followed by the two frequency variables logRELATIVEFREQUEN-CY (importance value: 0.67) and logWORDFORMFREQUENCY (importance value: 0.64). With importance values of 0.32 (ENVIRONMENT) and 0.26 (BASEFINALSTRESS), the other variables are far less important for predicting consonant duration with -*ly*.

6.3.7.2 Relative duration

To test whether the affix -*ly* geminates in terms of relative duration, it was necessary to create a subset. This subset only includes words which feature a vowel

at the end of their base, i.e. words in which the suffix is preceded by a vowel. This was necessary because relative duration (as a measure of gemination) is calculated by dividing consonant duration by preceding vowel duration. Since the complete data set includes words which are preceded by a consonant, the creation of the subset was necessary. However, the subset only includes 48 tokens, i.e. it is very small. Furthermore, in this data set only the difference in duration between two of the pertinent environments can be tested, i.e. l#l and #l. The items featuring a syllabic double (`syllabic l#l`) had to be excluded since they are always preceded by a consonant.

The small size of the data set restricted the number of noise variables included in the model, i.e. to avoid overfitting, only few variables were tested in the model. I chose to only include those noise variables which showed a significant effect in the absolute duration model, i.e. LOCALSPEECHRATE and logWORDFORMFREQUENCY. The variable PRECEDINGSEGMENT was not included because all of the preceding segments in the data set were vowels. The model was then fitted according to the same procedure as the absolute duration model. The residuals were normally distributed, and thus no transformation of the dependent variable was necessary and no items had to be excluded.

The final model reached an adjusted R-squared of 0.184, i.e. it explains 18% of the variance in the data. As with the other affixes, the relative duration model is thus worse than the absolute duration model. Only one variable remained in the final model, logWORDFORMFREQUENCY. The higher the frequency, the shorter the relative duration of the consonant ($estimate = -0.129$, $t\text{-}value = -3.533$, $p = 0.001$). In contrast to the absolute duration model, LOCALSPEECHRATE did not reach significance. Crucially, ENVIRONMENT did not reach significance. Doubles are not longer than singletons in terms of relative duration, i.e. they degeminate. Because of the low number of observations in the subset, no multi-model inferencing analysis was conducted.

6.3.7.3 Summary

Two models predicting consonant duration with *-ly* were fitted, one predicting absolute consonant duration and one predicting relative consonant duration. The usefulness of the relative duration model is, however, restricted by the small number of observations which were taken into account in this model. In both models only noise variables showed significant effects. In the absolute duration model, the expected effects of LOCALSPEECHRATE, PRECEDINGSEGMENT and logWORDFORMFREQUENCY were found. In the relative duration model only logWORDFORMFREQUENCY had a significant effect on consonant duration. Cru-

cially, the variable SMALLCAPS ENVIRONMENT did not have a significant effect on consonant duration, neither in absolute nor in relative duration. Multi-model inferencing also showed that this variable is not an important predictor for consonant duration with -*ly*. Thus, the suffix -*ly* clearly degeminates. The decomposability variables did not prove to be significant in either model.

6.3.8 Summary

The first durational analyses looked at the distribution of duration across environments to get a first impression of whether the affixes under investigation geminate, and if so, whether gemination is a gradient or a categorical phenomenon. For the prefixes, the analyses revealed that the distribution of duration in the data sets is bimodal, with the double consonants being significantly longer than the singletons. This indicates that the prefixes geminate, and that gemination is categorical. The distribution for -*ly* is not bimodal, and doubles are not longer than singletons. This suggests degemination with -*ly*. The impression that the prefixes geminate and that the suffix degeminates was validated in the linear models.

For all data sets at least two linear models were fitted, one predicting absolute consonant duration and one predicting relative consonant duration. In general, both models reveal very similar results, i.e. for the most part the same variables are significant in both models. However, the relative models feature fewer variables than the absolute models. Importantly, the effects of the variables of interest do not differ between the models, i.e. the same variables of interest are significant in both models. The relative duration models only feature less noise variables. The variable SMALLCAPS ENVIRONMENT shows stronger effects in absolute than in relative duration. Furthermore, for all affixes, the absolute duration models explain more of the variance in the data than the relative duration models. The results thus suggest that absolute consonant duration is a better measure of gemination than relative consonant duration. In other words, gemination in English is not expressed by a particular consonant-vowel ratio but by absolute consonant duration. In what follows I will therefore concentrate on absolute consonant duration.

Table 6.24 shows an overview of the variables which show significant effects on absolute consonant duration in the subsets. Only variables which are significant in at least one of the absolute duration models are listed. For *dis*-, the table indicates all variables which are significant in at least one of the two absolute duration models fitted, i.e. the model including all *dis*-prefixed items and the model including only voiceless items.

Let us first discuss the noise variables. The variable SMALLCAPS LOCALSPEECHRATE shows

Table 6.24: Overview of significant variables in absolute duration models

Variable	un-	im-	un- & im-	dis-	-ly
ENVIRONMENT	✓	✓	✓	✓	n.s.
AFFIX	-	✓	✓	-	-
SEMANTICTRANSPARENCY	-	n.s.	-	(✓)	-
LOCALSPEECHRATE	✓	✓	✓	✓	✓
BASEINITIALSTRESS	n.s.	✓	✓	✓	-
VOICING	-	-	-	✓	-
PRECEDINGSEGMENT	-	-	-	-	✓
logWORDFORMFREQUENCY	n.s.	n.s.	n.s.	n.s.	✓

✓ : significant in at least one of the absolute duration models
n.s. : not significant in any of the absolute duration models
- : not included in the absolute duration models

the expected effect in all models. The higher the speech rate, the shorter the duration of the consonant. For *dis-*, the variable VOICING is significant in the complete model. As expected, voiced items are shorter than voiceless items. The variable is irrelevant for all other models.

BASEINITIALSTRESS affects consonant duration in the *in-*model and in the *un-* and *in-* model. Before unstressed base-initial syllables the consonant is shorter than before stressed base-initial syllables. In the *dis-*models, BASEINITIALSTRESS forms interactions with ENVIRONMENT and SEMANTICTRANSPARENCY. As discussed■ thoroughly in §6.3.6, these interactions were caused by only a few items and their interpretation is therefore unclear. Either only *dis-*prefixed items with a double consonant are affected by BASEINITIALSTRESS, or only semantically opaque items are affected by stress.

BASEINITIALSTRESS does not affect consonant duration with *un-*. The reason for the absence of the effect might be due to the distribution of BASEINITIAL-STRESS across *un-*prefixed words. Most *un-*words feature a stressed base-initial syllable. This includes all items with a double consonant. This lack of variation in stress might have caused the absence of the effect with *un-*. It is also possible

that other factors, such as prefixal stress, might have interfered with the effect of BaseInitialStress. As discussed in Chapter 3, the stress status of *un-* is yet unclear. Further research on prefixal stress is necessary to explore the relation of stress and prefixal consonant duration further. This is, however, beyond the scope of this study.

For the suffix *-ly*, the two variables PrecedingSegment and logWordFormFrequency show the expected effects. The consonant is shorter after a consonant than after a vowel, and more frequent words display shorter consonant durations than less frequent words. While PrecedingSegment was only coded for *-ly*, the variable logWordFormFrequency was included in all models. Only for *-ly* it affected consonant duration. There are various possible explanations for this. One possibility is that word form frequency only affects consonant duration in suffixes but not in prefixes. Another possibility is that, while the effect exists for all affixes, it only reaches significance in the *-ly*-model. There are two arguments for this explanation. First, the *-ly*-model is the worst of all models, i.e. it explains the least variance. It is therefore easier for weaker effects to reach significance. Second, out of all data sets the *-ly*-data set features the highest number of different types. Different types have different frequencies, i.e. the distribution of logWordFormFrequency shows more variation in the *-ly*-data set than in the other data sets. As already discussed above, the lack of variation in the distribution of a variable might make it impossible to find effects. In other words, there might not be enough variation in logWordFormFrequency in the prefixal data sets to find effects.

Let us now turn to the variables of interest. Out of the five decomposability measures only SemanticTransparency showed an effect on consonant duration. This effect was, however, only found in one model, i.e. the model predicting absolute consonant duration with voiceless *dis-*. Furthermore, the effect was only found in interaction with the variable BaseInitialStress. Since the interaction was caused by only a few items in the data set, its validity is yet unclear. In none of the other models decomposability directly played a role. However, in two of the models decomposability indirectly influenced duration. As discussed in §6.2, the five affixes differ in their overall segmentability. This difference is mirrored in consonant duration in the *in-* and in the *un-* and *in-* model. In these models, the variable Affix significantly affects consonant duration. The affix *un-* features the longest nasal, negative *in-* features a significantly shorter nasal, and locative *in-* has the shortest nasal out of the three. This decline in duration resembles the decline in segmentability of the affixes. The prefix *un-* is the most segmentable affix of the three, followed by negative *in-*, followed by locative *in-*. Thus, while

word-specific decomposability measures did not affect consonant duration in the models, decomposability influenced consonant duration in terms of the overall segmentability of the affix.

The variable ENVIRONMENT significantly affected consonant duration in all prefixal models. For *-ly*, there was no significant difference in duration between the three tested environments. Table 6.25 summarizes the predicted durations for all environments for all affixes. The table also shows the significant differences in predicted durations between doubles and singletons for the prefixes, as well as singleton-geminate ratios. The conditions of the predicted values are the same as in the partial effects plots, i.e. numerical variables are held constant at their median and categorical variables are held constant at the most common category. For *dis-*, two predicted values are shown, one for the complete data set (for words with a stressed base-initial syllable), and one for the voiceless data set.

Table 6.25: Overview of predicted durations in the corpus studies

Predicted durations

	Phonological Doubles		Phonological Singletons	
	(non-syllabic)	(syllabic)	(consonant-adjacent)	(vowel-adjacent)
un-	90 ms	NA	63 ms	43 ms
im-	95 ms	NA	68 ms	NA
dis-[1]	136 ms	NA	98 ms	113 ms
dis-[2]	142 ms	NA	98 ms	97 ms
-ly	35 ms	43 ms	39 ms	

Durational difference   **Singleton–Geminate Ratio**

	Double-Singleton (consonant-adjacent)	Double-Singleton (vowel-adjacent)	Singleton-Double (consonant-adjacent)	Singleton-Double (vowel-adjacent)
un-	27 ms	47 ms	1:1.4	1:2.1
im-	27 ms	NA	1:1.4	NA
dis-[1]	38 ms	23 ms	1:1.4	1:1.2
dis-[2]	44 ms	45 ms	1:1.5	1:1.5

[1] stressed base-initial syllable
[2] voiceless

For *un-*, *in-* and *dis-*, double consonants are significantly longer than corresponding singletons. The only exception might be the double consonant in the

word *dissolution.* In the *dis*-model with all words, i.e. voiced and voiceless *dis*-prefixed words, a significant interaction between Environment and BaseInitialStress was found. Double consonant items with an unstressed base-initial syllable, i.e. words of the type *dissolution*, degeminated. However, it is yet unclear whether this effect is universal, i.e. whether it applies to all *dis*-prefixed words with a double consonant and an unstressed base-initial syllable, or whether the effect is caused by other factors, such as the word's opacity or type-specific effects. In Chapter 7, we will turn back to this question after considering the results of the experimental study.

The durational differences between doubles and singletons range from 23 ms to 47 ms. The singleton-double ratio for singletons followed by a consonant is comparable across all prefixes. The ratio is 1:1.4 for *un*-, *in*-, and *dis*- (in the complete data set). For voiceless *dis*- the ratio is 1:1.5. For singletons followed by a vowel, the absolute durational difference between doubles and singletons is practically the same for *un*- and voiceless *dis*- (47 ms and 45 ms), but smaller for voiced and voiceless *dis*- (23 ms). The absolute durational differences for *dis*- and *un*- suggest that both prefixes geminate to the same degree. However, the double-singleton ratios suggest the opposite. The ratio for *un*- is much bigger than the ones for *dis*-. This indicates that *un*- geminates to a higher degree than *dis*-.

In addition to the difference between doubles and singletons, a significant difference between the two singleton levels was found for *un*-. Consonants followed by a vowel are significantly shorter than consonants followed by a consonant. This difference between the two singleton environments was not found for *dis*-. For *-ly* the variable Environment never reached significance, i.e. singletons, non-syllabic doubles and syllabic doubles are predicted to be of the same duration.

A comparison of the singleton-double ratios found in this study with the ones in former studies is limited. Only two studies have investigated gemination with the investigated affixes, i.e. Kaye (2005) and Oh & Redford (2012). Out of the two, only Oh & Redford (2012) provide enough data for a sensible comparison. The comparison with Oh & Redford's study is, however, also restricted as they only investigated the prefixes *un*- and *in*- under experimental condition. Furthermore, consonant duration was only analyzed in intervocalic position, which forbids a straightforward comparison.

For *un*-, Oh & Redford (2012) find a ratio of 1:1.6 in normal speech, and a ratio of 1:2.0 in careful speech. For *in*-, they find a ratio of 1:1.3 in normal and a ratio of 1:1.2 in careful speech. As normal speech can be assumed to be more similar to conversational speech than careful speech, the singleton-geminate ratios of the present corpus study should be compared to the ones for normal speech. For both

un- and *in-*, the comparison shows that the ratios found in the present study are bigger than the ones in Oh & Redford (2012). One reason for this difference might be the different conditions under which the items were recorded, i.e. natural conversation vs. experimental reading. The experimental study on gemination will shed more light on this issue.

To summarize, the prefixes *un-*, *in-* and *dis-* geminate. The suffix *-ly* degeminates. No word-specific effect of decomposability was found, but the segmentability of the affix affected consonant duration in the expected way. Less segmentable affixes have shorter consonant durations. The effect of affix is independent of gemination, i.e. all prefixes geminate. The comparison of singleton-geminate ratios indicates, however, that the prefix *un-* geminates to a higher degree than the prefix *dis-*.

6.3.9 Discussion

The corpus study shows that gemination is affix-specific and categorical. While the prefixes geminate, the suffix *-ly* degeminates. This result falsifies common assumptions about the gemination pattern of English affixes as found in the literature (cf. §2.4.1). With regard to the theoretical approaches discussed in Chapter 4, the result falsifies word-specific approaches, and supports approaches which predict categorical gemination that depends on the affix. However, as thoroughly discussed in Chapter 4, there are important differences in the predictions of the affix-specific approaches. In other words, the affix-specific approaches make different predictions about which affixes should geminate and which should degeminate. Therefore, it is necessary to look at the predictions of each of these approaches individually, and to compare them to the gemination pattern found in the study.

Let us first look at the formal linguistic approaches discussed. Even though all of them are categorical in nature and predict gemination to be affix-specific, none of them is supported by the data. The fact that the level 2 affix *-ly* degeminates, while the level 1 affixes *dis-* and *in-* geminate, falsifies the predictions made by stratal approaches. That *in-* and *dis-* geminate furthermore falsifies the predictions made by the Prosodic Word Approach according to which one should find variation in their gemination pattern.

With regard to the psycholinguistic approaches discussed, only the two affix-specific approaches can be supported by data, i.e. the affix-specific Segmentability Approach and the affix-specific Morphological Informativeness Approach. These two approaches are based on the two lexical segmentability hierarchies introduced in Chapter 3 (see Figure 6.3 for the two hierarchies). While the affix-

specific Morphological Informativeness Approach predicts gemination to pattern according to the Semantic Segmentability Hierarchy, the affix-specific Segmentability Approach does not specify according to which of the two segmentability hierarchies gemination should pattern. In other words, while the affix-specific Morphological Informativeness Approach is only supported if gemination patterns according to the Semantic Segmentability Hierarchy, the affix-specific Segmentability Approach is also supported if gemination patterns according to the Non-Semantic Segmentability Hierarchy.

The corpus study reveals that gemination, at least partly, patterns according to the Semantic Segmentability Hierarchy (*un-* > {*dis-*, *in-*$_{\text{Neg}}$} > *in-*$_{\text{Loc}}$ > *-ly*). The least segmentable and least informative affix *-ly* degeminates, while all other affixes geminate. The most segmentable affix *un-* displays a higher singleton-geminate ratio than the less segmentable affix *dis-*, i.e. *un-* geminates to a higher degree than *dis-*. However, the data does not suggest a difference in degree of gemination between *in-* and *un-*, or between the two *in-*prefixes. It is not yet clear whether such differences are non-existent, or whether, due to methodological reasons, they were not detected in the data set. In contrast to *un-* and *dis-*, only two environments were investigated for *in-*, i.e. m#mV and m#C. The difference in degree of gemination between *un-* and *dis-* was only detected by comparing the singleton-geminate ratio of doubles and singletons followed by a vowel, a ratio which could not be calculated for *in-*. Thus, there might be differences between the degree of gemination of *un-* and *in-* which could not be detected in this study. Furthermore, the model directly comparing *un-*, negative *in-* and locative *in-* indicates that the difference in segmentability and informativeness between the three prefixes affects their phonetic realization in the predicted direction. The model revealed that the nasal in *un-* is longer than the one in negative *in-*, which in turn is longer than the one in locative *in-*. Thus, while there is no difference in the degree of gemination between the three prefixes, there is a general difference in their nasal duration.

To conclude, the corpus study provides evidence that the phonetic realization of the investigated affixes, including the realization of morphological geminates, patterns according to the Semantic Segmentability Hierarchy. The pattern of gemination is, however, not yet clear and further evidence is needed. Up until now, the affix-specific Segmentability Approach (which is based on the Semantic Segmentability Hierarchy) and the affix-specific Morphological Informativeness Approach are the only two approaches not falsified by the data, i.e. they are the two approaches which, based on the corpus study, explain gemination in English affixation the best.

Apart from evidence for the gemination pattern of English affixes, the corpus study also revealed important insights with regard to the phonetic realization of alleged homophones. The data shows that there is a durational difference between the two *in*-prefixes. This has important implications for current models of speech production. Models which do not allow for morphological information to be visible on the phonetic level need to be revised with regard to this aspect.

While the corpus study has revealed some important insights for theories of the morpho-phonological interface, a number of questions remain unanswered. Due to the limited number of types and tokens, some effects could not be investigated properly in the study, i.e the distribution of the data did not allow for valid investigations of some factors. Furthermore, one must complement the corpus data by experimental data, which is more controlled and therefore less prone to distorting effects.

7 Experimental study

The experimental study was conducted to replicate and complement the results of the corpus study. In contrast to the corpus study, the experimental study did not investigate natural, conversational speech but read speech. As thoroughly discussed in Chapter 5, this has the advantage of including a great number of types and tokens and investigating durational properties of words under controlled conditions.

The data was collected in two experiments carried out in October 2015 and October 2016 at the University of Cambridge's Phonetic Laboratory. While in the first experiment the *un-* and *in-*data was collected, the second experiment was conducted to record the *dis-* and *-ly*-data. The set-up of both experiments was identical. The experiments consisted of two parts, a reading task and a decomposability rating. Both experimental tasks will be described in detail in the first section of this chapter. The data was analyzed with regard to two aspects, decomposability and duration. First, I will describe the decomposability analysis. Then, I will lay out the durational analyses and their results. At the end of the chapter, I will summarize the results. A thorough discussion of the results and their theoretical implications will be conducted alongside with the results from the corpus study in Chapter 8.

7.1 Methodology

7.1.1 Stimuli

Five different structures were investigated in the experimental studies. The structures and their environments are shown in Table 7.1. For each of the investigated affixes, the included environments and examples are listed (see §5.2.2 for a detailed description of the investigated structures and environments). In contrast to the corpus study, the experimental study did not only investigate the allomorph /ɪm/ for the prefix *in-*, but also /ɪn/. Note that no distinction between locative and negative *in-* is made in the table but that this distinction was taken into consideration when selecting the experimental items, as well as when analyzing the data.

Table 7.1: Overview of the investigated structures in the experimental study

	Phonological double in complex word	Singleton in complex word		Singleton in base	Orthographic double in simplex word
{un-}	*unnatural* (n#nV)	*uneven* (n#V)	*untold* (n#C)	*natural* (#nV)	NA
{in-}	*immortal* (m#mV) *innumerous* (n#nV)	NA *inefficient* (n#V)	*impossible* (m#C) *intolerant* (n#C)	*mortal* (#mV) *numerous* (#nV)	NA
{dis-}	*dissatisfy* (s#sV)	*disarm* (s#V)	NA	*satisfy* (#sV)	*dissertation* (sV)
{-ly}	*really* (Vl#l)	*truly* (V#l)	NA	*real* (Vl#)	*belly* (Vl)

The experimental stimuli were selected from a word list that features all words with the desired structures attested in COCA (Davies 2008–2014). This list was generated using the speech corpus management system Coquery (Kunter 2016). The selection of stimuli was then guided by several criteria. The first criterion was that only types which are attested in the *OED* (OED 2013) were considered to serve as stimuli in the experiments. The second criterion concerns morphological structure. All complex stimuli in the experiments had to feature different roots.[1] For example, since the two derivatives *unnatural* and *unnaturally* feature the same root, i.e. {nature}, only one of them was included in the study. Complex words with fewer morphemes were preferred over complex words with more morphemes, e.g. the word *unnatural* was preferred over the word *unnaturally*. The third criterion concerns word form frequency. For each environment, types of different frequencies were selected, i.e. the same number of types with low, mid and high frequency were selected. A low frequency word was defined to have a frequency lower than 10, a mid frequency word was defined to have a frequency between 10 and 100, and a high frequency word was defined to have a frequency higher than 100. Frequency measures were taken from COCA (Davies 2008–2014). The fourth criterion concerns prosodic structure. When possible, the

---

[1]As in the corpus study, morphological structure was coded by using the criteria described in §3.1.2.

Draft of May 27, 2019, 18:04

number of syllables and the stress pattern of the stimuli was kept similar across the environments of each affix. In other words, the distribution of stress and syllable number in a particular environment should not be unique to that environment. To test whether the prosodic structure of the stimuli differed significantly between the environments of a particular affix, Wilcoxon Signed-Rank tests (see, for example, Crawley 2012, chapter 5) were applied. In case prosodic features differed significantly, additional stimuli were added until there was no significant difference between environments.

In addition to the four criteria just mentioned, which applied to all stimuli, some criteria only applied to words with specific affixes. For *dis-*, only words featuring voiceless /s/ were included in the study. For *in-* and *dis-*, when possible, the same number of semantically transparent and opaque words was included for each environment. Since *un-* and *-ly*-affixed words do not show variation in semantic transparency (see §6.2.2 for discussion), this criterion was irrelevant for these data sets. Furthermore, when possible, the same number of derivatives with negative and locative *in-* was included for each of the *in*-environments.

The application of the selection criteria led to different numbers of included stimuli for each affix and each environment. There are two reasons for these differences. The first reason is that for some environments only few types are attested, i.e. the number of included stimuli is restricted by the number of existing types. The second reason is methodological in nature. In some environments certain features, for example, base-initial stress and high word form frequency, tend to appear together. This raises the problem that certain effects cannot be tested independently from each other. To alleviate the problem, i.e. to ensure the independence of effects, it was sometimes necessary to add more stimuli with a certain combination of features to an environment. As a result, for some environments more stimuli were included than for others. In the following, I will discuss the stimuli selection for each affix. In these discussions, I will also point out for which environments the selection criteria were not, or only partially, met.[2]

7.1.1.1 un-

Table 7.2 shows the distribution of stimuli in the *un*-data set across environments. 22 *un*-prefixed stimuli with a phonological double (n#nV) were included. These 22 stimuli represent all *un*-prefixed types with a phonological double and different roots attested in COCA (Davies 2008–2014) and the *OED* (OED 2013). Due to the small number of available types, only the first two selection criteria (attesta-

---

[2]A list of all experimental stimuli can be found in appendix E.

tion in the *OED* (OED 2013) and different roots) were taken into account when selecting these stimuli. Considering the other selection criteria would have resulted in a fatal decrease of available stimuli.

The number of stimuli with singleton environments in complex words (n#C and n#V) roughly matches the number of stimuli with a double consonant. 21 types with a singleton followed by a consonant (n#C), and 26 types with a singleton followed by a vowel (n#V) were included. Of the 21 *un*-prefixed words with a following consonant, seven were of high, seven of mid, and seven were of low frequency. For the following-vowel words, eight items were of low frequency, eight of mid frequency and 10 of high frequency. The reason for including more stimuli of high frequency was due to the distribution of prosodic structure across the environments. To make the prosodic structure of *un*-prefixed words with a singleton comparable to the one of *un*-prefixed words with a double consonant, some additional words with a specific prosodic structure needed to be added to the data set. These words were of high frequency.

The selection of the base words (#nV) was determined by the selection of words with a phonological double, i.e. for each word with a phonological double the pertinent base word was included in the experiment. Two base words were not included in the study due to a mistake in the experimental set-up.

Table 7.2: Distribution of un-types in experimental study

Environment	Example	Number of Types
n#nV	*unnatural*	22
n#C	*untold*	21
n#V	*uneven*	26
#nV	*natural*	20
Total		89

7.1.1.2 in-

Table 7.3 shows the distribution of stimuli in the /ɪn/-data set across environments. Four stimuli with a phonological double (n#nV) were included. As with *un*-, these four stimuli represent all /ɪn/-prefixed types with a phonological double and different roots attested in COCA (Davies 2008–2014) and the *OED* (OED 2013). Only the first two selection criteria (attestation in the *OED* (OED 2013) and different roots) were taken into account when selecting these stimuli.

Only 19 /m/-prefixed types with a following /t/ (n#C) were attested in COCA (Davies 2008–2014) and the *OED* (OED 2013). Eight of them feature locative *in-* and 11 negative *in-*. All of them were included in the study irrespective of their frequencies, semantic transparency and prosodic structure.

For /m/ followed by a vowel (n#V), only three types with locative *in-* were attested. All three types were included irrespective of their frequency, semantic transparency and prosodic structure. To keep the number of stimuli comparable across affixes, and to account for the lack of locative *in-*prefixed words, 24 negative *in-*prefixed types with a singleton followed by a vowel were included. They are equally distributed among the three frequency categories.

Three base words with a singleton /n/ (#nV) were included. The selection was determined by the selection of words with a phonological double (n#nV). The base of one of the four /m/-prefixed words with a phonological double is bound and was therefore not included. The prosodic structure of /m/-prefixed words is comparable across environments.

Table 7.3: Distribution of /m/-types in experimental study

Environment	Example	Number of Types
n#nV	*innumerous*	4
n#C	*intolerant*	19
n#V	*inefficient*	27
#nV	*numerous*	3
Total		83

7.1.1.3 im-

Table 7.4 shows the distribution of stimuli in the /ɪm/ -data set across environments. 19 stimuli with a phonological double (m#mV) were included. These are all /ɪm/ -prefixed types with a phonological double and different roots attested in COCA (Davies 2008–2014) and the *OED* (OED 2013). As with *un-* and /ɪn/, only the first two selection criteria (attestation in the *OED* (OED 2013) and different roots) were taken into account when selecting these stimuli.

28 complex stimuli with a singleton (m#C) were included in the study. Only eight of these stimuli feature locative *in-*. These eight are the only types with a singleton /m/ and locative *in-* attested in COCA (Davies 2008–2014) and the *OED* (OED 2013). Two of them are of low frequency, and six of them are of high

frequency. To account for the uneven distribution of frequency with locative *in-*, eight of the 20 included negative *in*-items are of mid frequency (six are of low and six are of high frequency).

For 17 of the 19 stimuli with a phonological double, the pertinent base word (#mV) was included in the experiment. Two stimuli with a phonological double feature bound roots as bases. Their bases were therefore not included as stimuli. The prosodic structure of /ɪm/ -prefixed words is comparable across environments.

Table 7.4: Distribution of /ɪm/-types in experimental study

Environment	Example	Number of Types
m#mV	*immortal*	19
m#C	*impossible*	28
#mV	*mortal*	17
Total		64

7.1.1.4 dis-

Table 7.5 shows the distribution of stimuli in the *dis*-data set across environments. 15 *dis-* stimuli with a phonological double (s#sV) were included. As with the other prefixes, these stimuli represent all types with a phonological double and different roots attested in COCA (Davies 2008–2014) and the *OED* (OED 2013), and only the first two selection criteria (attestation in the *OED* (OED 2013) and different roots) were taken into account when selecting these stimuli.

30 complex stimuli with a singleton (s#V) were included. They are equally distributed among the different frequency ranges, i.e. 10 words are of high frequency, 10 are of mid frequency and 10 are of low frequency. Only two of the *dis*-prefixed stimuli with a phonological singleton are semantically opaque.

The selection of the base words (#sV) was determined by the selection of words with a phonological double. Nine bases were included. Six bases are bound and were therefore not included.

Five simplex words with orthographic doubles (sV) were included. These were the only types attested in COCA (Davies 2008–2014) and the *OED* (OED 2013). All of them were included irrespective of their frequency and their prosodic structure.

With regard to prosodic structure, it should be noted that double consonant words and base words significantly differ from words featuring a singleton in that double consonant words and base words bear base-initial stress whereas words with a singleton often do not. Due to the lack of double consonant words and base words without base-initial stress, as well as the lack of singleton words with base-initial stress, it was impossible to even out this bias in the distribution of stress.

Table 7.5: Distribution of dis-types in experimental study

Environment	Example	Number of Types
s#sV	*dissatisfy*	15
s#V	*disarm*	30
#sV	*satisfy*	9
sV	*dissertation*	5
Total		59

7.1.1.5 –ly

Table 7.6 shows the distribution of stimuli in the *-ly*-data set across environments. 31 stimuli with a phonological double (Vl#l) were included. They are of four different types. These four types differ with regard to their orthography and with regard to whether *-ly* is preceded by an additional suffix. The four types are shown with examples in Table 7.7.

The first type are complex words ending in the orthographic string <lly> with

Table 7.6: Distribution of -ly-types in experimental study

Environment	Example	Number of Types
Vl#l	*really*	31
V#l	*truly*	30
Vl#	*real*	31
Vl	*belly*	11
Total		103

no other suffix preceding *-ly* (e.g. *really*). These words are extremely rare. Only five types are attested in COCA (Davies 2008–2014) and the *OED* (OED 2013). All were included. None of them is of high frequency.

The second type of double consonant stimuli are complex words ending in the orthographic string <lely> (e.g. *solely*). These words are also quite rare, i.e. only eight types are attested and were included as stimuli. Three of them are of low frequency, two are of mid frequency and two are of high frequency.

Most of the *-ly*-words with a double consonant are preceded by either the suffix *-al* (e.g. *educationally*), i.e. the third type of *-ly*-stimuli with a phonological double, or by the suffix *-ful* (e.g. *successfully*), i.e. the fourth type of *-ly*-stimuli with a phonological double. Nine types of <ally>-words and nine types of <fully>-words were randomly selected from a list of available words and included in the study. Three words of each type were of low frequency, three were of mid frequency and three were of high frequency.

30 types with a singleton in complex words (V#l) were included. Nine words are of low frequency, eight of mid frequency and 13 of high frequency. The reason for the uneven distribution of frequency is the phonological structure of the included types. Most *-ly*-suffixed words with a phonological singleton are preceded by the high-front vowel /ɪ/. This is different from *-ly*-suffixed words with a phonological double for which there is no specific bias with regard to their preceding vowel. To ensure the comparability of singleton and double consonant *-ly*-suffixed words, all attested *-ly*-words with a singleton which are not preceded by /ɪ/ were included in the data set (9 types). Since most of these types are of high frequency, there are more high frequency words than mid and low frequency words in the data set.

The selection of the base words (Vl#) was determined by the selection of words with a phonological double, i.e. for each word with a phonological double the pertinent base word was included in the experiment.

Table 7.7: The four types of -ly-stimuli with a phonological double

Type	Example	Number of Stimuli
<-lly>	*really*	5
<-lely>	*solely*	8
<-ally>	*educationally*	9
<-fully>	*successfully*	9

Eleven simplex words with orthographic doubles (Vl) were included. These were the only types attested in COCA (Davies 2008–2014) and the *OED* (OED 2013). All of them were included irrespective of their frequency and their prosodic█ structure.

7.1.2 Experimental set-up

In the first experiment, the *un-* and *in-*data was collected. In the second experiment, the *dis-* and *-ly-*data was collected. Furthermore, some additional *un-*prefixed words were collected in the second experiment. This was because initial analyses of the *un-*data from the first experiment called for some additional items.

The same experimental set-up was used in both experiments. First, the participants fulfilled a reading task, then they completed a rating task in which they rated all stimuli for their decomposability. At the beginning of the rating task some biographical information was recorded. This information consisted of the participants' age, sex, profession, education, the region they grew up in, and their knowledge of linguistics and Latin.[3] In the following I will first describe the reading task, then I will describe the rating task in further detail.

7.1.2.1 Reading task

The stimuli were presented to the participants in carrier sentences of two different types. While the first type was constructed to encourage the participants to read the stimulus with pitch accent, the second type was constructed to read the stimulus without accent. Each participant read each stimulus once, either in accented or in unaccented condition. The two conditions are illustrated in examples 1 and 2 for the stimulus *unnatural*.

(1)   Accented condition: *John said UNNATURAL again.*

(2)   Unaccented condition: *It is JOHN who said unnatural again, NOT HENRY.*

Accentuation was controlled in two ways. First, the syntactic position of the stimulus in the sentence led the reader's accentuation implicitly. Second, the use of capital letters explicitly indicated what to emphasize, i.e. participants were instructed to put emphasis on words written in capital letters. While in accented condition, the experimental item was written in capital letters, in unaccented position, the focus was taken away from the experimental item by capitalizing

---

[3]The full questionnaire can be found in appendix A.

the two agents of the sentence (see, for example, Plag et al. (2011) for a similar methodology).

The carrier sentences for both conditions differ slightly across affixes and environments. Table 7.8 gives an overview of all carrier sentences. In the upper part of the table the sentences for the accented condition are shown, in the lower part the sentences for the unaccented condition are shown. The first line lists the sentences for the *un*-words, the second the sentences for the /ɪn/-, /ɪm/ - and *dis*-words, and the third the sentences for the *-ly*-words. The first column lists the sentences for the base words, and the second column lists the sentences for all other structures, i.e. complex words and simplex words with orthographic doubles.

Table 7.8: Carrier sentences

Accented Condition

	Base words	Complex words and simplex words with orthographic doubles
un-	John tells you ⎯⎯ again.	John says ⎯⎯ again.
in-, im-, dis-	John tells me ⎯⎯ again.	John says ⎯⎯ again.
-ly	John said ⎯⎯ into the microphone.	It is JOHN who said ⎯⎯ to the janitor.

Unaccented Condition

	Base words	Complex words and simplex words with orthographic doubles
un-	It is JOHN who tells you ⎯⎯ again, NOT HENRY.	It is JOHN who says ⎯⎯ again, NOT HENRY.
in-, im-, dis-	It is JOHN who tells me ⎯⎯ again, NOT HENRY.	It is JOHN who says ⎯⎯ again, NOT HENRY.
-ly	It is JOHN who said ⎯⎯ into the microphone, NOT HENRY.	It is JOHN who said ⎯⎯ to the janitor, NOT HENRY.

All carrier sentences are of similar structure. The stimulus is always followed by at least one other word to avoid effects of phrase-final lengthening, and for

each affix, the number of syllables in the carrier sentences is kept constant across environments. In the carrier sentences for the base words, an additional syllable was added to compensate for the missing affix.

To ensure a reliable annotation of the recorded data, the carrier sentences were created in such a way that the crucial sounds could be segmented without difficulty. The crucial sounds for the prefix-data are the first sounds of the stimuli, and the crucial sounds for the suffix-data are the last sounds of the stimuli. For the prefixed data, the word preceding the stimulus was chosen to end in a sound which is easily recognizable in the speech signal, i.e. which can be easily segmented from the initial sound of the prefix. For the suffixed data, the word following the stimulus was chosen to start in an easily recognizable sound, respectively. The prefixed stimuli start with either a vowel (*un-* and *in-*) or a plosive (*dis-*). As these two types of sounds can be easily distinguished from fricatives, the word *says*, which ends in a fricative, was chosen to precede the prefixed stimuli. The *-ly*-stimuli are followed by the word *to*, which starts with a stop consonant. Stops can be easily segmented from preceding vowels.

In the sentences for base words, words with a particular phonological make-up were chosen to precede (in case of prefixes), or follow (in case of *-ly*), the stimulus. This was to ensure that the environment of the base-initial consonant (for the prefixes), and the environment of the base-final consonant (for *-ly*), is comparable to the environment of the pertinent double consonant. For *un-*, the segment preceding the prefixal nasal is a back vowel, and so is the last segment of the word *you*, which precedes the base-initial nasal in the experimental sentences for the base words. For *in-* and *dis-*, the segment preceding the prefixal consonant is a front vowel, and so is the last segment of the word *me*, which precedes the base-initial consonant in the experimental sentences. For *ly-*, the segment following the affixational consonant is a front vowel, and so is the first segment of the word *into*, which follows the final lateral of the bases in the experimental sentences.

After the stimuli were put into the carrier sentences, unrelated filler sentences were added to the experimental sentences. The filler sentences were extracted from the *Corpus of American Soap Operas* (Davies 2011-), and were included to avoid a list-reading effect, i.e. to ensure that the experimental sentences were read as naturally as possible. The first experiment included 183 experimental sentences and 127 filler sentences. The second experiment included 205 experimental sentences and 130 filler sentences.

The sentences were presented to the participants on a screen in random order. Participants were instructed to read the sentences as naturally as possible and to

repeat a sentence if they made a mistake. They were told to read the sentences at their own pace, and to go on to the next sentence whenever they were ready. Before the experimental sentences were presented, participants were presented with four example sentences. The reading took place in a sound-proof booth, and the recordings were made using a free-standing microphone and a digital recorder. After participants completed the reading task, they took a short break and then proceeded to the rating task.

7.1.2.2 Decomposability rating

After completing the reading task, all participants were asked to rate the stimuli for their decomposability. The rating was carried out using the software LimeSurvey (LimeSurvey Project Team & Carsten Schmitz 2015). Participants of experiment 1 rated all complex stimuli included in experiment 1, and participants of experiment 2 rated all complex stimuli included in experiment 2.

The rating was designed in the same way as the one conducted for the corpus study (see §6.1.2 for a detailed description). After an explanation of the complexity of words, participants were asked to rate how decomposable a word is on a 4-Point-Likert scale. Participants were able to indicate that they did not know a word, i.e. only words which are known by the participants were rated. In addition to the complex stimuli, participants also rated simplex words starting or ending with the same graphemes as the affixed stimuli, e.g. *uncle* and the word *family*. Including simplex words served the purpose of testing whether participants correctly understood the task. Simplex words should be rated as very difficult to decompose.

7.1.3 Participants

51 native speakers of British English participated in the experiments, 29 in the first experiment and 22 in the second experiment. The participants had no or only little knowledge of linguistics and were naive to the study's purpose. 26 participants were male and 27 female. Their age ranged between 18 and 65 with a median age of 21. None of the participants reported any hearing or speech impediments.

7.1.4 Processing of the sound files

Overall 9590 experimental sentences were recorded, 5130 in the first experiment and 4460 in the second experiment.[4] Some of the recorded sentences were excluded from the phonetic annotation due to misreadings and unnatural productions. Furthermore, some items were excluded because a valid segmentation was impossible. In the first experiment 266 tokens were excluded from segmentation, in the second experiment 376 were excluded from segmentation.

The sentences were segmented and phonetically annotated as described in §5.3. During the segmentation process all tokens were coded for their environment (Environment). For *-ly*-words this included coding for syllabicity. Furthermore, *-ly*-words were coded for the type of /l/ they featured (TypeOfL), i.e. `approximant`, `tap` or `vocalized` (see §5.5.1 for a detailed description of the coding).

The acoustic analysis revealed that for a few of the complex tokens, a pause was produced between the affix and its base, or that some part of the affix was deleted. These tokens were excluded from the study. In the first experiment, 43 items were excluded because of a missing vowel in the affix (10 *un*-prefixed tokens, 9 /ɪm/ -prefixed tokens, 22 /ɪn/-prefixed tokens), and 48 were excluded because a pause was produced between prefix and base (30 *un*-prefixed tokens, 6 /ɪm/ -prefixed tokens, 12 /ɪn/ -prefixed tokens). In the second experiment, 41 tokens were excluded because of a missing vowel in the affix (all *dis*-prefixed), and 6 tokens were excluded because of a pause between affix and base (4 *dis*-prefixed tokens, 2 *un*-prefixed tokens).

Furthermore, 353 recorded *-ly*-tokens were excluded from the study. All of these tokens are base words which feature a vocalized /l/ (cf. variable TypeOfL). As explained in detail in §5.3, /l/ was coded as vocalized whenever the boundary between /l/ and its preceding vowel was not clearly detectable in the signal. As in these cases the boundary between /l/ and its preceding vowel was not segmented, the duration of /l/ in these tokens is not comparable to the duration of /l/ in the other tokens. Therefore, all items featuring vocalized /l/ were excluded.

The final data set only includes tokens known by the pertinent speaker. In other words, tokens which were marked as unknown by the speaker in the rating task were excluded from the analyses. In the first experiment 153 tokens were marked as unknown, in the second experiment 63 tokens were marked as unknown.

---

[4]Note that some sentences were skipped by the participants, i.e. some sentences were not recorded. In the first experiment, 177 sentences were not recorded. In the second experiment, 50 sentences were not recorded.

Overall 8241 tokens entered the analyses, 4620 were recorded in the first experiment, and 3621 were recorded in the second experiment. Table 7.9 gives an overview of the final number of tokens for each environment.

7.1.5 Processing of the rating data

To test the validity of the rating, it was checked whether participants rated the simplex words included in the ratings as significantly more difficult to decompose than the complex words. If so, one can assume that the task was understood correctly and that the rating is valid. Except for one participant in the first experiment, all other participants made a clear distinction between complex and simplex words. This indicates that participants understood the task correctly and were able to recognize differences in decomposability between the words.

As laid out in §5.5.1, the rating of each token was coded in the variable SemanticTransparencyRating. Note that this coding of SemanticTranspa-rencyRating■ is different from the coding of SemanticTransparencyRating in the corpus study. In the corpus study, the variable was coded by computing the median rating for each investigated type, i.e. the variable represented the average rating for an item. In the experimental study, in contrast, the variable is not coded by computing an average. Instead each recorded token was rated by the speaker of that particular token, and this rating was used to code the variable. This has the great advantage that it allows us to directly link each participant's rating to his/her production of the token.[5]

7.1.6 Variable coding

After segmentation, all tokens were annotated with regard to factors possibly influencing consonant duration. These factors were described in detail in §5.5. Two of the variables described in §5.5 were not coded for in the experimental study, i.e. LSAScore and Voicing. LSAScore was not coded for because the results of the corpus study suggest that this variable does not form a good operationalization of decomposability. Voicing was not coded for because all pertinent fricatives in the experimental study, i.e. the /s/ in the *dis*-words, are voiceless.

In addition to the factors described in §5.5, the variable Order was coded. This variable codes for the order in which the tokens were presented to the participants during the reading task. It was included in the statistical models to control

---

[5]Note that tests of inter-rater reliability are unnecessary for the experimental rating. The reason is that, different from in the corpus study, SemanticTransparencyRating does not represent an average in the experimental study, and that the coding of the variable does therefore not rely on similar ratings across participants.

Table 7.9: Distribution of tokens in experimental study

un-

Environment	Example	Number of Tokens
n#nV	*unnatural*	966
n#C	*untold*	427
n#V	*uneven*	674
#nV	*natural*	548
Total		2615

in- | | | | **im-** | | |

Environment	Example	Number of Tokens	Environment	Example	Number of Tokens
n#nV	*innumerous*	88	m#mV	*immortal*	488
n#V	*inefficient*	630			
n#C	*intolerant*	437	m#C	*impossible*	689
#nV	*numerous*	77	#mV	*mortal*	458
Total		1232			1635

dis-

Environment	Example	Number of Tokens
s#sV	*dissatisfy*	242
s#V	*disarm*	587
#sV	*satisfy*	191
sV	*dissertation*	94
Total		1114

-ly

Environment	Example	Number of Tokens
l#l	*really*	464
syll. l#l	*ment(a)lly*	132
#l	*truly*	609
l#	*real*	218
syll. l#	*ment(a)l*	21
l	*belly*	201
Total		1645

for possible training effects, i.e. for the possibility that participants became faster in their production of the tokens throughout the reading task.

The levels of the variable Environment were recoded in the *dis-* and in the *-ly*-data set.[6] For *dis-*, the variable was enriched with information about base-initial stress. As already mentioned in §7.1.1.4, in the *dis*-data set the variable BaseInitialStress is unevenly distributed across environments. While all types with a phonological double (`s#sV`) and all base words (`#sV`) feature a stressed base-initial syllable, all types with an orthographic double (`sV`) feature an unstressed base-initial syllable. Only in complex words with a singleton (`s#V`) variation is found, i.e. only here we find types with an unstressed base-initial syllable and types with a stressed base-initial syllable.

The interconnectedness of the two variables Environment and BaseInitial-Stress might cause problems in the statistical modeling of the data. To avoid this problem, the two variables were collapsed, i.e. the variable Environment was enriched by information about the stress status of the base-initial syllable. Since only singletons in complex words (`s#V`) show variation with regard to base-initial stress, the lumping of the variables Environment and BaseInitialStress resulted in only one additional level. The five levels of the variable Environment in the *dis*-data set are thus phonological doubles with a stressed base-initial syllable (`s#sV-stressed`), singletons in complex words with a stressed base-initial syllable (`s#V-stressed`), singletons in complex words with an unstressed base-initial syllable (`s#V-unstressed`), singletons in base-words with a stressed base-initial syllable (`#sV-stressed`), and orthographic doubles with an unstressed base-initial syllable (`sV-unstressed`).

For *-ly*, the variable Environment was enriched with information about orthography. In the *-ly*-data set, /l/ and /ll/ are represented by four different orthographic strings: <l> as in *truly*, *real* or *mental*, <ll> as in *really*, *mentally* or *belly*, <lel> as in *solely* and <le> as in *sole*. As orthography might influence consonant duration (see discussion in §5.2.2), one should test the influence of the orthographic string /l/ and /ll/ are represented by in the statistical models.

However, while some orthographic strings represent more than one environment, others only represent one. Table 7.10 shows the distribution of the experimental tokens with the different orthographic strings across the five environments of *-ly*-words. For each environment, examples with the different orthographic structures are given. The table shows that while <l> and <ll> represent three environments each, <le> and <lel> only represent one. The variable Environment and the orthographic string /l/ or /ll/ is represented by are thus inter-

---

[6]See §5.5.1 for initial coding of the variable Environment.

Draft of May 27, 2019, 18:04

connected.

Table 7.10: Distribution of *-ly*-tokens across ENVIRONMENT and ORTHOGRAPHY

		<l>	<ll>	<le>	<lel>
l#l	*really, solely*	0	313	0	151
syll. l#l	*ment(a)lly*	0	132	0	0
#l	*truly*	609	0	0	0
l#	*real, sole*	115	0	103	0
syll. l#	*ment(a)l*	21	0	0	0
l	*belly*	0	201	0	0

Due to this interconnectedness, it was not reasonable to code the factor orthography as an independent variable and test its effect independently from ENVIRONMENT. To nevertheless account for possible effects of orthography, the levels of the variable ENVIRONMENT were enriched with information about orthography. This resulted in two additional factor levels. The eight levels of the variable ENVIRONMENT in the *-ly*-data set are: l#l-<ll> (*really*), syll.l#l-<ll> (*ment(a)lly*), l#l-<lel> (*solely*), #l-<l> (*truly*), l#-<l> (*real*), syll.l#-<l> (*ment(a)l*), l#-<le> (*sole*), l-<ll> (*belly*).

Overviews of all variables initially included in the models predicting absolute consonant duration are given in tables F.1 - F.5 in appendix F.

7.2 Decomposability

Decomposability was analyzed similarly as in the corpus study. On the one hand, I investigated the relations between the different decomposability measures. This is, as discussed in §6.2.1, important with regard to their suitability in the study. Only when the relation between the different decomposability measures is clear, their possible effects on duration can be interpreted adequately. On the other hand, I compared the affix's segmentability to test whether the segmentability hierarchies proposed in §3.2 are borne out by the data. As the hierarchies serve as the basis of two of the theoretical predictions made in Chapter 4, i.e. the predictions of the affix-specific Segmentability Approach and the predictions of the affix-specific Morphological Segmentability Approach, their validation is of high importance (see §6.2.2 for details).

In the segmentability comparison of the affixes, I concentrated on the variable SEMANTICTRANSPARENCYRATING. The reason is that the distribution of the other three decomposability variables is highly influenced by the choice of the included stimuli. As laid out in §7.1.1, the experimental stimuli were (as far as possible) controlled for their semantic transparency and their word form frequency. Both factors highly correlate with the three decomposability variables SEMANTICTRANSPARENCY, TYPEOFBASE and logRELATIVEFREQUENCY. Only the variable SEMANTICTRANSPARENCYRATING is independent from the stimuli selection.

7.2.1 The relation between decomposability measures

To investigate the relation between the decomposability measures in the experimental study, hierarchical cluster analyses were conducted. As explained in §6.2.1, this type of analysis can successfully be applied to investigate the similarity between different variables, and is thus well suited to provide insights about the relation between the decomposability variables (see also Baayen 2008: 200f.).

The type of cluster analysis applied here first computes Spearman's rank correlations between all included variables, then squares them, and then puts them into a correlation matrix (see §6.2.1 for a detailed description of cluster analyses). Since correlations can only be calculated for numerical variables, the categorical variable SEMANTICTRANSPARENCY was recoded into the numerical variable NUM-SEMANTICTRANSPARENCY, and the categorical variable TYPEOFBASE was recoded into the numerical variable NUMTYPEOFBASE. The variable NUMSEMANTICTRANS-PARENCY featured the two levels 1 (= opaque) and 2 (= transparent), the variable NUMTYPEOFBASE the levels 1 (= word as base) and 2 (= bound root). After the correlation matrix was created, a dendrogram in which the correlations are displayed was generated. Variables which cluster together in the dendrogram are more similar, i.e. have a higher correlation, than variables which are more distant in the dendrogram.

As in the corpus study, I conducted three cluster analyses, one for the whole data set, one for the *in*-prefixes and one for *dis*-. For *un*- and *-ly*, it was not reasonable to conduct cluster analyses. This is due to the distribution of the decomposability variables in the *un*- and the *-ly*-data set. The two decomposability variables SEMANTICTRANSPARENCY and TYPEOFBASE do not show enough variability to be investigated in the subsets.

In contrast to the corpus study, the cluster analyses in the experimental study were based on tokens, not on types. This has two reasons. First, in contrast to the corpus study, in the experimental study each type is represented by approx-

imately the same number of tokens, i.e. there is no problem of specific types being underrepresented. Second, the variable SEMANTICTRANSPARENCYRATING is token-based. To investigate the correlations between this variable and the other decomposability variables, it is thus necessary to consider every token. All cluster analyses were generated in R using the Hmisc package (Harrell Jr 2017).

The first cluster analysis investigated all tokens of the experimental study. The correlation matrix created in the analysis is shown in Table 7.11. The highest correlation is found between the two variables NUMSEMANTICTRANSPARENCY and NUMTYPEOFBASE. All other correlations are much lower. A separate analysis of the non-squared correlations revealed that all correlations go in the expected direction.

Table 7.11: Correlation matrix for decomposability measures in experimental study

	SEMANTICTRANS- PARENCYRAT	NUMSEMANTIC- TRANSPARENCY	NUMTYPE- OFBASE	logRELATIVE- FREQUENCY
SEM.TRANSP.RAT.	1.00			
NUMSEM.TRANSP.	0.12	1.00		
NUMTYPEOFBASE	0.12	0.44	1.00	
logREL.FREQ.	0.05	0.08	0.10	1.00

Figure 7.1 displays the relation between the variables in a dendrogram. On the y-axis the squared Spearman correlation score between the variables is displayed. The figure shows two splits which structure the variables into three clusters. The lower the split in the figure is, the higher is the correlation between the variables of the pertinent cluster. The first split is positioned in the upper part of the figure and separates the variable logRELATIVEFREQUENCY from all other variables. That logRELATIVEFREQUENCY forms its own cluster in the upper part of the figure indicates the dissimilarity of this variable to the other decomposability variables. The second split, which is also positioned in the upper part of the figure, separates the variable SEMANTICTRANSPARENCYRATING from NUMSEMANTICTRANSPARENCY and NUMTYPEOFBASE. This indicates that SEMANTICTRANSPARENCYRATING is also rather dissimilar from the other decomposability variables. NUMSEMANTICTRANSPARENCY and NUMTYPEOFBASE are much more similar to each other. This is indicated by the cluster they form in the lower part of the figure.

The results of the cluster analyses for *in-* and *dis-* are displayed in the dendrograms in Figure 7.2. The results resemble the result of the first cluster analysis. For both prefixes, the variables NUMSEMANTICTRANSPARENCY and NUMTYPEOF-
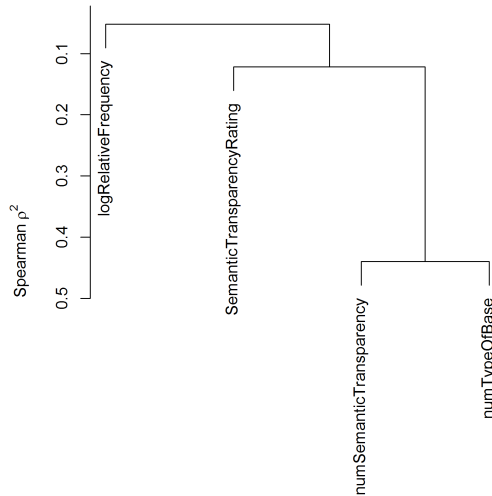
Figure 7.1: Dendrogram of the four decomposability measures for all words in the experimental study
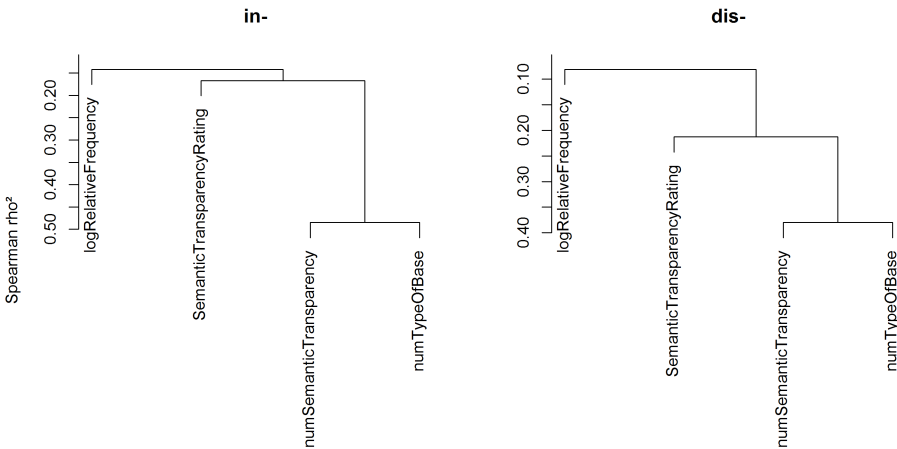
Figure 7.2: Dendrogram of the four decomposability measures for *in-* and *dis-*prefixed words in the experimental study

Base cluster together in the lower part of the figure. This means that the correlations between these two variables are rather high. LogRelativeFrequency and SemanticTransparencyRating do not correlate to a high degree with any other variable. For *dis-*, SemanticTransparencyRating is a little more similar to numSemanticTransparency and numTypeOfBase than for *in-*. Overall, logRelativeFrequency and SemanticTransparencyRating are, however, not very similar to the other decomposability variables in both data sets.

To sum up, the cluster analyses have revealed that the two variables SemanticTransparency and TypeOfBase are very similar. The two variables logRelativeFrequency and SemanticTransparancyRating, in contrast, barely correlate with any other decomposability variable. This outcome only partly resembles the outcome of the corpus study. While both studies found that SemanticTransparency and TypeOfBase are very similar, and that the variable logRelativeFrequency is different, the outcome for SemanticTransparencyRating differs between the studies. In the corpus study, SemanticTransparencyRating is very similar to SemanticTransparency and to TypeOfBase. In the experimental study, the correlations between SemanticTransparencyRating and TypeOfBase, and between SemanticTransparencyRating and SemanticTransparency are not very high. They are barely higher than the ones between logRelativeFrequency and the three other decomposability variables. This means that while ratings of the corpus study were highly influenced by the semantic transparency and the type of base of a derivative, the ratings of the experimental study were less influenced by these factors.

One possible explanation for this difference between the corpus and the experimental ratings is that the two groups of raters might have differed with regard to their definition of decomposability, and that this difference might have led to different rating strategies. This explanation especially makes sense regarding the fact that the experimental raters were younger and mostly students, while the corpus raters represent a random selection of people of different ages. It might be the case that the younger raters, who are still in school, used a rule-based rating strategy which is based on their knowledge about word-formation, while the older raters, who might not have such knowledge, relied on information about semantic transparency and the type of base of a word.

The relation between the decomposability measures has important implications for the interpretation of possible decomposability effects on duration. As in the corpus study, possible effects of the decomposability variables SemanticTransparency and TypeOfBase can be assumed to be caused by the same underlying property. Effects of logRelativeFrequency and SemanticTransparency-

Rating on duration are probably caused by different underlying properties (see also §6.2.1 for a discussion of the concept decomposability and its operationalization in the study).

	Segmentability hierarchy	Additional assumption
Semantic Hierarchy	*un-* > {*dis-*, *in-*$_{\text{Neg}}$}> *in-*$_{\text{Loc}}$ > *-ly*	lexical meaning over productivity, transparency and type of base
Non-Semantic Hierarchy	*un-* > *-ly* > {*dis-*, *in-*$_{\text{Neg}}$}> *in-*$_{\text{Loc}}$	productivity, transparency and type of base over lexical meaning

Figure 7.3: Lexical segmentability hierarchies of affixes

7.2.2 The segmentability of the affixes: A comparison

Figure 7.3 displays the segmentability hierarchies proposed in §3.2. To see whether the hierarchies are borne out by the data, it is necessary to compare the segmentability of the five investigated affixes as found in the data. If the hierarchies are valid, the segmentability of the affixes should pattern according to the hierarchies. As explained above, in the experimental study, only the distribution of the variable SemanticTransparencyRating across affixes was compared to investigate the segmentability of the affixes.

Table 7.12 shows the distribution of SemanticTransparencyRating for each affix. Next to the total number of tokens, the percentage of tokens with the pertinent rating per affix is given.

Overall most items were rated as quite easy to decompose, i.e. the majority of items was rated with 1. This distribution supports the suspicion that experimental raters might have used a rule-based approach in their rating, i.e. they categorically rated items as either decomposable or not decomposable. However, the table also shows that there is variation in the ratings. Crucially, there are differences in the distribution of ratings between affixes. Kruskal-Wallis tests ($p < 0.05$) revealed that all differences between affixes, except the one between negative *in-* and *dis-*, are significant.

Table 7.12: Semantic Transparency Rating by affix

SEMANTICTRANS-PARENCYRATING	*in*-$_\text{Loc}$		*-ly*		*dis-*		*in*-$_\text{Neg}$		*un-*	
1 - most decomposable	201	(35%)	747	(62%)	590	(71%)	1225	(71%)	1868	(92%)
2	81	(14%)	213	(18%)	119	(14%)	244	(14%)	129	(6%)
3	100	(17%)	182	(15%)	69	(8%)	148	(9%)	37	(2%)
4 - least decomposable	194	(34%)	63	(5%)	51	(6%)	100	(6%)	5	(< 1%)

The prefix *un-* is rated as the most segmentable affix. Locative *in-* is rated as the least segmentable affix, and the other three affixes pattern in between. This pattern partly resembles the pattern found in the corpus study. In both studies, *un-* was rated the most segmentable affix, and locative *in-* was rated the least segmentable affix. However, differently from the corpus study, in the experimental study, the suffix *-ly* is rated as slightly less decomposable than negative *in-* and *dis-*. In the corpus study, *-ly* was rated as the second most segmentable affix after *un-*.

The difference in the rating of the suffix *-ly* between corpus and experimental rating is very interesting with regard to the segmentability status of the suffix and the segmentability hierarchies. The placement of *-ly* in the segmentability hierarchies highly depends on the definition of decomposability. On the one hand, *-ly* is very segmentable in terms of its productivity, its transparency and the types of bases it takes, on the other, the suffix does not feature a clear lexical meaning and its status as a derivational suffix is contested in the literature (see, for example, Zwicky 1995; Plag 2003; Giegerich 2012; Bauer et al. 2013 for discussion). Depending on one's definition of decomposability, *-ly* is either a very segmentable affix (Non-Semantic Segmentability Hierarchy) or an affix with very low segmentability (Semantic Segmentability Hierarchy) (see also discussion in §3.1.6).

The different segmentability patterns found in the corpus and the experimental study mirror the ambiguous segmentability status of *-ly*. In turn, they can be interpreted to provide support for the validity of both segmentability hierarchies. In the corpus study, the suffix *-ly* is the second most segmentable affix. This is in line with the Non-Semantic Segmentability Hierarchy. In the experimental study, the suffix *-ly* is rated as one of the least segmentable affixes. This is in line with the Semantic Segmentability Hierarchy.[7] The segmentability pattern of the

---

[7] Note, however, that according to the Semantic Segmentability Hierarchy, locative *in-* is expected to be less segmentable than *-ly*. This is not the case.

prefixes provides further support for the validity of the two hierarchies. In both studies, the segmentability of the prefixes patterns according to both hierarchies.

7.2.3 Summary

The investigation of the relation of the decomposability measures has revealed that while the two variables SemanticTransparency and TypeOfBase are very similar to each other, the two variables logRelativeFrequency and Semantic-TransparencyRating are different from all other decomposability measures. This has important implications for the interpretation of possible decomposability effects on duration. While possible effects of SemanticTransparency and TypeOfBase can be assumed to be caused by the same underlying property, possible effects of logRelativeFrequency and SemanticTransparencyRating on duration are probably caused by different underlying properties.

With regard to the segmentability of the affixes, the experimental rating shows a similar pattern as the corpus study. As in the corpus study, the prefix *un-* is rated as the most segmentable affix, and locative *in-* is rated as the least segmentable affix. However, the suffix *-ly* is rated differently in the experimental rating, i.e. it is the second least segmentable affix, whereas it is rated the second most segmentable affix in the corpus. The different ratings for *-ly* in the corpus and the experimental study mirror the ambiguous segmentability status of the affix and its different placements in the segmentability hierarchies.

7.3 Duration

7.3.1 Analyses

As in the corpus study, each affix (and each allomorph, if there was more than one) was investigated separately, i.e. five subsets were created, one for the *un*-words, one for the /ɪn/-words, one for the /ɪm/ -words, one for the *dis*-words, and one for the *-ly*-words.

To get a first impression of the gemination pattern, and to test whether gemination is a categorical or a gradient phenomenon, the first durational analysis consisted of investigating the raw distribution of consonant duration in each subset (cf. *Nature of gemination: Predictions* in §4.3.1). To see whether the distributions differ between environments, I generated boxplots for each environment of each subset. If the boxplots indicate a bimodal distribution with doubles having a higher mean than singletons, one can assume gemination to be categorical (see §6.3.1 for detailed description of the analyzing strategy).

After investigating the raw distributions across environments, I fitted two linear mixed effects regression models to each subset. The first model predicts absolute consonant duration with all complex words of a given subset (*complex model*). These models are very similar to the models fitted in the corpus study, i.e. they include complex words with a phonological double (e.g. *unnatural*) and complex words with a phonological singleton (e.g. *uneasy*). The second model, the *complete model,* predicts absolute consonant duration with all tokens of a pertinent subset, i.e. the models also include base words with a singleton (e.g. *natural* or *real*), and simplex words with an orthographic double (e.g. *dissertation* or *belly*). The complete models were fitted to test whether phonological doubles are longer than corresponding singletons in base words, and whether gemination is affected by the presence of orthographic doubles.

In addition to the two models for each subset, one model which directly compares *un-* and /ɪn/-prefixed words was fitted. This model was fitted to test whether the three prefixes *un-*, locative *in-* and negative *in-* deviate in their durational patterns. No other affixes were directly compared in one model as inherent durational differences between different types of consonants are too severe to be investigated in one model. Note that in the model featuring *un-* and /ɪn/-prefixed words, some variables cannot be investigated because there are systematic differences in the distribution of variables between the prefixes. Details will be given in the pertinent section.

The dependent variable of all models was absolute consonant duration. Based on the findings of the corpus study, no models with relative duration as the dependent variable were fitted. The corpus study revealed that relative consonant duration is a much weaker measure of morphological gemination in English than absolute consonant duration (see §6.3.8).

Only the complex models tested the effects of the decomposability measures on consonant duration, since the decomposability variables are not applicable to simplex words. Furthermore, not all decomposability measures were tested for all affixes. As in the corpus study, only in the complex *in-* and *dis-*models all decomposability variables were included. In the complex *-ly*-model, only the effects of logRELATIVEFREQUENCY and SEMANTICTRANSPARENCYRATING were tested. In the complex *un-*model, only the effect of logRELATIVEFREQUENCY was tested. The reason is that the two variables SEMANTICTRANSPARENCY and TYPEOFBASE do not show enough variation for *un-* and *-ly*. For *un-*, SEMANTICTRANSPARENCYRATING does not show enough variation either.

In the complex models predicting consonant duration with /ɪn/, /ɪm/ and *dis-*, collinearity problems had to be addressed. As discussed in the previous section,

the decomposability variables SEMANTICTRANSPARENCY and TYPEOFBASE highly correlate, and there are also correlations between the other decomposability variables. It is thus problematic to test all of them simultaneously in the model. Therefore, the effect of these variables was tested by including them individually in the model, and by conducting principal component analyses (see §5.4 on principal component analyses).

In the complete models for the prefixes, the noise variable PRECEDINGSEG-MENTDURATION was not included because the base-initial consonant does not feature a preceding segment. In all models, the two variables SPEAKER and TYPE were included as random effects.

As in the corpus study, two types of interactions were tested in each model. First, I tested for interactions which are predicted to affect gemination according to the theoretical approaches discussed in Chapter 4. Then, I tested for interactions which, based on previous empirical work and theoretical considerations, can be assumed to affect affixational consonant duration (see §6.3.1 for a more detailed description of the two types of interactions). All interactions tested in the experimental study are listed in appendix G.

All models were fitted according to the modeling strategy described in §5.4. The models were generated using the `lme4 package` (Bates et al. 2014), and the plots of the regression models were generated with the `visreg package` (Breheny & Burchett 2015).

After fitting the complex models for each subset, I computed each variable's contribution to the goodness of fit for the final model by checking how the absence of each significant term affects the AIC of the model. The higher the increase of the AIC without a specific term, the more variance is explained by the pertinent term in the model, and the higher is its contribution to the goodness of fit.

7.3.2 Overview

Figure 7.4 depicts the distribution of consonant duration for each environment in each subset using boxplots. In the upper row, the distribution for *un-* is shown in the left panel, the one for /ɪn/ is shown in the middle panel, and the one for /ɪm/ is shown in the right panel. The distribution for *dis-* is shown in the lower left panel, and the one for *-ly* in the lower right panel. The y-axis of each plot displays the duration of the consonant in milliseconds.

The boxes in each plot represent the distribution of consonant duration for the different environments in each data set. The dot in the middle represents the median duration, and the box itself represents the interquartile range of conso-

nant duration. In each plot, the left box(es) represent the environment(s) with a phonological double. They are followed by the box(es) for complex words with singleton environments and the box(es) for base words with a singleton. For *dis-* and *-ly*, the last box (from the left) represents the consonant duration in simplex words with an orthographic double.
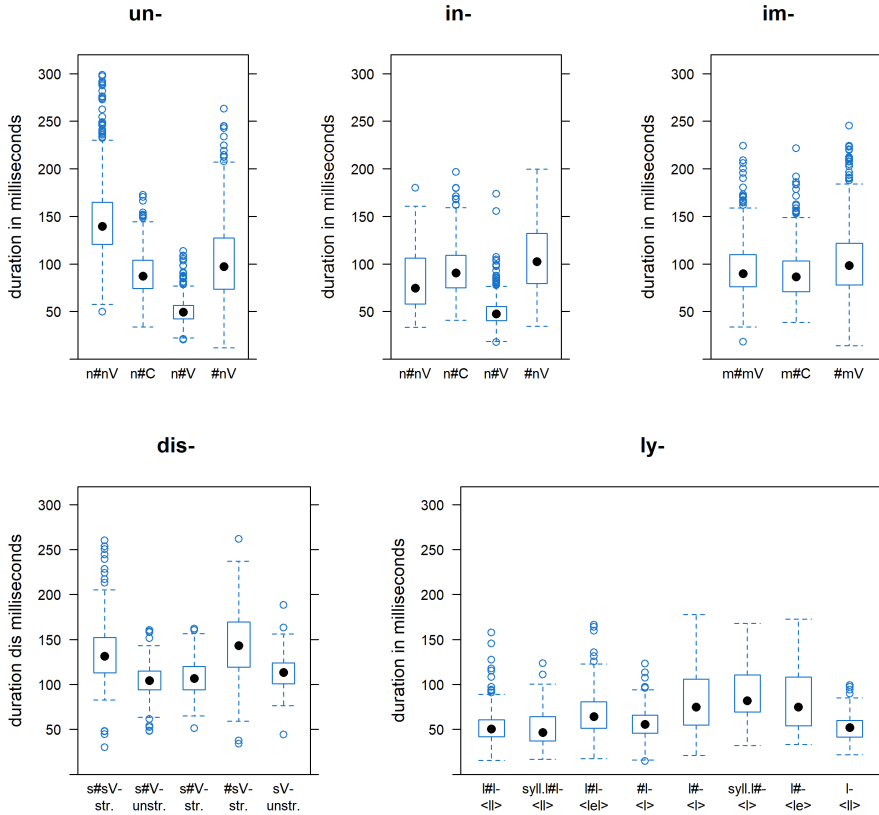
Figure 7.4: Distribution of consonant duration in the five data sets

For *un-*, the figure shows a clear difference in duration between double and single consonants. Doubles (n#nV) are longer than singletons in complex words (n#C, n#V) and singletons in base words (#nV). The figure also shows that there is no overlap of the boxes for the doubles and the boxes for the singletons. This indicates a bimodal distribution in the data set with doubles being longer than singletons. As the data from the corpus study, the plot thus suggests that *un-*

geminates, and that gemination is a categorical phenomenon.

For /ɪn/, the plot suggests that doubles (n#nV) are as long as the singletons in complex words followed by a consonant (n#C) and singletons in base words (#nV). Only singletons in complex words followed by a vowel (n#V) are shorter than doubles. On the one hand, this durational difference speaks for gemination with *in-*, on the other, there is no durational difference between the other singleton levels and the double consonant. This is different from what was found for *un-*. With regard to the question whether gemination is a categorical or a gradient phenomenon, the plot shows that there is no overlap between the box of the doubles (n#nV) and the box of the singletons with a following vowel (n#V) for /ɪn/. In other words, the distribution of duration of doubles and singletons with a following vowel seems to be bimodal. This suggests that, if there is gemination with /ɪn/, it is probably categorical.

For /ɪm/ , no difference in consonant duration can be seen between the three environments (m#mV, m#C, #mV). The plot thus suggests degemination with /ɪm/ . However, it is yet unclear whether the allomorph /ɪm/ really degeminates, or whether gemination with /ɪm/ depends on additional factors which are not taken into account when comparing the raw durations of doubles and singletons.

For *dis-*, the plot shows that doubles (s#sV-str.) are longer than singletons in complex words (s#V-unstr., s#V-str.) and singletons in simplex words with an orthographic double (sV-unstr.). However, singletons in base words (#sV) are as long as doubles. As for /ɪn/, there is thus some evidence for gemination, but also some evidence against it. The boxes for the double consonants and for the singleton environments (except for the one for singletons in base words) hardly overlap. Thus, the distribution of duration of doubles and singletons in complex words seems to be bimodal. This suggests that if there is gemination with *dis-*, it is categorical.

For *-ly*, there is a big overlap in the distribution of all environments. Only singletons in base words (l#-<l>, syll.l#-<l>, l#-<le>) seem to have slightly longer durations than singletons in all other environments. Crucially, the plot does not suggest doubles to be longer than singletons of any category. One might thus suspect degemination for *-ly*.

The overview of the durations of all environments reveals some similarities across all subsets, as well as some differences. One similarity is that in all subsets, the consonant in base words is relatively long. This might be due to its word-initial (or word-final) position. Furthermore, for the prefixes with a nasal, the nasal is longer before consonants than before vowels. Another similarity is that if there are differences between double and singleton consonants, their duration

seems to be bimodally distributed. As in the corpus study, the data thus suggest gemination to be a categorical phenomenon.

With regard to the question of gemination, the affixes seem to behave quite differently. For *un-*, the distributional analysis clearly suggests gemination, doubles are longer than all singleton levels. For the other affixes, it is less clear whether we find gemination. For *in-*, i.e. /ɪn/ and /ɪm/ , only one singleton environment features shorter consonants than the double environment. For *dis-*, the singleton consonants of all but one environment are shorter than the double consonants. For *-ly*, none of the three double environments is longer than the singleton environments. Further analyses which take more variables into account are needed to clarify the gemination pattern of the affixes. In the next subsections I will present such analyses for each subset.

7.3.3 The prefix *un-*

7.3.3.1 Complex model

The model predicting consonant duration with all complex *un*-words (N = 2067) was fitted according to the modeling procedure described in §5.4. Due to an uneven distribution of the residuals in the initial model, the dependent variable AbsoluteConsonantDuration was Box-Cox-transformed (λ = 0.101) and 50 outliers were removed (2.4% of the data). After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions).

After model simplification five variables remained in the model, Environment, Accentuation, LocalSpeechRate, PrePause and PrecedingSegmentDuration. The two variables Environment and Accentuation interact. The final model is summarized in Table H.1 which can be found in appendix H.

The four noise variables show the expected effects. As in the corpus study, the nasal in *un-* becomes shorter with increasing speech rate. The nasal in tokens which are preceded by a pause is longer than the nasal in tokens without a preceding pause. This effect of PrePause can be attributed to word-initial strengthening. Furthermore, consonant duration depends on the duration of the preceding segment. The longer the preceding segment, the shorter the nasal.

The fourth significant noise variable Accentuation forms an interaction with the variable of interest Environment. The interaction is depicted in Figure 7.5. The light blue lines in the figure show the effect of Environment for items in accented position, and the dark blue lines show the effect for items in unaccented
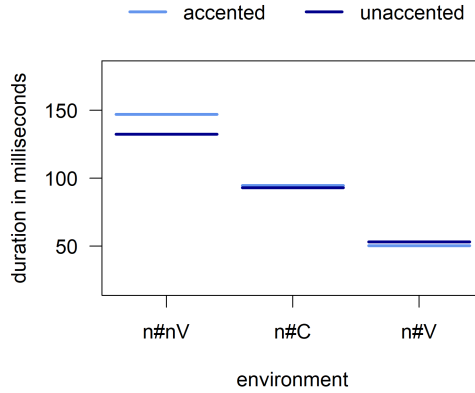
Figure 7.5: Effect of accentuation by environment on consonant duration in complex *un*-data set

position. In both conditions, i.e. in the accented and in the unaccented condition, there is a significant durational difference between doubles (n#nV) and singletons (n#C, n#V). In accented position, doubles are predicted to be 53 ms longer than singletons followed by a consonant, and 97 ms longer than singletons followed by a vowel. In unaccented position, doubles are also predicted to be longer than both types of singletons but the durational differences are smaller. The differences are 39 ms for singletons followed by a consonant (n#C), and 80 ms for singletons followed by a vowel (n#V). The difference between the two singleton levels is roughly the same in both conditions (45 ms in accented position, 40 ms in unaccented position).

The results clearly show that *un-* geminates. Phonological doubles are predicted to be more than twice as long as phonological singletons followed by a vowel, and the predicted singleton-double ratios for consonant adjacent singletons are, depending on accentuation, 1:1.4 and 1:1.6. In comparison to former studies, these durational differences are very large (see sections 2.1 and 6.3.3 for a discussion of durational differences between geminates and their corresponding singletons).

To test the contribution of each variable for the model's goodness of fit, I checked how the absence of each term affects the AIC of the model. Figure 7.6 displays the increase of the model's AIC without each factor. The higher the increase, the more variance is explained by the pertinent factor in the model.

The figure clearly shows that the variable Environment explains most of
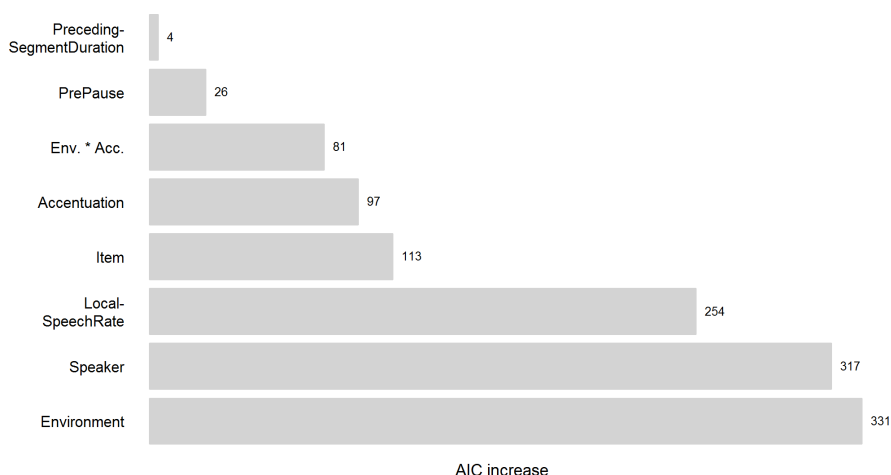
Figure 7.6: AIC increase for each variable of the final *un*-model, AIC final model = -11398

the variation found in the data. In other words, the absence or presence of a phonological double consonant explains a large portion of the durational differences found in the data. Furthermore, there are speaker-dependent differences, i.e. different speakers produce the nasal in *un*-prefixed words with different durations. However, it is important to note that while there are overall differences in the duration of the nasal between speakers, all speakers show the same pattern with regard to gemination, i.e. all of them produced the double consonants with longer durations than the singletons. The third most important variable is LocalSpeechRate. The other noise variables, as well as the interaction, are much less important for the model, i.e. they explain much less of the variance found in the data.

To sum up, the analyses have shown that *un*- clearly geminates, and that the duration of the nasal in *un*-prefixed words is furthermore influenced by phonetic and prosodic factors. The variables Environment and LocalSpeechRate are two of the most important determiners for consonant duration with *un*-. In addition to Speaker, they explain most of the variance in the data. This result fits in well with the findings of the corpus study, where these two variables were the only two significant predictors for nasal duration with *un*-. Furthermore, the two prosodic variables Accentuation and PrePause are rather important predictor variables. The variable PrecedingSegmentDuration, even though significant in the final model, is of less importance.

7.3.3.2 Complete model

The second *un*-model investigates nasal duration in all tokens of the *un*-data set (*N* = 2615), i.e. it investigates nasal duration in prefixed words and in base words. As in the model predicting nasal duration with only complex words, the dependent variable AbsoluteConsonantDuration was Box-Cox-transformed (*λ* = 0.182) and outliers (*N* = 66, i.e. 2.52% of the data) were removed to achieve a normal distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions). The final model is displayed in Table H.2 in appendix H.

The final model features the variables Environment, LocalSpeechRate, BaseInitialStress, Accentuation and PrePause. The two noise variables LocalSpeechRate and BaseInitialStress show the expected effects. The faster the speech rate, the longer the nasal, and when the base-initial syllable is unstressed the nasal is shorter than when the base-initial syllable is stressed. The two other noise variables Accentuation and PrePause interact with the variable of interest Environment. Note that there is no three-way interaction between Accentuation, PrePause and Environment.

Figure 7.7 shows the effect of Accentuation by Environment. The light blue lines represent the estimated durations for words in accented position, the dark blue lines show the estimated durations for words in unaccented position. In both conditions, i.e. in accented and unaccented condition, double consonants are predicted to be significantly longer than all types of singletons. This means *un-* geminates, independent of accentuation.

For doubles (n#nV) and singletons in base words (#nV), nasals in accented words are longer than nasals in unaccented words. For singletons in complex words (n#C and n#V), consonant duration is not affected by accentuation. This has the effect that durational differences between doubles and singletons in complex words are bigger in accented than in unaccented position. The durational difference between doubles and singletons in base words is not affected by accentuation.

Figure 7.8 shows the effect of PrePause by Environment. The light blue lines represent the estimated durations for words with no preceding pause, the dark blue lines show the estimated durations for words with a preceding pause. The figure shows that all doubles are longer than all types of singletons, i.e. *un-* geminates independent of whether a pause precedes a word or not. As with Accentuation, only the two environments n#nV and #nV are affected by the variable PrePause. For items with a double consonants (n#nV), the nasal is longer when a pause precedes the word. For base words, the opposite is the case, i.e. the nasal is shorter after a preceding pause. Crucially, gemination does not depend on
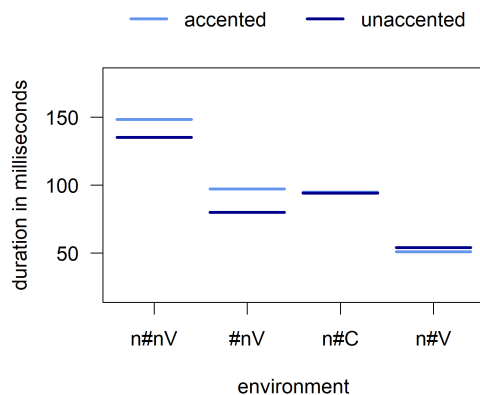
Figure 7.7: Effect of accentuation by environment on consonant duration in complete *un*-data set

PREPAUSE.
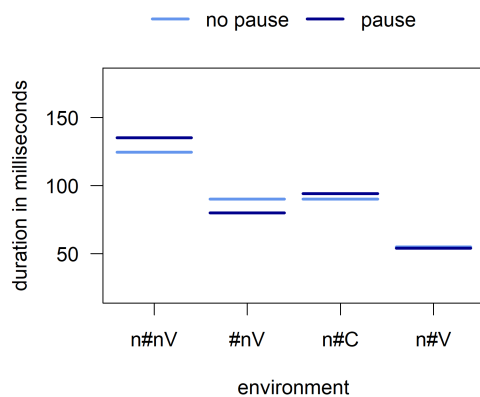
Figure 7.8: Effect of pause before item by environment on consonant duration in complete *un*-data set

7.3.3.3 Summary

The prefix *un-* clearly geminates. Phonological doubles (n#nV) are longer than singletons in complex words (n#C, n#V) and singletons in base words (#nV). There

are also significant durational differences between the singleton levels with the singletons in base words being the longest, the singletons in prefixed words followed by a consonant being the second longest and the singletons in prefixed words followed by a vowel being the shortest. This pattern resembles the one found in the corpus study. The durational differences between doubles and singletons are, however, much bigger in the experimental study than in the corpus study, i.e. gemination seems to be more extreme in the experimental data than in the corpus data. As in the corpus study, the decomposability measure logRELATIVEFREQUENCY does not affect nasal duration with *un-*.

In both *un*-models, the noise variables show the expected effects. In the complex model, the noise variables ACCENTUATION, LOCALSPEECHRATE, PREPAUSE and PRECEDINGSEGMENTDURATION proved to be significant. In the complete model, the noise variables ACCENTUATION, LOCALSPEECHRATE, PREPAUSE and BASEINITIALSTRESS proved to be significant. The noise variable ACCENTUATION interacts with the variable ENVIRONMENT in both models. While doubles and singletons in base words are longer when accented, singletons in base words are not. In the complete model, there is an interaction between PREPAUSE and ENVIRONMENT. Again, only doubles and singletons in base words are affected by this variable. Crucially, whether *un-* geminates or not does not depend on accentuation or a preceding pause.

7.3.4 The prefix *in-*

7.3.4.1 The allomorph /ɪn/: Complex model

The model predicting consonant duration with all complex /ɪn/-words (N = 1232) was fitted according to the modeling procedure described in §5.4. Due to an uneven distribution of the residuals in the initial model, the dependent variable ABSOLUTECONSONANTDURATION was Box-Cox-transformed (λ = 0.061) and 22 outliers were removed (1.9% of the data). After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions). The decomposability variables were tested individually.

The final model features five variables, ENVIRONMENT, BASEINITIALSTRESS, ACCENTUATION, LOCALSPEECHRATE and PRECEDINGSEGMENTDURATION. There are two interactions in the model, one between ENVIRONMENT and BASEINITIALSTRESS, and one between ENVIRONMENT and ACCENTUATION. The final model is summarized in Table H.3 in appendix H.

The two noise variables LOCALSPEECHRATE and PRECEDINGSEGMENTDURATION behave as expected. The higher the speech rate, the shorter the consonant, and the longer the preceding segment, the shorter the nasal.

Figure 7.9 shows the interaction between BASEINITIALSTRESS and ENVIRONMENT. For each environment, the estimated consonant durations for words with a stressed base-initial syllable are indicated by light blue lines, and the estimated durations for words with an unstressed base-initial syllable are indicated by dark blue lines. The plot shows that only when the base-initial syllable of a word is stressed, doubles are predicted to be clearly longer than singletons followed by a vowel (33 ms). When the base-initial syllable of a word is unstressed, doubles (n#nV) are predicted to be only 10 ms longer than singletons followed by a vowel (n#V). Doubles are never predicted to be longer than singletons followed by a consonant (n#C). When the base-initial syllable of a word is stressed, doubles are predicted to be as long as singletons followed by a consonant (n#C). When the base-initial syllable of a word is unstressed, doubles are predicted to be 41 ms shorter than singletons followed by a consonant (n#C).
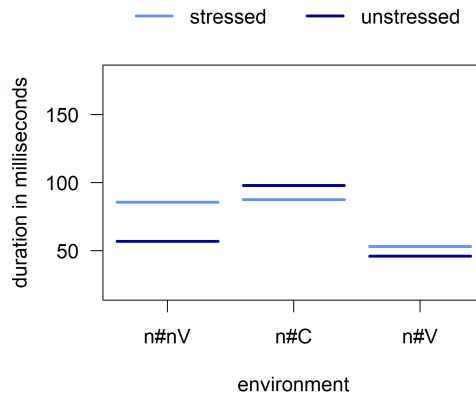
Figure 7.9: Effect of base-initial stress by environment on consonant duration in complex /ɪn/-data set

Figure 7.10 shows the interaction between ACCENTUATION and ENVIRONMENT. Light blue lines represent the estimates for accented items, and dark blue lines represent the estimated for unaccented items. Note that the figure shows the estimates for items with an unstressed base-initial syllable, and that there is no three-way-interaction between BASEINITIALSTRESS, ACCENTUATION and ENVIRONMENT.
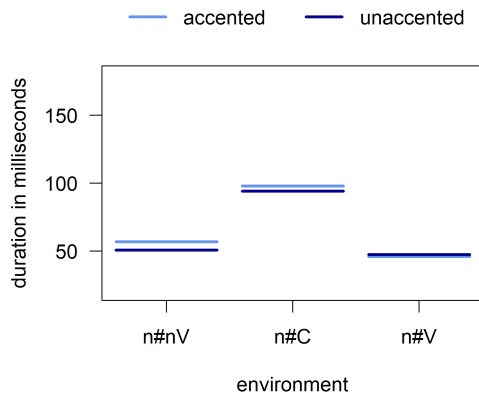
Figure 7.10: Effect of accentuation by environment on consonant duration in complex /ɪn/-data set

The figure shows that while double consonants (n#nV) and singletons followed by a consonant (n#C) are slightly longer when the word is accented, singletons followed by a vowel (n#V) are not affected by ACCENTUATION. Crucially, the durational pattern of the three environments does not change depending on whether a word is accented or unaccented. With unstressed base-initial syllables, singletons followed by a consonant (n#C) are the longest, followed by doubles (n#nV), and singletons followed by a vowel (n#V) are the shortest. With stressed base-initial syllables, doubles (n#nV) and singletons followed by a consonant (n#C) are of the same duration, and singletons followed by a vowel (n#V) are shorter.[8]

The effect size of each significant term in the model is displayed in Figure 7.11. As in the *un*-model, the variables SPEAKER, ENVIRONMENT, LOCALSPEECHRATE and ITEM explain most of the variance found in the data. Crucially, the variable ENVIRONMENT is very important for the model, i.e. the number of consonants at the morphological boundary, as well as the segment following the nasal, highly influence the duration of the nasal. In contrast to *un*-, ENVIRONMENT does, however, not explain most of the variance found in the data. It thus seems that it is less important for predicting consonant duration with /ɪn/-prefixed words than with *un*-prefixed words.

None of the decomposability measures proved to be significant in the final model. On the one hand, this might mean that decomposability does not affect

---

[8]Note that the durational differences for words with stressed base-initial syllables cannot be seen in Figure 7.10.

Draft of May 27, 2019, 18:04

consonant duration with /ɪn/. On the other, there is the possibility that the individual decomposability measures are not strong enough to reach significance on their own. As a combined decomposability measure, they might, however, become significant in the model (see also discussions in sections 7.3.1 and 6.3.4).
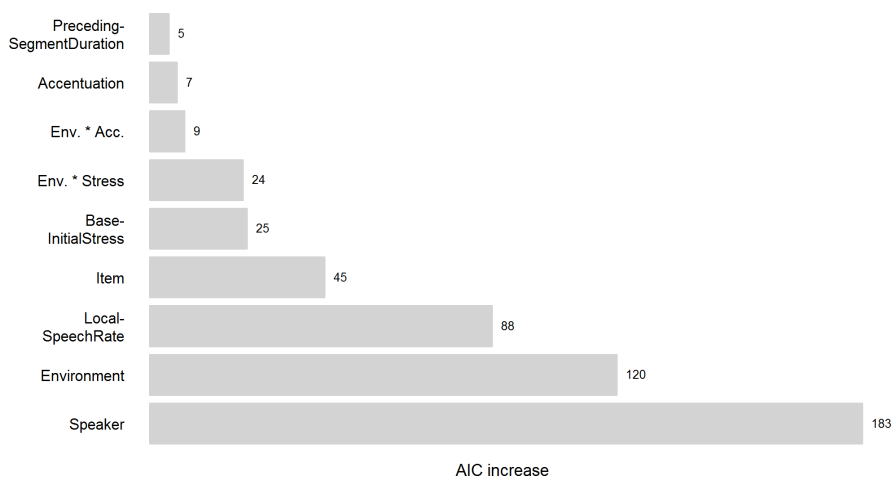
Figure 7.11: AIC increase for each variable of the final /ɪn/-model, AIC final model = -6974

To test this idea, an additional model with combined decomposability measures (as opposed to individual decomposability measures) was fitted. The combined measures were created by means of a principal component analysis (see §5.4 on principal component analyses). As in the corpus study, the principal component analysis was fitted with the variables logRELATIVEFREQUENCY, SEMANTIC-TRANSPARENCY, SEMANTICTRANSPARENCYRATING, TYPEOFBASE and AFFIX. The variable AFFIX features the two levels inLoc (locative *in-*) and inNeg (negative *in-*). All variables were recoded into numerical variables and scaled before the analysis was conducted. Table 7.13 summarizes the analysis by showing the composition of each principal component, i.e. the loading of each variable for each principal component, and by displaying the proportion of variance covered by each component.

The principal component analysis revealed that the first three components can account for most of the variance expressed by the five decomposability variables (90%). An inspection of the rotation matrix shows that the first component is composed of all five measures, that the second is mainly dominated by the variables scaledAFFIX and scaledRELATIVEFREQUENCY, and that the third is mainly dominated by the variable scaledSEMANTICTRANSPARENCYRATING. The first three

principal components were included as predictor variables in the mixed model.

Table 7.13: Summary of principal components

Composition of principal components					
	PC1	PC2	PC3	PC4	PC5
scaledAFFIX	-0.457	0.591	0.022	-0.324	-0.580
scaledRELATIVEFREQUENCY	-0.390	-0.724	-0.347	-0.391	-0.225
scaledSEMANTICTRANSPARENCYRATING	-0.391	-0.252	0.882	0.036	0.065
scaledTYPEOFBASE	-0.469	-0.037	-0.242	0.836	-0.145
scaledSEMANTICTRANSPARENCY	-0.516	0.250	-0.206	-0.205	0.767
Variance explained by principal components					
	PC1	PC2	PC3	PC4	PC5
Proportion of Variance	0.626	0.150	0.121	0.076	0.027

The model was fitted similarly to the model with the individual decomposability measures. After model simplification, none of the principal components remained in the model. This means, the simplification of the model resulted in the same final model as the simplification of the model with the individual decomposability measures. Decomposability does not affect consonant duration with /ɪn/.

To sum up, the analyses have shown that prefixal nasal duration in /ɪn/- prefixed words is influenced by the variables ENVIRONMENT, BASEINITIALSTRESS, ACCENTUATION, LOCALSPEECHRATE and PRECEDINGSEGMENTDURATION. None of the decomposability variables influences nasal duration with /ɪn/, neither as independent measures nor as combined measures. While the noise variables ACCENTUATION, LOCALSPEECHRATE and PRECEDINGSEGMENTDURATION behave as expected and do not influence the gemination pattern of /ɪn/, the variable BASEINITIALSTRESS affects gemination with /ɪn/. Only when the base-initial syllable of a prefixed word is stressed, double consonants (n#nV) are predicted to be clearly longer than singletons with a following vowel (n#V). In this case, i.e. when the base-initial syllable is stressed, doubles are predicted to be as long as singletons followed by a consonant (n#C). When the base-initial syllable of a word is unstressed, doubles (n#nV) are predicted to be only marginally longer than singletons with a following vowel (n#V). In this case, they are predicted to be shorter than singletons followed by a consonant (n#C).

7.3.4.2 The allomorph /ɪn/: Complete model

The model predicting consonant duration with all /ɪn/-words (*N* = 1232) was fitted according to the modeling procedure described in §5.4. Due to an uneven distribution of the residuals in the initial model, the dependent variable ABSOLUTE-CONSONANTDURATION was Box-Cox-transformed (λ = 0.202) and 22 outliers were removed (1.79% of the data). After the transformation, the model showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions).

The final model features the five variables LOCALSPEECHRATE, ENVIRONMENT, BASEINITIALSTRESS, ACCENTUATION and PREPAUSE (see Table H.4 in appendix H for a summary of the final complete model). The noise variable LOCALSPEECHRATE shows the expected effect. The other three noise variables interact with the variable ENVIRONMENT, i.e. there are three two-way interactions in the model, one between BASEINITIALSTRESS and ENVIRONMENT, one between ACCENTUATION and ENVIRONMENT, and one between PREPAUSE and ENVIRONMENT. Note that all pertinent three-way interactions were tested but none proved to be significant.
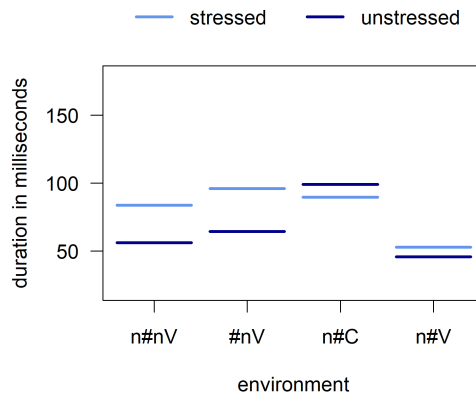
Figure 7.12: Effect of base-initial stress by environment on consonant duration in /ɪn/-data set

Figure 7.12 shows the effect of BASEINITIALSTRESS on ENVIRONMENT. Light blue lines indicate estimates for words with stressed base-initial syllables, and dark blue lines indicate estimates for words with unstressed base-initial syllables. The plot shows the predicted durations for accented items.

As the complex /ɪn/-model, the complete model reveals that stress affects gem-

ination with /ɪn/. The plot shows that only when doubles (n#nV) are part of a word with a stressed base-initial syllable, they are as long as singletons with a following consonant (n#C) and longer than singletons with a following vowel (n#V). When part of a word with an unstressed base-initial syllable, doubles are shorter than singletons followed by a consonant and only slightly longer than singletons followed by a vowel. Independent of BASEINITIALSTRESS, doubles are slightly shorter than singletons in base words.

Figure 7.13 shows the interaction between ACCENTUATION and ENVIRONMENT. Estimates for accented items are indicated by light blue lines, estimates for unaccented items by dark blue lines. The figure shows the estimates for words with an unstressed base-initial syllable.
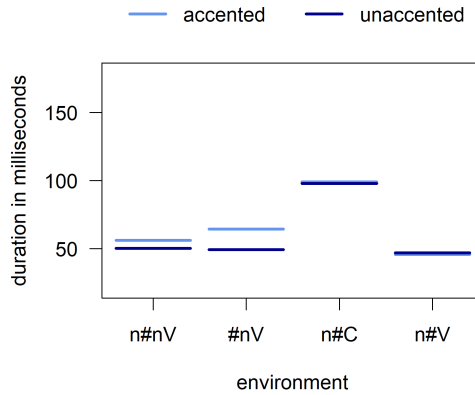
Figure 7.13: Effect of accentuation by environment on consonant duration in /ɪn/-data set

The plot shows that only the durational pattern of doubles and singletons in base words is affected by accentuation. In unaccented condition, singletons in base words and doubles are of the same duration. In accented condition, singletons in base words are longer. Crucially, doubles are never longer than singletons in base words. The durational pattern of doubles and singletons in complex words is not affected by accentuation. For words with an unstressed base-initial syllable, doubles are shorter than singletons followed by a consonant, and slightly longer than singletons followed by a vowel. For words with a stressed base-initial syllable, doubles are as long as singletons followed by a consonant and longer than singletons followed by a vowel.[9]

---

[9]Note that the durational differences for words with stressed base-initial syllables cannot be

Figure 7.14 shows the interaction between PrePause and Environment. Light blue lines indicate the estimated nasal durations for words with no preceding pause, and dark blue lines show the estimated nasal durations for words with a preceding pause. The figure shows that only the duration of singletons in base words (#nV) is affected by a preceding pause. When a pause precedes a base word, nasals are shorter than when there is no preceding pause. This effect was also observed with *un-*, and is therefore not surprising. The effect of PrePause does not affect gemination with /ɪn/.
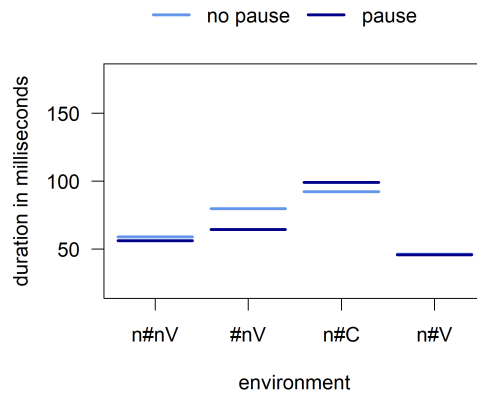
Figure 7.14: Effect of pause before item by environment on consonant duration in /ɪn/-data set

As with the complex model, the results of the complete /ɪn/-model show that only when the base-initial syllable of a complex word is stressed, phonological doubles with /ɪn/ are longer than singletons in complex words followed by a vowel. Phonological doubles are never longer than singletons in complex words followed by a consonant or singletons in base words. This suggests that gemination with *in-* depends on the stress-pattern of the prefixed word. Only when the base-initial syllable of a double-consonant word is stressed, the prefix geminates. It, furthermore, suggests that gemination with *in-* is weaker than gemination with *un-*.

In contrast to morphological geminates with *un-*, morphological geminates with /ɪn/ are not longer than all types of singletons, i.e. they are not longer than singletons in base words and they are not longer than singletons in prefixed

---

seen in Figure 7.13.

words followed by a consonant. That doubles with /ɪn/ are not longer than single-tons followed by a consonant could be attributed to a lengthening effect caused by the following consonant. It is expected that nasals followed by a consonant are longer than nasals followed by a vowel (see, for example, Umeda 1977, see also discussion in §5.5.1). Thus, singletons followed by a consonant might only be as long as doubles because they are lengthened by the following consonant. However, for *un-* this lengthening effect did not result in a lack of durational difference between doubles and singletons followed by a consonant, but only in smaller durational differences between doubles and singletons.

That *un-* geminates to a higher degree than *in-* is also suggested by the fact that the durational differences between doubles and singletons for /ɪn/ are much smaller than the ones for *un-* (singleton-geminate ratio for *un-*: 1:3.0, singleton-geminate ratio for /ɪn/: 1:1.6). The AIC increase analysis also supports the idea that gemination is less strong with /ɪn/. It revealed that the variable Environment is of less importance in the /ɪn/-model than in the *un*-model.

However, before coming to a conclusion with regard to the gemination pattern of *in-*, one must take a few additional facts into consideration. First, there are only four types of /ɪn/-prefixed words with a double consonant in the data set (*innervate*, *innocuous*, *innominate*, *innumerable*). Three of them feature a stressed base-initial syllable. This means that there is only one /ɪn/-prefixed type that does not geminate. It seems rather bold to make generalizations based on this one type. Second, as /ɪn/ is only one of several allomorphs of the prefix *in-*, and as the gemination of the prefix *in-* is expected to follow similar patterns across allomorphs, a final conclusion with regard to the gemination pattern of *in-* can only be drawn after more empirical facts are revealed, i.e. after the results of the /ɪm/ -models are discussed.

7.3.4.3 The allomorph /ɪm/: Complex model

The model predicting consonant duration with all complex /ɪm/ -words (N = 1177) was fitted according to the modeling procedure described in §5.4. Because the initial model showed a non-normal distribution, 24 outliers (2.04% of the data) were removed and the dependent variable was Box-Cox-transformed (λ = 0.101). After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions). The decomposability variables were tested individually.

The final model features five variables, Environment, BaseInitialStress, Accentuation, LocalSpeechRate and GlobalSpeechRate. There are two interac-

tions in the model, one between BASEINITIALSTRESS and ACCENTUATION, and one between BASEINITIALSTRESS and ENVIRONMENT. The model is shown in Table H.5 in appendix H.

The two noise variables LOCALSPEECHRATE and GLOBALSPEECHRATE show the expected effects. With increasing speech rate, the nasal becomes shorter. The other two noise variables BASEINITIALSTRESS and ACCENTUATION interact. In accented position, nasals are longer when the base-initial syllable of a word is stressed. In unaccented position, the opposite is the case, i.e. words with an unstressed base-initial syllable feature longer nasals than words with a stressed base-initial syllable.

As in the /m/-models, the variable of interest ENVIRONMENT interacts with BASEINITIALSTRESS. The effect is shown in Figure 7.15. The light blue lines represent the estimated nasal durations for items with a stressed base-initial syllable, the dark blue lines represent the estimated durations for items with an unstressed base-initial syllable. The figure clearly shows that only when the base-initial syllable is stressed, doubles are longer than singletons. When the base-initial syllable is unstressed, doubles are shorter than singletons. Thus, as with /m/, gemination seems to depend on the stress status of the base-initial syllable of a word. Only when it is stressed, the prefix *in-* geminates.
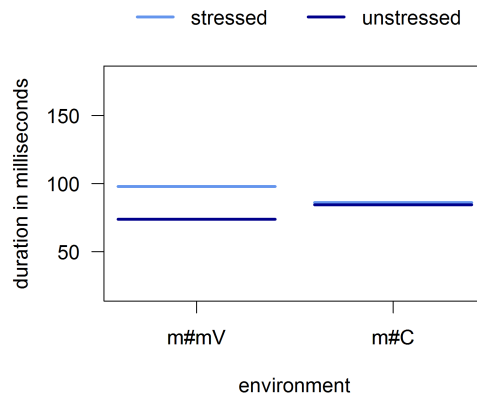
Figure 7.15: Effect of base-initial stress by environment on consonant duration in complex /m/-data set

The durational differences between doubles and singletons are rather small. When the base-initial syllable is stressed, doubles are 12 ms longer than singletons. When the base-initial syllable is unstressed, doubles are 10 ms shorter than

singletons. One could explain these small differences by referring to the difference in the following segment between doubles and singletons in the /ɪm/ -data set. Doubles are always followed by a vowel, and singletons are always followed by a consonant. As shown in the corpus and the experimental study with *un-* and /ɪn/, following vowels shorten the preceding nasal, i.e. the double, and following consonants lengthen the preceding nasal, i.e. the singleton. One might thus suspect that if one kept the environment constant for doubles and singletons, the durational difference between doubles and singletons in words with a stressed base-initial syllable would be larger, and doubles in words with an unstressed base-initial syllable would not be shorter than singletons.

However, if one compares the results with the ones of the /ɪm/ -corpus study, one can see that the durational difference between doubles and singletons followed by a consonant was much larger in the corpus study (27 ms). Furthermore, in the corpus study gemination was not dependent on the stress status of the base-initial syllable. It thus seems that gemination with /ɪm/ differs between corpus and experimental data. In the corpus, gemination with /ɪm/ is stronger and independent of prosodic factors, whereas in the experimental data gemination is weaker and depends on base-initial stress.
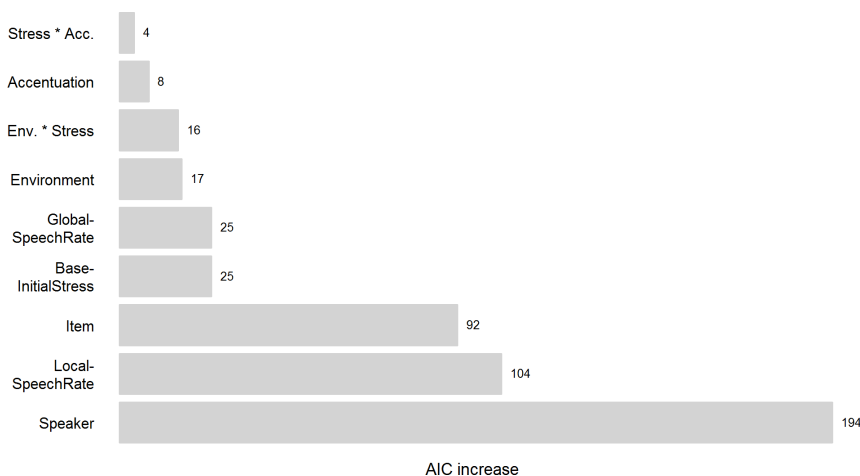
Figure 7.16: AIC increase for each variable of the final /ɪm/-model, AIC final model = -6389

To test the relative importance of each term in the final model, I looked at each term's contribution to the AIC of the final model. Figure 7.16 displays the increase of the model's AIC without each term. The figure shows that the noise variables SPEAKER, LOCALSPEECHRATE and ITEM explain most of the variation found in

the data. They are followed by the two noise variables GLOBALSPEECHRATE and BASEINITIALSTRESS. The variable of interest ENVIRONMENT explains much less of the variance found in the data. The analysis of the AIC increase thus shows that most of the variation in the data is explained by noise variables, i.e. not by the variable ENVIRONMENT. This supports the idea that gemination with /ɪm/ is weaker in the experimental study than in the corpus study, and that, as proposed in the previous section, gemination with the prefix *in-* is weaker than gemination with the prefix *un-*.

None of the decomposability measures showed a significant effect in the final model when tested individually. To test possible effects of a combined decomposability measure, an additional mixed model with combined decomposability measures was fitted to the data set. To attain these combined decomposability measures, I fitted a principal components analysis to the scaled variables scaledAFFIX, scaledRELATIVEFREQUENCY, scaledSEMANTICTRANSPARENCYRATING, scaledTYPE-OFBASE and scaledSEMANTICTRANSPARENCY.

Table 7.14 summarizes the principal components analysis. In the upper part of the table the loadings of each principal component are shown. The lower part of the table displays the proportion of variance accounted for by each component. Most of the variance is explained by the first principal component. The second, third and fourth component explain much less of the variance and the fifth hardly any. The first four components were tested in the model.

Table 7.14: Summary of principal components

Composition of principal components					
	PC1	PC2	PC3	PC4	PC5
scaledAFFIX	0.437	-0.407	0.154	-0.785	0.056
scaledRELATIVEFREQUENCY	0.314	0.808	0.407	-0.147	0.247
scaledSEMANTICTRANSPARENCYRATING	0.431	0.275	-0.851	-0.076	-0.090
scaledTYPEOFBASE	0.526	-0.070	0.291	0.335	-0.722
scaledSEMANTICTRANSPARENCY	0.497	-0.317	0.038	0.494	0.637
Variance explained by principal components					
	PC1	PC2	PC3	PC4	PC5
Proportion of Variance	0.542	0.185	0.117	0.102	0.054

This first component is composed of all five measures. This can be as seen by its loadings which are roughly the same for all variables. The second component

is dominated by the variables scaledRelativeFrequency and scaledAffix, the third by scaledSemanticTransparencyRating and the fourth by scaledAffix, scaledTypeOfBase and scaledSemanticTransparency.

The model with the principal components was fitted similarly to the model with the individual decomposability measures. After I simplified the model, it turned out that all terms which are significant in the final model with the individual decomposability measures are also significant in the final model with the principal components. On top of the effects significant in the model with the individual measures, the principal component model shows an effect of the fourth principal component (see Table H.6 in appendix H for a summary of the final model). Figure 7.17 shows the effect of PC4.
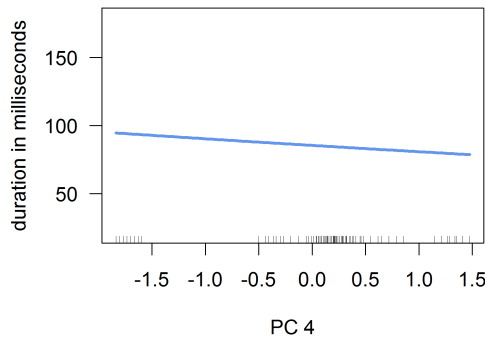
Figure 7.17: Effect of PC4 on consonant duration in complex /ɪm/-data set

The figure shows that with an increasing PC4-value, the nasal in /ɪn/-prefixed words becomes shorter. The size of the effect is, however, rather small. As described above, PC4 is mostly composed of the variables Affix, SemanticTransparency and TypeOfBase. As indicated by the loadings shown in Table 7.14, the component negatively correlates with Affix, meaning that a higher PC4-value represents negative *in-*, and a lower value represents locative *in-*. For SemanticTransparency and TypeOfBase, the component shows positive loadings. This means that a higher PC4-value represents opaque derivatives with a bound base, and that a lower PC4-value represents transparent derivatives with words as bases. The effect of PC4 on nasal duration can thus be interpreted as follows: nasals in opaque negative *in*-prefixed words with a bound base tend to be shorter

than nasals in transparent locative *in*-prefixed words with words as bases. This interpretation is supported by the fact that all tokens with a PC4-value lower than –1.5 are locative *in*-prefixed words with a word as a base and transparent meaning (e.g. *implant, imprison*). It is yet unclear why these words feature a particularly long nasal. It seems though that PC4 reflects an effect on prefixal consonant duration that is restricted to a small number of types with a particular feature combination. This feature combination does not appear to directly translate to decomposability. Instead, it might be related to the existence of two different locative *in*-prefixes.

As discussed in §3.1.4, there might be two distinct locative *in*-prefixes: native and non-native locative *in*-. The two types of locative *in*- are argued to differ in their origin, the type of base they take and their productivity. The items with a low PC4-value seem to represent items with native locative *in*-, i.e. locative *in*-items with a transparent meaning and a native and free base. (e.g. *implant, imprison*). It remains unclear why they feature particularly long nasals.

To summarize, doubles with the allomorph /ɪm/ are only longer than singletons in complex words when the base-initial syllable of the derivative is stressed. This is similar to what was found with the allomorph /ɪn/. The singleton-geminate ratio for /ɪm/ in the experimental study is smaller than the one found in the corpus study, and also smaller than the one found for *un*-. The model which tested combined decomposability measures furthermore shows a significant effect of one principal component on nasal duration with /ɪm/ . However, the effect is relatively small and not clearly interpretable in terms of decomposability. Crucially, the effect does not affect gemination.

7.3.4.4 The allomorph /ɪm/: Complete model

The model predicting consonant duration with all /ɪm/ -words (*N* = 1635) was fitted according to the modeling procedure described in §5.4. Because the initial model showed a non-normal distribution, 37 outliers (2.63% of the data) were removed and the dependent variable was Box-Cox transformed (λ = 0.343). After the transformation, the model showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions).

The final model features the five variables Environment, BaseInitialStress, PrePause, LocalSpeechRate and GlobalSpeechRate (see Table H.7 in appendix H for a summary of the final model). Both speech rates show the expected effects. With increasing speech rates, the nasal becomes shorter. The three variables Environment, BaseInitialStress, and Pause form a three-way interaction. The

interaction is shown in Figure 7.18.
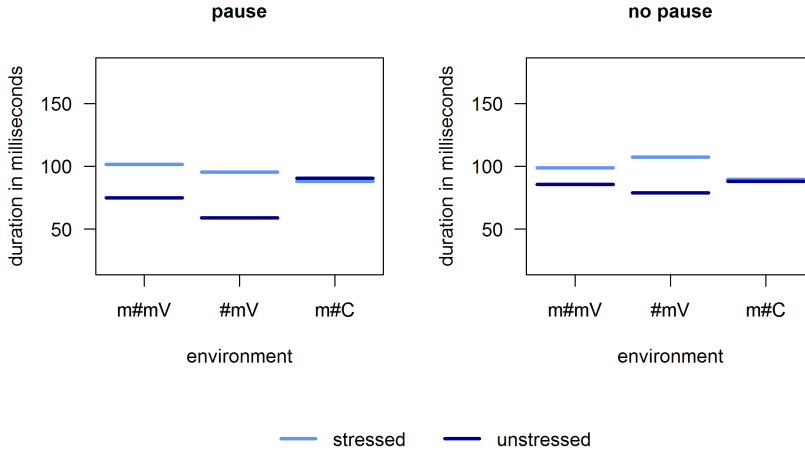
Figure 7.18: Effect of base-initial stress by environment on consonant duration in /ɪm/-data set

The left panel of the figure shows the effect of BASEINITIALSTRESS by ENVIRONMENT for items produced with a pause before the item. The right panel shows the effect for items produced without a preceding pause. For each environment, light blue lines indicate the durations for items with a stressed base-initial syllable, and dark blue lines indicate the durations for items with an unstressed base-initial syllable.

In the left panel, i.e. in items with a preceding pause, doubles (m#mV) are longer than singletons in base words (#mV). This is independent of the stress status of the base-initial syllable. Doubles are, however, only longer than singletons in complex words (m#C) when the base-initial syllable is stressed. In the right panel, i.e. in items without a preceding pause, doubles (m#mV) are only longer than singletons in base words (#mV) when the base-initial syllable is unstressed. In this case, they are as long as singletons in complex words (m#C). For items with a stressed base-initial syllable, doubles (m#mV) are shorter than singletons in base words (#mV) but longer than singletons in complex words (m#C).

To summarize, independent from a preceding pause, doubles in words with a stressed base-initial syllable are predicted to be longer than singletons in complex words. Depending on the absence or presence of a preceding pause, doubles in

words with an unstressed base-initial syllable are either as long as or shorter than singletons in complex words. They are never predicted to be longer than singletons in complex words. Except for when the double consonant is part of a word which is not preceded by a pause and which features a stressed base-initial syllable, doubles are longer than singletons in base words with the same condition.

7.3.4.5 Summary

In all *in*-models, the noise variables behaved as expected. Decomposability does not seem to influence nasal duration with *in-*. Only in one model, one decomposability measure affected nasal duration (PC4). Its effect was very weak and not clearly interpretable in terms of decomposability. It did not affect gemination.

For both allomorphs of *in-*, the analyses have shown that doubles are only longer than singletons when the base-initial syllable of the derivative is stressed. These doubles are, however, only longer than some types of singletons, and durational differences between doubles and singletons are often rather small. In case of /ɪn/, doubles in words with a stressed base-initial syllable are only longer than singletons in complex words with a following vowel. Doubles are not predicted to be longer than singletons in complex words with a following consonant or singletons in base words. For /ɪm/ , doubles in words with a stressed base-initial syllable are longer than singletons with a following consonant but the durational difference between doubles and singletons is rather small. Furthermore, doubles are slightly longer than singletons in base words.

The results suggest that gemination with *in-* depends on stress and that the degree of gemination with *in-* is rather weak. Only when the base-initial syllable of a word is stressed, doubles are longer than some types of singletons, i.e. for /ɪn/, they are longer than singletons in complex words followed by a vowel, and for /ɪm/ , they are longer than singletons in complex words followed by a consonant. Durational differences between doubles and singletons are rather small. This result is quite different from what was found for *un-*. For *un-*, doubles are longer than all types of singletons and the durational differences between doubles and singletons are much larger. One can thus state that the prefix *in-* geminates, but that it geminates to a lesser degree than the prefix *un-*.

Interestingly, the gemination pattern found for *in-* in the experimental study deviates from the one found in the corpus study. In the corpus study, *in-* geminates independent of stress, durational differences between doubles and singletons are bigger, and the degree of gemination is the same as for *un-*. In other

words, gemination with *in-* is weaker in the experimental study than in the corpus study.

One possible explanation for the difference between corpus and experimental data is that the degree of semantic processing differs between the two types of investigated speech. It can be hypothesized that in natural, conversational speech, the semantic processing of words is deeper than in read speech, and that processing depth, in turn, affects gemination. With deeper processing, the meaning of the affix is more present in the production of the derivative, leading to less reduction, i.e. to gemination. In the corpus study, the meaning of the prefix is deeply processed, and it therefore clearly geminates. In the experimental study, in contrast, the affix's meaning is not processed deeply. In turn, gemination is weaker and governed by a non-semantic factor, i.e. the prosodic factor stress. This explanation is supported by the fact that the variable AFFIX (locative *in-* vs. negative *in-*) only affects consonant duration with *in-* in the corpus study, not in the experimental study. The meaning of the affix only affects consonant duration in the corpus data, where it is semantically processed. It does not affect consonant duration in the experimental data, where no deep semantic processing takes place.

7.3.5 The prefixes *un-* and *in-*

The model predicting consonant duration with all complex *un-* and /ɪn/-words (*N* = 3237) was fitted to directly compare gemination with *un-* and *in-*, and to further investigate the observed differences between the affixes. As already discussed in §6.3.5, in a model which investigates both prefixes, the decomposability variables cannot be tested in an interesting way. This is because *un-*, as described in sections 6.2.2 and 7.2, does not vary in most of the decomposability measures, and because relative frequency measures are not well comparable across *un-* and *in-*. The prefix *in-* has very many bound roots, which is problematic with regard to computing relative frequency measures that are comparable to the relative frequency measures of affixes with hardly any or no bound roots, i.e. in this case *un-*. Therefore, none of the decomposability measures was tested in the model.

The model was fitted according to the modeling procedure described in §5.4. Due to an uneven distribution of the residuals in the initial model, the dependent variable ABSOLUTECONSONANTDURATION was Box-Cox-transformed (λ = 0.061) and 67 outliers were removed (2.07% of the data). The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions).

The final model features the five variables ENVIRONMENT, AFFIX, PREPAUSE, ACCENTUATION and LOCALSPEECHRATE. The variable LOCALSPEECHRATE has the

expected effect. The three variables ENVIRONMENT, AFFIX and PREPAUSE interact. Furthermore, there is an interaction between ENVIRONMENT and ACCENTUATION. The final model is summarized in Table H.8 in appendix H.

Figure 7.19 shows the interaction between ENVIRONMENT and ACCENTUATION. Estimates for accented items are shown in light blue, and estimates for unaccented items are shown in dark blue. The figure shows that, independent of accentuation, doubles (n#nV) are clearly longer than singletons (n#C, n#V). In accented position, doubles are even longer. This has the effect that the durational difference between doubles and singletons increases in accented condition.
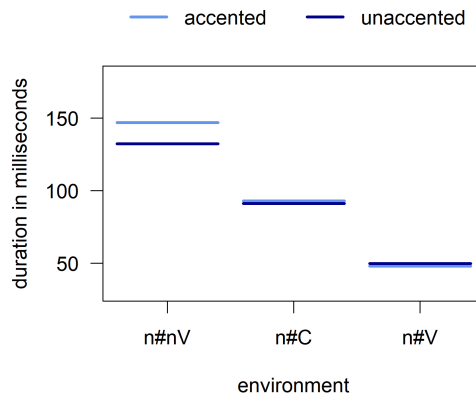
Figure 7.19: Effect of accentuation by environment on consonant duration in complex *un*- and /ɪn/-words

However, to really interpret the gemination pattern of the affixes, one must consider the three-way interaction between ENVIRONMENT, AFFIX and PREPAUSE. The interaction is displayed in Figure 7.20. In the left panel of the figure, the estimates for items with a preceding pause are shown. In the right panel of the figure, the estimates for items without a preceding pause are shown. For each affix, estimates for double consonants are shown in light blue, estimates for singletons followed by a consonant are shown in green, and estimates for singletons followed by a vowel are shown in dark blue.

In both panels of the figure, it is clearly visible that while the durations of the singletons do not differ much across affixes, the durations of the double consonants differ a lot across affixes. The double consonant in *un*-prefixed words is much longer than the double consonant in both *in*-prefixed words. The double consonant in negative *in*-prefixed words is longer than the one in locative
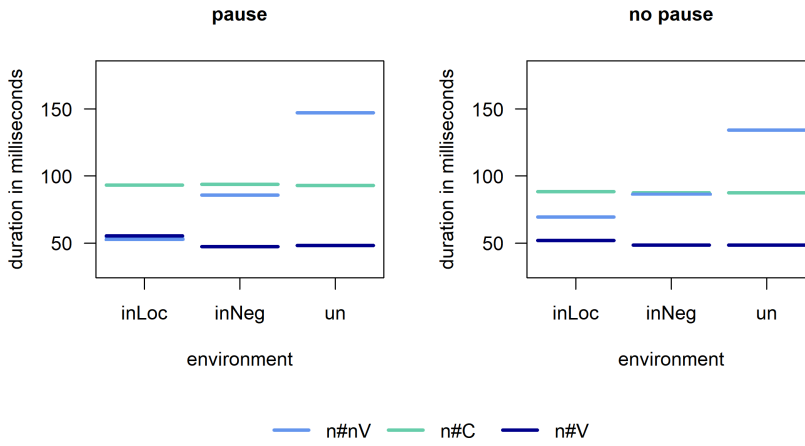
Figure 7.20: Effect of environment by affix on consonant duration in complex *un-* and /ɪn/-words with and without a preceding pause

in-prefixed words. As a consequence, there are differences in gemination between the affixes. The prefix *un-* clearly geminates, i.e. doubles are clearly longer than both singleton levels. For negative *in-*, gemination is weaker. Doubles are longer than singletons with a following vowel (n#V), and about as long as singletons with a following consonant (n#C). The double consonant in locative *in*-words is the shortest. It is only longer than singletons which are followed by a vowel (n#V) and preceded by a pause. It is debatable whether this difference between doubles and singletons with locative *in-* can be interpreted as gemination at all.

At first glance, the results suggest that the degree of gemination decreases from *un-* to negative *in-*, to locative *in-*. However, one must be cautions with this interpretation. As already mentioned before, there are only four /ɪn/-prefixed types with a double consonant in the data set. Only one of these types features locative *in-*, i.e. the conclusion that locative *in-* geminates to a lesser degree than negative *in-* is based on only one type. Furthermore, the type featuring locative *in-* is the only type with an unstressed base-initial syllable in the data set. It might thus be the case that the difference between locative and negative *in-* is actually caused by a difference in the prosodic structure of the words, i.e. by a difference in the stress status of the base-initial syllable. There are two arguments for this explanation. First, the analysis of all *in*-data has already shown that gemination

Draft of May 27, 2019, 18:04

with *in-* depends on the stress status of the base-initial syllable. The effect of stress was observed for both allomorphs of *in-*. Only when the base-initial syllable is stressed, *in-* geminates. Second, in none of the *in-*analyses an effect of Affix was found.

One can conclude that there is a difference in the degree of gemination between *un-* and the *in-*prefixes: gemination with *un-* is stronger than gemination with *in-*. This might be explained with the fact that *un-* is more segmentable and more informative than both *in-*prefixes. However, there is no clear evidence for different degrees of gemination between locative *in-* and negative *in-*, even though the two *in-*prefixes differ in segmentability and informativeness. Gemination with *in-* depends on prosodic structure.

This result is different from what was found in the corpus study. In the corpus study, gemination was the same across *un-* and *in-* but there was a general durational difference between the affixes. The prefix *un-* featured the longest nasal, followed by negative *in-*. Locative *in-* featured the shortest nasal. Even though the results differ, they can be interpreted similarly: the more segmentable the affix, or the more informative, the less reduction. In the corpus study, reduction affected singletons and doubles. In the experimental study, only doubles were affected and there was no distinction between negative and locative *in-*.

7.3.6 The prefix *dis-*

7.3.6.1 Complex model

The model predicting consonant duration with all complex *dis-*words ($N = 829$) was fitted according to the modeling procedure described in §5.4. Due to an uneven distribution of the residuals in the initial model, the dependent variable AbsoluteConsonantDuration was Box-Cox-transformed ($\lambda = 0.263$) and 16 outliers were removed (1.9% of the data). After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions). The decomposability variables were tested individually.

The final model features three variables, Environment, LocalSpeechRate and Accentuation. The variable LocalSpeechRate behaves as expected: the higher the speech rate, the shorter the consonant. The variables Environment and Accentuation form an interaction. The final model is summarized in Table H.9 in appendix H.

The interaction between Environment and Accentuation is depicted in Fig-
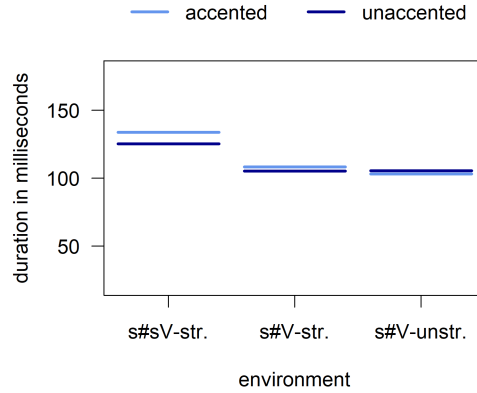
Figure 7.21: Effect of accentuation by environment on consonant duration in complex *dis*-data set

ure 7.21. For each environment, the estimates for accented items are shown by light blue lines, and the estimates for unaccented items are shown by dark blue lines. The figure shows that, independent from accentuation, doubles (s#sV-str.) are longer than both types of singletons in complex words (s#V-str., s#V-unstr.) The durational differences between doubles and singletons are larger in the accented than in the unaccented condition. In accented condition, doubles are 25 ms longer than singletons with a stressed base-initial syllable and 30 ms longer than singletons with an unstressed base-initial syllable. In unaccented condition, doubles are 20 ms longer than both types of singletons. The durational difference between the two singleton levels is not significant.

As doubles are longer than both types of singletons, one can state that *dis-* geminates. The durational differences between doubles and singletons are smaller than the ones found for *un-*. They are in approximately the same range as the ones for /ɪn/. This suggests that *dis-* geminates to a similar degree as *in-*.

Figure 7.22 shows the contribution of each variable to the final model's goodness of fit. The figure clearly shows that the variable Speaker explains most of the variance found in the data. The variable Environment, i.e. the crucial variable with regard to gemination, explains much less of the variance. This fits in well with the interpretation that the prefix *dis-* geminates, but that it does not geminate to a high degree.

None of the individual decomposability variables proved to be significant in the final complex *dis*-model. To check whether a combined decomposability mea-
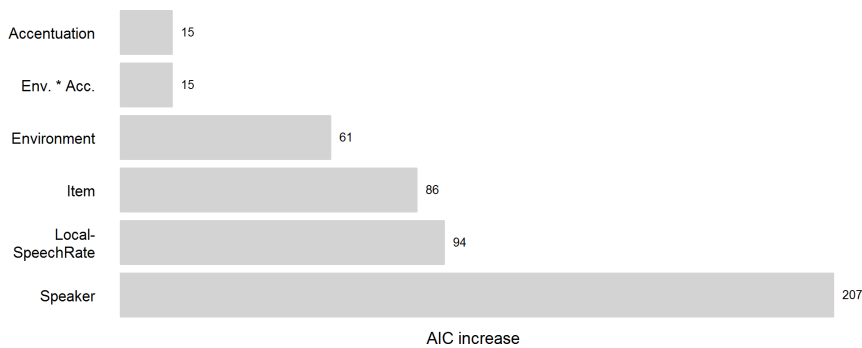
Figure 7.22: AIC increase for each variable of the final *dis*-model, AIC final model = -4031

sure affects consonant duration with *dis*-, I conducted a principal component analysis and tested the effect of the principal components in an additional model. The principal component analysis included the four variables logRELATIVEFRE-QUENCY, SEMANTICTRANSPARENCYRATING, TYPEOFBASE and SEMANTICTRANSPA-RENCY. Categorical variables were recoded as numerical, and all variables were scaled. Table 7.15 summarizes the principal components.

Table 7.15: Summary of principal components

Composition of principal components				
	PC1	PC2	PC3	PC4
scaledRELATIVEFREQUENCY	-0.465	0.817	0.344	0.005
scaledSEMANTICTRANSPARENCYRATING	-0.489	-0.559	0.670	0.002
scaledTYPEOFBASE	-0.523	-0.098	-0.462	-0.710
scaledSEMANTICTRANSPARENCY	-0.522	-0.104	-0.469	0.705
Variance explained by principal components				
	PC1	PC2	PC3	PC4
Proportion of Variance	0.644	0.148	0.112	0.097

Most of the variance is accounted for by the first component, the second and the third component explain much less variance, and the last principal component explains barely any variance. The first principal component is composed of all decomposability measures. The second component is mainly dominated

by logRᴇʟᴀᴛɪᴠᴇFʀᴇǫᴜᴇɴᴄʏ, the third component mostly represents TʏᴘᴇOꜰBᴀsᴇ, SᴇᴍᴀɴᴛɪᴄTʀᴀɴsᴘᴀʀᴇɴᴄʏRᴀᴛɪɴɢ and SᴇᴍᴀɴᴛɪᴄTʀᴀɴsᴘᴀʀᴇɴᴄʏ, and the fourth is mostly composed of TʏᴘᴇOꜰBᴀsᴇ and SᴇᴍᴀɴᴛɪᴄTʀᴀɴsᴘᴀʀᴇɴᴄʏ. The first three principal components were included in the model.

The model was fitted similarly to the model with the individual decomposability measures. After model simplification, none of the principal components remained in the model. This means, the simplification of the model resulted in the same final model as the simplification of the model with the individual decomposability measures. Decomposability does not affect consonant duration with *dis-*.

7.3.6.2 Complete model

The model predicting consonant duration with all *dis*-words (*N* = 1114) was fitted according to the modeling procedure described in §5.4. Due to an uneven distribution of the residuals in the initial model, the dependent variable AʙsᴏʟᴜᴛᴇCᴏɴsᴏɴᴀɴᴛDᴜʀᴀᴛɪᴏɴ was Box-Cox-transformed (λ = 0.343) and 24 outliers were removed (2.15% of the data). After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions).

The final model features four variables: LᴏᴄᴀʟSᴘᴇᴇᴄʜRᴀᴛᴇ, Eɴᴠɪʀᴏɴᴍᴇɴᴛ, Aᴄᴄᴇɴᴛᴜᴀᴛɪᴏɴ and PʀᴇPᴀᴜsᴇ. The two noise variables PʀᴇPᴀᴜsᴇ and LᴏᴄᴀʟSᴘᴇᴇᴄʜRᴀᴛᴇ behave as expected. With increased speech rate, the fricative becomes longer, and the fricative is longer when a pause precedes the item. The two variables Eɴᴠɪʀᴏɴᴍᴇɴᴛ and Aᴄᴄᴇɴᴛᴜᴀᴛɪᴏɴ interact. The final model is summarized in Table H.10 in appendix H.

Figure 7.23 shows the interaction between Eɴᴠɪʀᴏɴᴍᴇɴᴛ and Aᴄᴄᴇɴᴛᴜᴀᴛɪᴏɴ. For each environment, the estimates for accented items are indicated by light blue lines, and the estimates for unaccented items are indicated by dark blue lines. The plot reveals that doubles are longer than singletons in complex words (`s#V-unstr.`, `s#V-str.`) and singletons in simplex words (`sV-unstr.`). In accented condition, the durational differences between doubles and these two types of singletons is larger than in unaccented condition. Independent from accentuation, doubles are shorter than singletons in base words (`#sV-str.`).

The figure, furthermore, indicates that singletons in simplex words (`sV-unstr.`) are not significantly longer than singletons in complex words (`s#V-unstr.`, `s#V-str.`). This, and the fact that they are shorter than phonological doubles, shows that gemination with *dis-* is not an orthographic phenomenon but a morpho-
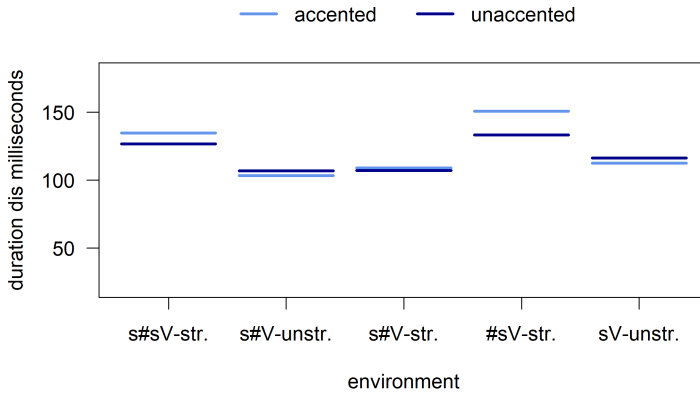
Figure 7.23: Effect of accentuation by environment on consonant duration in *dis*-data set

phonological one. In other words, phonological doubles are longer than phonological singletons because of the presence of two underlying identical consonants, not because of the presence of two identical graphemes.

7.3.6.3 Summary

Both *dis*-models have shown the expected effects of the noise variables. With regard to the variables of interest, the complex model has revealed that decomposability does not affect consonant duration with *dis*-. The variable Environ-ment is significant in all models. Its effect shows that *dis*- geminates. The double in *dis*-prefixed words is longer than the singleton in complex words and the singleton in simplex words. In both models, there is an interaction between Envi-ronment and Accentuation, indicating that gemination is stronger when the double consonant word is accented. This is similar to what was found for *un*-.

That singletons in simplex words are shorter than phonological doubles shows that gemination does not depend on orthography but is a morpho-phonological phenomenon. If the lengthening of the phonological double was caused by its orthography, i.e. by the fact that it is spelled with two graphemes, the singleton in simplex words, which is also represented by an orthographic double, should also be lengthened. This is not the case.

The data suggests that the degree of gemination with *dis*- is weaker than the degree of gemination with *un*-. Gemination with *dis*- seems to be similar in its

degree to gemination with *in-*. This is indicated by the durational differences between doubles and singletons for *dis-*. They are smaller than the ones found for *un-* and similar to the ones found for /m/. Furthermore, in contrast to what was found for *un-*, and similar to what was found for *in-*, doubles with *dis-* are only longer than some types of corresponding singletons, i.e. doubles are not longer than singletons in base words.

A comparison of the durations of the experimental study with the ones of the corpus study reveals that gemination with *dis-* is weaker in the experimental study than in the corpus study. Differences between doubles and singletons are larger in the corpus study than in the experimental study. The same pattern was observed for the prefix *in-*.

7.3.7 The suffix *-ly*

7.3.7.1 Complex model

The model predicting consonant duration with all complex *-ly*-words (N = 1205) was fitted according to the modeling procedure described in §5.4. Due to an uneven distribution of the residuals in the initial model, the dependent variable AbsoluteConsonantDuration was Box-Cox-transformed (λ = 0.263) and 27 outliers were removed (2.24% of the data). After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions).

The final model features three variables of interest: Environment, logRelativeFrequency and SemanticTransparencyRating. It features five noise variables: LocalSpeechRate, TypeOfL, PostPause, Accentuation and PrecedingSegmentDuration. There are two interactions in the model, one between Environment and logRelativeFrequency, and one between Environment and Accentuation.

Note that there is no suppression effect with the two decomposability variables in the model, i.e. the effects of SemanticTransparencyRating and logRelativeFrequency do not negatively affect each other in the model. Fitting the final model with only one of the two decomposability variables at a time, furthermore, revealed that their effect sizes do not change much in the presence of the other. Their effects are thus interpretable. The final model is summarized in Table H.11 in appendix H.

The noise variables LocalSpeechRate, TypeOfL, PostPause and PrecedingSegmentDuration behave as expected. The higher the speech rate, the shorter

the lateral.; a tap /l/ is shorter than an approximant /l/; items which are followed by a pause feature a longer /l/ than items with no following pause; and with increased preceding segment duration, the duration of the lateral becomes shorter.

The interaction between ENVIRONMENT and ACCENTUATION is shown in Figure 7.24. For each environment, the estimates for accented items are shown by light blue lines, and the estimates for unaccented items are shown by dark blue lines. The estimated duration for singletons in complex words (#l-<l>) is shown on the left, the estimated durations for the three double consonant environments l#l-<lel>, l#l-<ll> and syll.l#l-<ll> are shown on the right. If *-ly* geminates, the estimated duration for the singleton should be shorter than the estimated durations for the three double environments.
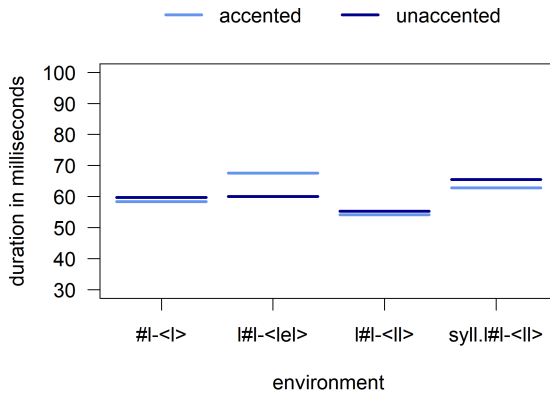
Figure 7.24: Effect of accentuation by environment on consonant duration in complex *-ly*-data set

For items in accented position, doubles of the l#l-<ll>-environment feature the shortest durations of all environments. This means that singletons (#l-<l>) are estimated to have longer durations than this type of double consonant. This clearly speaks against gemination with *-ly*. Doubles of the l#l-<lel>-environment feature the longest durations of all environments, and syllabic doubles (syll.l#l-<ll>) pattern in between singletons (#l-<l>) and doubles of the l#l-<lel>- environment.

For items in unaccented position, doubles of the l#l-<ll>-environment again feature the shortest durations, i.e. they are estimated to be shorter than the singletons in the data set. Syllabic doubles (syll.l#l-<ll>) feature the longest lateral in this condition. Singletons in complex words (#-<l>) and doubles of the

l#l-<lel>- environment pattern in between.

Crucially, in both conditions, i.e. accented and unaccented, singletons are not estimated to be shorter than all three types of double consonants. In fact, they are consistently estimated to be longer than doubles of the l#l-<ll>-environment. This speaks against gemination with *-ly*.

One could, however, argue that the fact that doubles of the l#l-<lel>- environment and doubles of the syll.l#l-<ll>-environment are longer than singletons speaks for gemination with *-ly*. However, there is no evidence that these durational differences between doubles and singletons are caused by the presence of two underlying laterals. If that was the case, doubles of the l#l-<ll>-environment would also be longer than singletons. Instead, the durational difference between singletons and doubles of the l#l-<lel>-environment and doubles of the syll. l#l-<ll> -environment can be attributed to the factors syllabicity and orthography. It can be assumed that doubles of the syll.l#l-<ll>-environment are longer than singletons because they are syllabic, and that doubles of the l#l-<lel>-environment are longer because they are spelled with the orthographic sequence <lel>. Overall, there is no evidence for gemination with the suffix *-ly*.
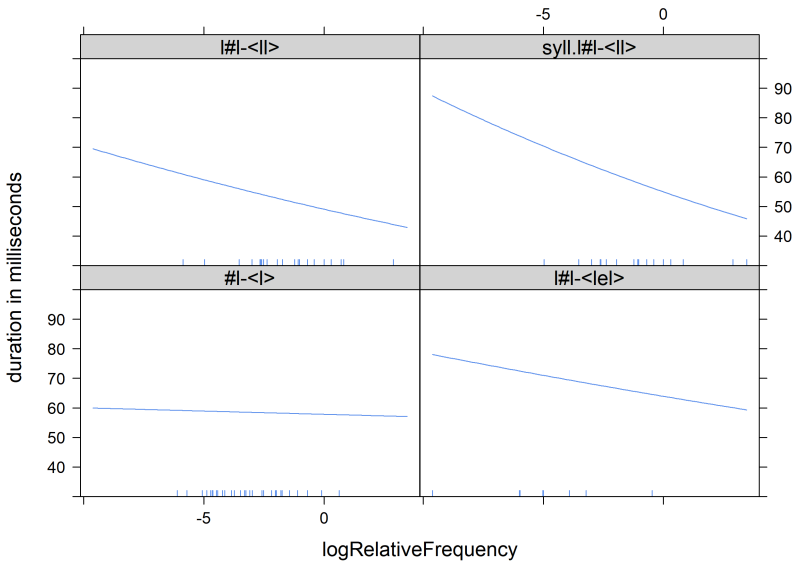
Figure 7.25: Effect of relative frequency by environment on consonant duration in complex *ly*-data set

The variable Environment also forms an interaction with the variable log- Re-lativeFrequency. Figure 7.25 shows the interaction. For each environment, the effect of relative frequency is shown. The figure suggests that for the three double environments (`syll.l#l- <ll>`, `l#l-<ll>`, `l#l-<lel>`), consonant duration de-creases with increasing relative frequency. The higher the relative frequency, i.e. the less decomposable a word, the shorter the /l/. For the singleton environment, the predicted value does not change depending on relative frequency (`#l-<l>`).

However, it is very important to note that the effect of relative frequency is only significant for syllabic doubles (`syll.l#l-<ll>`). The effect is not significant for the other two double environments (`l#l-<ll>`, `l#l-<lel>`). Furthermore, it is unclear how trustworthy the effect for the syllabic doubles really is. Looking at the rugs in the panel for the syllabic doubles, it becomes obvious that for a large portion of the predicted frequency range no observations exist. There is no item with a logRelativeFrequency below −5. Most types in the data set feature a logRelativeFrequency between −5 and −0.9, and there are only two types with a very high relative frequency, i.e. a relative frequency above 2.5. It can be assumed that the observed relative frequency effect is caused by these two types (*aerobically* and *therapeutically*).

To conclude, even though there seem to be tendencies for a relative frequency effect on the duration of the two double environments `l#l-<ll>` and `l#l-<lel>`, and even though there is a significant effect of relative frequency for syllabic doubles (`syll.l#l-<ll>`), the model does not provide convincing evidence for an effect of relative frequency on the duration of phonological doubles.

Figure 7.26 shows the effect of the decomposability measure SemanticTrans-parencyRating. The figure shows that with a higher rating, i.e. with decreasing decomposability, the lateral in *-ly*-suffixed words becomes shorter. This effect is expected. However, the size of the effect is quite small, i.e. durational differences are minimal.

To sum up, the complex model does not provide evidence for gemination with *-ly*. Doubles are not systematically longer than singletons. The model shows two significant effects of decomposability. Syllabic doubles with a low relative fre-quency are predicted to be longer than syllabic doubles with a high relative fre-quency, and items which are rated as highly decomposable are predicted to fea-ture a longer lateral than items which are rated as less decomposable. However, both effects are rather weak and the effect of relative frequency might be caused by only a few types in the data set.

The analysis of the AIC increase for each term in the model supports the analy-sis that *-ly* does not geminate. As can be seen in Figure 7.27, the AIC of the model
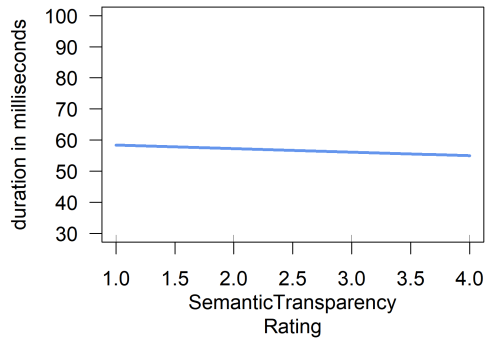
Figure 7.26: Effect of semantic transparency rating on consonant duration in complex *-ly*-data set
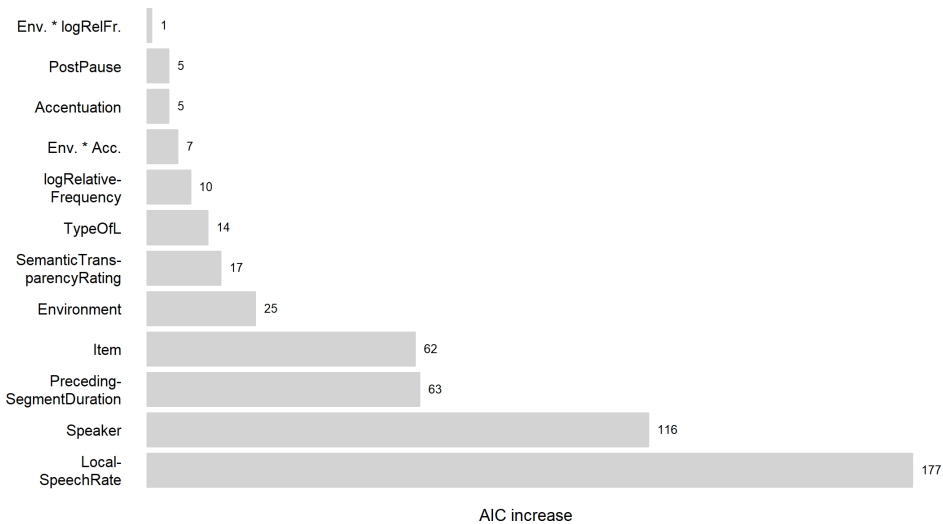
Figure 7.27: AIC increase for each variable of the final *-ly*-model, AIC final model = -4915

increases by only 25 if the variable Environment is taken out of the model, i.e. the variable does not explain much of the variation in the model. The figure furthermore shows that the decomposability measures do not explain much variance either. One should therefore be cautious to not over-interpret their effect. Most of the variance in the complex *-ly*-data is explained by noise variables, i.e. by LocalSpeechRate, Speaker, PrecedingSegmentDuration and Item. This fits in well with the corpus results. In the corpus study, all durational differences between *-ly*-suffixed words were explained by noise variables, i.e. no effect of Environment was found.

7.3.7.2 Complete model

The complex model suggests that *-ly* does not geminate. The complete model was fitted to provide some further evidence for this conclusion. Phonological doubles in words like *really*, *educationally* and *solely* were compared to phonological singletons in simplex words which are represented by orthographic doubles, such as the /l/ in *belly*, and to phonological singletons in base words, such as the /l/ in *real*, *educational* and *sole*. If *-ly* does not geminate, as indicated by the complex model, phonological singletons in simplex words should be as long as phonological doubles. Furthermore, singletons in base words should be longer than double consonants. This is because word-final consonants are usually longer than word-internal consonants (see, for example, Berkovits 1993; Oller 1973; Umeda 1977).

The model with all *-ly*-words (*N* = 1645) was fitted according to the modeling procedure described in §5.4. Due to an uneven distribution of the residuals in the initial model, the dependent variable AbsoluteConsonantDuration was Box-Cox-transformed (λ = 0.061) and 30 outliers were removed (1.82% of the data). After the model was refitted with the transformed dependent variable, it showed a satisfactory distribution of residuals. The model was then simplified and interactions were tested (see appendix G for a list of all tested interactions).

The final model features seven variables: Environment, logWordFormFrequency, LocalSpeechRate, TypeOfL, PostPause, Accentuation and PrecedingSegmentDuration. There are two interactions in the model, one between Environment and PostPause, and one between Environment and Accentuation. The final model is summarized in Table H.12 in appendix H.

The noise variables LocalSpeechRate, TypeOfL, and PrecedingSegmentDuration behave as expected. The higher the speech rate, the shorter the lateral. A tap /l/ is shorter than an approximant /l/. And, the duration of the lateral becomes shorter when the duration of the preceding segment increases. The variable logWordFormFrequency is only marginally significant in the model

but shows the expected effect: with increasing frequency, the lateral becomes shorter.

Figure 7.28 shows the effect of ACCENTUATION by ENVIRONMENT. For each environment, the predicted consonant duration for items in accented position is indicated by light blue lines, and the predicted consonant duration for items in unaccented position is indicated by dark blue lines. For convenience, the estimates for the four different structures, i.e. phonological singletons in complex words (#l-<l>), phonological doubles in complex words (l#l-<lel>, l#l-<ll>, syll.l#l-<ll>), phonological singletons in base words (l#-<le>, l#-<l>, syll. l#-<ll>), and phonological singletons in simplex words (l-<l>), are separated by vertical lines in the figure.
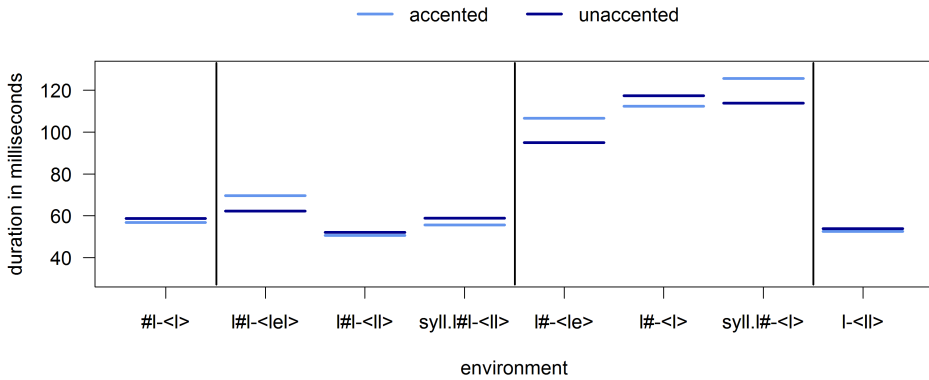
Figure 7.28: Effect of accentuation by environment on consonant duration in complete *-ly*-data set

The figure shows the expected pattern. The durational differences between singletons in complex words and doubles in complex words resemble the ones of the complex model (cf. Figure 7.24 in the previous section). Furthermore, the laterals in base words are predicted to be longer than the laterals in all other environments. This is independent of accentuation. As expected, singletons in simplex words pattern with the phonological doubles: they are predicted to be slightly shorter than doubles of the l#l-<lel>-environment, slightly longer than doubles of the l#l-<ll>-environment, and as long as syllabic doubles (syll.l#-<l>). This is independent of accentuation.

Figure 7.29 shows the effect of POSTPAUSE by ENVIRONMENT. For each environment, the predicted consonant duration for items without a following pause

is indicated by light blue lines, and the predicted consonant duration for items with a following pause is indicated by dark blue lines.
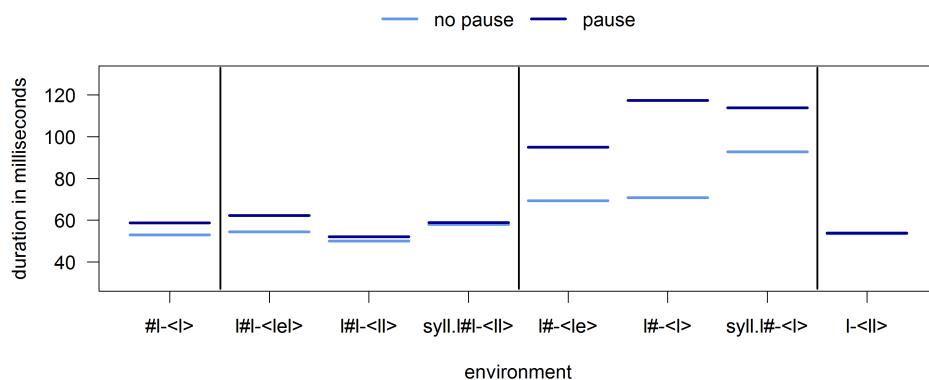
Figure 7.29: Effect of post pause by environment on consonant duration in complete *-ly*-data set

Overall, the figure shows that same durational pattern as Figure 7.24. The three base environments (l#-<l>, syll.l#-<l>, l#-<le>) clearly feature the longest lateral, and there are only minor durational differences between the lateral durations of the other environments. With a following pause, the base-final /l/ becomes even longer, and the durational differences between the base-environments and all other environments increase. As only in base words /l/ is immediately followed by the pause, this is expected.

To sum up, the complete model has supported the result of the complex model. The suffix *-ly* does not geminate. The double consonant is shorter than the word-final consonant in base words, and about as long as the singleton /l/ represented by an orthographic double. There is thus no indication that two underlying /l/s in *-ly*-suffixed words are realized with a longer duration than one underlying /l/.

7.3.7.3 Summary

In both *-ly*-models, the noise variables Accentuation, LocalSpeechRate, PrecedingSegmentDuration, TypeOfL and PostPause have shown expected effects. Additionally, in the complete model, the variable logWordFormFrequency affected consonant duration in the expected direction.

Both models have revealed that phonological doubles in *-ly*-suffixed words are not systematically longer than phonological singletons. Only under certain

conditions, are some types of phonological double consonants, i.e. syllabic ones and phonological doubles represented by the orthographic string <lel>, longer than some types of phonological singletons. The longer duration of the doubles in those cases can be attributed to their syllabicity and their orthography, not to the fact that they feature two underlying consonants. One can conclude that the suffix *-ly* degeminates, and that syllabicity and orthography affect the acoustic realization of /l/ in *-ly*-suffixed words.

The complex model revealed effects of decomposability with *-ly*. Items which were rated as less decomposable were produced with shorter consonant durations. Furthermore, for the syllabic double consonants, relative frequency affected consonant duration. With a higher relative frequency, i.e. with less decomposable words, the consonant becomes shorter. These effects are in line with the assumption that less decomposable units are reduced, while more decomposable units are not reduced. However, the effect sizes are quite small and the effect of relative frequency might be caused by just a few types in the data set.

It is important to note that gemination with *-ly* does not depend on relative frequency. In other words, it is not the case that words with high relative frequency degeminate and words with low relative frequency geminate, or vice versa. If that was the case, relative frequency would be significant for all double consonant environments. This is not the case. The effect of relative frequency is independent of gemination.

7.3.8 Summary

The first durational analyses looked at the distribution of duration across environments to get a first impression of whether the affixes under investigation geminate, and if so, whether gemination is a gradient or a categorical phenomenon. For all affixes, the analysis revealed that if there is a durational difference between doubles and singletons, durations are bimodally distributed with the doubles being longer than the singletons. As discussed in §4.3.1 (*Nature of gemination: Predictions*), this indicates that gemination is a categorical phenomenon.

For all data sets two or more linear models were fitted. One model predicted consonant duration with only complex words, and one model predicted consonant duration with all words of the data set, i.e. complex words, base words and simplex words with an orthographic double. Both models revealed very similar results, i.e. for the most part the same variables are significant in both models. Table 7.16 shows an overview of the variables which show significant effects on absolute consonant duration in the subsets. Only variables which are significant in at least one of the models, as independent effects or as part of an interaction,

are listed.

Table 7.16: Overview of significant variables in experimental models

Variable	un-	in-	im-	un- & in-	dis-	-ly
Environment	✓	✓	✓	✓	✓	✓
Affix	-	n.s.	n.s.	✓	-	-
SemanticTransparencyRating	n.s.	n.s.	n.s.	-	n.s.	✓
logRelativeFrequency	n.s.	n.s.	n.s.	-	n.s.	✓
PC4	-	-	✓	-	-	-
LocalSpeechRate	✓	✓	✓	✓	✓	✓
Accentuation	✓	✓	✓	✓	✓	✓
PrePause	✓	✓	✓	✓	✓	n.s.
BaseInitialStress	✓	✓	✓	n.s.	-	-
PrecedingSegmentDuration	✓	✓	n.s.	n.s.	n.s.	✓
GlobalSpeechRate	n.s.	n.s.	✓	n.s.	n.s.	n.s.
PostPause	n.s.	n.s.	n.s.	n.s.	n.s.	✓
logWordFormFrequency	n.s.	n.s.	n.s.	n.s.	n.s.	✓
TypeOfL	-	-	-	-	-	✓

✓ : significant in at least one of the models

n.s. : not significant in the models

-   : not included in the models

Overall, the noise variables showed the expected effects in all models. As can be seen in the table, the variables LocalSpeechRate and Accentuation are significant with all affixes. Some variables, such as PrePause, only affect consonant duration with the prefixes, and some, such as PostPause, only affect consonant duration with the suffix *-ly*. Furthermore, there are some noise variables which only affect consonant duration in a few models, such as the variable GlobalSpeechRate or the variable PrecedingSegmentDuration.

With regard to the variables of interest, three of the decomposability variables proved to be significant in the models: PC4 for /ɪm/ , and SemanticTransparencyRating and logRelativeFrequency for *-ly*. All effects are, however, quite weak and do not allow for general conclusions about the effect of decomposability.

As discussed in §7.3.4, the effect of PC4 in the /ɪm/ -model is very weak and caused by only a few types with a particular feature combination. The effect is therefore not clearly interpretable in terms of decomposability. For *-ly*, the two decomposability variables SEMANTICTRANSPARENCYRATING and logRELATIVEFRE-QUENCY indicate that with increasing decomposability, /l/ in *-ly*-suffixed words becomes longer. This is in line with the assumption that less decomposable units are reduced, while more decomposable units are not reduced. However, the effect sizes are quite small and the effect of relative frequency is only significant for one environment, i.e. syllabic doubles. Furthermore, the effect might be caused by only two types in the data set.

The variable of interest ENVIRONMENT significantly affects consonant duration in all models. However, the effect of ENVIRONMENT does not indicate gemination for all affixes. Only for the prefixes, phonological doubles are systematically longer than phonological singletons, i.e. only the prefixes geminate. Double consonants with *-ly* are not systematically longer than corresponding singletons. The suffix *-ly* degeminates.

While all prefixes geminate, there are differences in the degree of gemination between them. While the prefix *un-* clearly and strongly geminates, the degree of gemination is much weaker for the two *in-*prefixes and for *dis-*. The weaker degree of gemination is indicated by four aspects: first, durational differences between doubles and singletons are smaller for *in-* and *dis-* than for *un-*.[10] Second, while the double consonant in *un-*prefixed words is longer than all types of singletons, the double consonant in *in-* and *dis-*prefixed words is not. For example, the double in *in-* and *dis-*prefixed words is not longer than the singleton in base words. Third, the variable ENVIRONMENT explains much less of the variance found in the *in-* and *dis-*data than in the *un-*data. Fourth, while gemination with *un-* is independent of prosodic factors, gemination with *in-* depends on stress. Only when the base-initial syllable of an *in-*prefixed word is stressed, the word geminates. For *dis-*, an interaction between BASEINITIALSTRESS and ENVIRONMENT could not be tested because of the distribution of stress in the data set. The experimental study did not feature a *dis-*prefixed word with an unstressed base-initial syllable and a phonological double.

It is important to note that, in contrast to the experimental study, the corpus study featured one *dis-*prefixed type with a phonological double and an unstressed base-initial syllable (*dissolution*). Interestingly, this was the only type which degeminated. As discussed in §6.3.8, it remained unclear whether the degem-

---

[10]See Table I.1 in appendix I for an overview of the predicted consonant durations of prefixed words in the experimental study.

ination of *dissolution* was caused by type-specific factors, its unstressed base-initial syllable or its semantic opacity. Since the experimental study has shown that gemination does not depend on semantic transparency, i.e. that semantically opaque *dis*-prefixed words geminate, it now seems plausible to assume that the degemination of *dissolution* was not caused by its semantic opacity but by its unstressed base-initial syllable. This means one could also assume that other *dis*-prefixed words with a double consonant and an unstressed base-initial syllable also degeminate. Due to the non-existence of additional relevant types, i.e. additional *dis*-prefixed words with a double consonant and an unstressed base-initial syllable, it is however impossible to further investigate the issue.

Further evidence for the conclusion that *in-* geminates to a lesser degree than *un-* can be gleaned from the model which directly compared the prefixes. In this model, the variable AFFIX interacts with the variable ENVIRONMENT. While the singletons of all three prefixes (*un-*, locative *in-* and negative *in-*) are of comparable length, there is a significant difference in the duration of double consonants between the prefixes. The affix *un-* features the longest double nasal, and the double nasal with negative *in-* is longer than the one with locative *in-*. This decline in duration resembles the decline in segmentability of the affixes. The prefix *un-* is the most segmentable affix of the three, followed by negative *in-*, followed by locative *in-*. However, as discussed in §7.3.4, the only type featuring a double consonant and locative *in-* is also the only /ɪn/-prefixed item with a double consonant featuring an unstressed base-initial syllable. It is therefore possible that the difference in gemination between locative and negative *in-* is actually caused by a difference in prosodic structure. This explanation is especially plausible as the variable AFFIX is not significant in any of the *in*-models, whereas the variable BASEINITIALSTRESS interacts with the variable ENVIRONMENT in all of the *in*-models. One can thus state that there is a difference in the degree of gemination between *un-* and *in-*, but that there is no difference in the degree of gemination between locative and negative *in-*.

The experimental study revealed one other important fact with regard to the nature of gemination. Gemination does not depend on orthography. As evidenced by the complete *dis-* and *-ly*-models, orthographic doubles in simplex words (e.g. *dissertation*, *belly*) are not longer than orthographic singletons. In the *dis*-model, it was furthermore revealed that orthographic doubles in simplex words are shorter than phonological doubles. Thus, the lengthening of phonological doubles is not caused by their orthography. Gemination in English is caused by the presence of two identical phonological consonants, not by the presence of two identical orthographic consonants.

To summarize, in all models the noise variables showed the expected effects. Furthermore, effects of word-specific decomposability were found for *-ly*. With regard to gemination, the experimental study revealed that the prefixes *un-*, *in-* and *dis-* geminate. The suffix *-ly* degeminates. The prefix *un-* geminates to a higher degree than the two *in*-prefixes and *dis-*. The pattern of gemination in the experimental study thus resembles the segmentability and informativeness of the affixes (see *Semantic Segmentability Hierarchy* in Figure 7.3 in §7.2.2). The most informative and segmentable affix *un-* clearly geminates, the least informative and least segmentable affix *-ly* degeminates, and the other affixes geminate to a rather low degree. Gemination does not depend on spelling.

8 Summary and Discussion

In this book, I have investigated gemination with the five English affixes *un-*, negative *in-*, locative *in-*, *dis-* and *-ly*. I conducted a corpus and an experimental study. Table 8.1 summarizes the main results of both studies by displaying the gemination pattern of each affix. Note that the gemination pattern of locative and negative *in-* is not displayed separately as there were no differences in gemination between the two affixes.

By investigating gemination with the different affixes, the predictions of various theories of the morpho-phonological and the morpho-phonetic interface were tested. On the one hand, I tested predictions of formal linguistic theories. On the other, I tested predictions which were deduced from psycholinguistic approaches of the morpho-phonological-phonetic interface. The psycholinguistic predictions center around two factors: decomposability and morphological informativeness. Below I will summarize and discuss the main findings of this book.

8.1 Decomposability

Decomposability is one of the main factors whose effect on gemination, and on the acoustic realization of words in general, was tested in this project. On the one hand, categorical decomposability in terms of the overall segmentability of the affix was investigated. On the other, the effects of word-specific decomposability measures were tested. Before testing the effect of the overall segmentability of an affix, it was necessary to find out how segmentable each affix is. This was done by first deriving segmentability hierarchies from the theoretical literature, and by then validating these hierarchies in the corpus and the experimental study. Below I will first give a summary of the segmentability analyses of the affixes. Then, I will summarize the effects of decomposability found in both studies.

8.1.1 The segmentability of the affixes

In Chapter 3, two segmentability hierarchies, which order the five investigated affixes in terms of their segmentability, were derived from the theoretical litera-

Table 8.1: Overview of gemination in corpus and experimental study

		Corpus	Experiment	Overall
un-	Doubles longer than singletons	yes	yes	Gemination■
	Durational difference	large	very large	
in-	Doubles longer than singletons	yes	mostly, stress-dependent■	■ Gemination
	Durational difference	large	small	
dis-	Doubles longer than singletons	mostly	mostly	Gemination■
	Durational difference	small	small	
-ly	Doubles longer than singletons	no	mostly not	Degemination■
	Durational difference	none	small	

ture. They are shown below in (1) and (2).

(1) Non-Semantic Segmentability
Hierarchy: *un-* > *-ly* > {*dis-*, *in-*$_{\neg\text{NEG}}$} > *in-*$_{\text{LOC}}$

(2) Semantic Segmentability
Hierarchy: *un-* > {*dis-*, *in-*$_{\neg\text{NEG}}$} > *in-*$_{\text{LOC}}$ > *-ly*

The two hierarchies differ with regard to the definition of decomposability they are based on. In the Non-Semantic Segmentability Hierarchy, decomposability is defined in terms of productivity, transparency and the type of base an affix takes. In the Semantic Segmentability Hierarchy, decomposability is defined in terms of the semantics of the affix, i.e. an affix with an independent clear

meaning is more segmentable than an affix without a clear meaning. Note that, as discussed in §4.3.2, the Semantic Segmentability Hierarchy does not only capture the segmentability of the affix but simultaneously represents its informativeness. An affix with a clear lexical meaning is more informative than an affix without a clear meaning.

The comparison of the affix's segmentability in both studies indicates that the two theoretically derived segmentability hierarchies are borne out by the data. The corpus study showed the same segmentability pattern across all decomposability measures. The prefix *un-* and the suffix *-ly* are the most segmentable affixes, followed by *dis-* and negative *in-*, and locative *in-* is the least segmentable affix. This pattern matches the Non-Semantic Segmentability Hierarchy.

In the experimental study, the distribution of the decomposability ratings revealed a similar picture. The participants of the experimental study rated *un-* as the most segmentable affix, locative *in-* as the least segmentable affix, and negative *in-* and *dis-* pattern in between. However, in contrast to the corpus study, the suffix *-ly* is rated as the second least segmentable affix after locative *in-*. This pattern resembles the Semantic Segmentability Hierarchy in that *-ly* is one of the least segmentable affixes.

To sum up, the data shows that the five investigated affixes differ in their segmentability. The prefix *un-* is very segmentable and very informative, the prefixes negative *in-* and *dis-* are less segmentable and informative, and locative *in-* is the least segmentable and informative prefix. The segmentability of the suffix *-ly* largely depends on the definition of decomposability. While it is very segmentable in terms of its productivity, its transparency and the type of base it takes, it is less segmentable in terms of its semantics. It is the least informative affix of the set of investigated affixes. The two segmentability hierarchies capture these findings. While the segmentability pattern of the prefixes is the same in both hierarchies, the position of *-ly* differs between the two hierarchies.

8.1.2 Effects on the acoustic realization of words

The studies revealed two types of decomposability effects on acoustic duration: categorical segmentability effects of the affix and gradient word-specific decomposability effects. Both types of effects go in the same direction: the more segmentable a unit is, the longer it is, i.e. the less reduced it is. That there are both categorical and gradient effects of decomposability is in line with former studies which also found different types of decomposability effects (see, for example, Schuppler et al. (2012), or the discussion in §4.3.1).

In the corpus study, the segmentability of the affixes affected the duration of

the nasal in *un-* and *in*-prefixed words. The most segmentable prefix *un-* featured a longer nasal than the less segmentable prefix negative *in-*, which in turn featured a longer nasal than the least segmentable affix locative *in-*.

In the experimental study, nasals in *un-* and *in*-prefixed words also differed in their duration. However, only double nasals were affected, i.e. the double nasal in *un*-prefixed words was longer than the double nasal in *in*-prefixed words. There was no clear difference between the duration of the double nasal in negative and locative *in-*.

The experimental study, furthermore, showed effects of word-specific decomposability. The two decomposability measures SemanticTransparencyRating and logRelativeFrequency affected consonant duration with *-ly*. Items which were rated as less decomposable featured a shorter /l/ than items which were rated as more decomposable. Furthermore, for *-ly*-words with a syllabic double consonant, derivatives with a higher relative frequency, i.e. less decomposable items, featured a shorter /l/ than derivatives with a lower relative frequency, i.e. more decomposable derivatives. However, as discussed in §7.3.7.3, the effects of word-specific decomposability are very small and the effect of relative frequency might be caused by a few items in the data set.

To sum up, the present study shows that categorical segmentability and word-specific decomposability may affect the acoustic duration of complex words. There are differences in the effect of categorical segmentability between corpus and experimental study. This suggests that segmentability might be affected by speech mode. Effects of word-specific decomposability were only found for the suffix *-ly*, not for the prefixes. This suggests that suffixes might be more affected by word-specific decomposability than prefixes. Crucially, the effects of word-specific decomposability do not interact with gemination.

8.2 Morphological gemination: The overall picture

Gemination is a categorical phenomenon. This is evidenced by the bimodal distribution of singletons and doubles found in the data. Furthermore, gemination is mostly governed by the affix, i.e. by a categorical factor. While some affixes geminate, others do not. In addition to the affix itself, for some affixes, the stress pattern of a derivative affects gemination. As shown in the experimental study, gemination does not depend on orthography.

Even though gemination is categorical in the sense that doubles are categorically longer than singletons, there are gradient differences in the degree of gemination between affixes. The degree of gemination, or the strength of gemina-

tion, is mainly indicated by the durational differences between phonological doubles and phonological singletons. Stronger gemination goes together with larger singleton-geminate ratios.

The prefix *un-* clearly geminates in both studies. In the corpus and the experimental study, phonological doubles (e.g. /nn/ in *unnatural*) are clearly longer than singletons in complex words (e.g. /n/ in *uneven* or *untold*). The experimental study furthermore showed that doubles are longer than singletons in base words (e.g. /n/ in *natural*). Singleton-geminate ratios are bigger in the experimental than in the corpus study.

The prefix *in-* also geminates but gemination is weaker than gemination with *un-*. While gemination with *in-* is comparable to gemination with *un-* in the corpus study, in the experimental study gemination with *in-* is weaker and depends on stress. In the corpus study, all doubles are longer than all singletons, and the durational differences between doubles and singletons are similar to the ones found for *un-*. In the experimental study, the singleton-geminate ratios for *in-* are smaller than the ones for *un-*. Furthermore, doubles are only longer than singletons when the base-initial syllable of a derivative is stressed, and doubles are only longer than some types of singletons. For /ɪn/, doubles (e.g. /nn/ in *innumerous*) are longer than singletons in complex words followed by a vowel (e.g. /n/ in *inefficient*), but not longer than singletons followed by a consonant (e.g. /n/ in *intolerant*). For /ɪm/, doubles (e.g. /mm/ in *immortal*) are longer than singletons in complex words followed by a consonant (e.g. /m/ in *impossible*). Doubles in derived words are never longer than initial singletons in base words (e.g. /n/ in *numerous* or /m/ in *mortal*).

The prefix *dis-* geminates in the corpus study as well as in the experimental study. In the corpus study, gemination with *dis-* is weaker than gemination with *un-* and *in-* in the sense that there is a smaller singleton-geminate ratio for *dis-* than for *un-* and *in-*. Furthermore, one *dis-*prefixed type did not geminate (*dissolution*), presumably because of its unstressed base-initial syllable.

In the experimental study, gemination with *dis-* is weaker than gemination with *un-* but similar to gemination with *in-*. All *dis-*prefixed words in the experimental data geminated. However, no morphological geminates with an unstressed base-initial syllable were included. The singleton-geminate ratio for *dis-* is smaller than that for *un-*, and similar to that for *in-*. Furthermore, in contrast to *un-*, and similar to *in-*, doubles are not longer than singletons in base words.

For the suffix *-ly*, no gemination was found. In the corpus study, double consonants with *-ly* were as long as singletons. In the experimental study, three different types of double consonants were investigated: syllabic ones (e.g. /ll/ in

ment(a)lly), non-syllabic ones spelled with the orthographic sequence <lel> (e.g. /ll/ in *solely*), and non-syllabic ones spelled with <ll> (e.g. /ll/ in *really*). While under certain conditions the syllabic doubles and the doubles spelled with <lel> were longer than some types of singletons, there was no systematic difference between doubles and singletons, i.e. two underlying consonants are not longer than one. The suffix *-ly* degeminates.

   Thus, the overall pattern of gemination is the following: the prefix *un-* geminates. The prefixes *in-* and *dis-* also geminate but gemination is weaker. For *in-*, gemination is only weaker than gemination with *un-* under experimental conditions. For *dis-*, gemination is weaker in both the corpus and the experimental study. The suffix *-ly* degeminates. The overall gemination pattern of the affixes is shown in (3) in form of a hierarchy.

   (3)   Gemination Hierarchy: *un-* > {*in-*_{NEG} , *in-*_{LOC}}> *dis-* > *-ly*

A comparison with former empirical studies reveals that the results are largely in line with previous research. Apart from the studies presented in this book, three studies looked at affixational gemination in English, Kaye (2005), Oh & Redford (2012) and Kotzor et al. (2016) (see Chapter 2.4.2 for a detailed discussion). Kaye (2005) and Oh & Redford (2012) looked at gemination with *un-* and *in-*. Their results that both prefixes geminate fit in well with the results presented in this book.

   Oh & Redford (2012) furthermore showed that the singleton-geminate ratio is smaller for *in-* than for *un-*, i.e. *in-* geminates to a lesser degree than *un-*, and that the difference in singleton-geminate ratios between *un-* and *in-* is even larger in careful speech than in normal speech. This fits in well with this study's result that there are only differences in the degree of gemination between *un-* and *in-* in the experimental data, which presumably is more similar to careful speech than the corpus data.

   Kotzor et al. (2016) looked at gemination with *-ness* and *-ly*. They claim that both suffixes geminate. Their result that *-ly* geminates is not in line with the results presented in this book. However, as thoroughly discussed in Chapter 2.4.2, Kotzor et al. (2016) do not provide separate analyses for the two affixes *-ness* and *-ly*. It is thus questionable how valid their result is.

8.3 Corpus study vs. experimental study

There is a peculiar difference between the affixes with regard to their behavior in the corpus study vs. their behavior in the experimental study. While for some

affixes, there is a difference in the degree of gemination between the corpus and the experimental study, for others no such difference was observed. For the prefix *un-*, gemination is stronger in the experimental study than in the corpus study. For *in-*, the opposite is the case: gemination is weaker in the experimental study than in the corpus study. For *dis-* and *-ly*, no difference was found in the degree of (de)gemination between the corpus study and the experimental study.

Before attempting to interpret the differences between the corpus results and the experimental results, it is important to note that all differences between the two studies are merely observed. In other words, no statistical analysis which directly compares the corpus and the experimental data was conducted. The reason for not conducting such an analysis is methodological in nature: the corpus data and the experimental data are too different to be analyzed in one statistical model. One major difference between the data sets is their size. The experimental data set features more than 13 times more observations than the corpus study. In a linear model, this would lead to serious problems, as the model's estimates would be largely based on the experimental data. Furthermore, the compositions of the data sets in the two studies are only partly comparable, i.e. the data sets feature items of slightly different environments (see also §5.2 for discussion). The fact that different variables were coded in the two studies also poses a problem for a direct comparison (see also §5.5 for discussion).

Returning to the observed differences, one might speculate that the different behavior of the affixes with regard to their gemination pattern in the corpus and the experimental study is related to their segmentability. It might, for example, be that highly segmentable affixes like *un-* display stronger gemination under experimental conditions, while less segmentable affixes like *in-* show weaker gemination under experimental conditions. However, this idea does not carry through. As described above, the prefix *dis-* is approximately as segmentable as negative *in-*, and presumably more segmentable than locative *in-*. If the differences between corpus and experimental study were related to segmentability, *dis-* should behave as both *in-*prefixes. It does not. It is thus highly questionable, whether the different behavior of the affixes with regard to their gemination in the corpus and the experimental study is related to their segmentability.

To conclude, there are differences between the results of the corpus and the experimental study but it is yet unclear how these differences can be explained. In order to find out what causes deviating results between corpus and experimental studies, it is necessary to better understand the differences between speech production in reading tasks and speech production in natural conversational speech. In order to do so, more research combining corpus and experimental studies is

necessary (see also §5.1, or Arppe & Järvikivi (2007) for discussion).

8.4 Implications for theory

To theoretically interpret the results of the studies, we need to reconsider the approaches to the morpho-phonological interface discussed in Chapter 4. By comparing the actual gemination pattern (as found in the studies) with the predictions these approaches make, we can find out which approach is supported by the data, and which approach is not.

Table 8.2 gives an overview of the discussed theoretical approaches. For each approach, the table shows the underlying concepts the approach assumes to govern gemination, and the variables which, according to the approach, are predicted to affect gemination in the studies. For example, according to Lexical Phonology, the stratum of the affix is decisive for its gemination. In turn, the variable affix is predicted to affect gemination in the studies. Variables which affected gemination in the studies are printed in black, variables which did not affect gemination are printed in gray.

The studies revealed that gemination in English is categorical and affix-specific. There are no word-specific effects on gemination.[1] Out of all the factors predicted to govern gemination, only one actually affected gemination in the studies, i.e. affix (see Table 8.2). In addition to the affix, the variable BaseInitialStress, which is not predicted to govern gemination, affected gemination with the prefix *in-* in the experimental study.

Only those approaches which predict gemination to be affix-dependent can be supported by the results. Word-specific approaches are not supported by the data. Thus, the two word-specific approaches, i.e. word-specific Morphological Segmentability and word-specific Morphological Informativeness, are not supported by the data.

To find out which of the five approaches that assume affix-specific gemination is supported by the data, we need to take a closer look at their predictions. Even though all five approaches expect gemination with some affixes and degemination with others, they differ with regard to the gemination pattern they predict. For example, while the three formal approaches are rather strict with re-

---

[1] The only exception might be the word *dissolution* which shows a particularly short fricative duration. As discussed thoroughly in §7.3.8, the analyses do not reveal what causes the short /s/ in *dissolution*, but one plausible explanation is its unstressed base-initial syllable. In other words, it is assumed that its degemination is not caused by type-specific factors but by more general prosodic factors.

Draft of May 27, 2019, 18:04

Table 8.2: Summary of underlying concepts and variables predicting gemination according to different theoretical approaches

Approach	Concept	Variable(s)
Lexical Phonology	stratum of affix	affix
Stratal OT	stratum of affix type of base for dual-level affixes	affix type of base
Prosodic Word	prosodic word status of affix	affix semantic transparency type of base
Morphological Segment-ability (word-specific)	decomposability of derivative	relative frequency semantic transparency type of base decomposability rating semantic similarity
Morphological Segment-ability (affix-specific)	segmentability of affix	affix
Morphological Informative-ness (word-specific)	word-specific informativeness	semantic transparency
Morphological Informative-ness (affix-specific)	affix-specific informativeness	affix

gard to the expected gemination or degemination of a certain affix, the two affix-specific psycholinguistic approaches predict an implicational gemination pattern, i.e. they predict gemination to follow a certain hierarchy. For example, if in a given hierarchy A<B<C affix A and affix C geminate, affix B is also expected to geminate. We need to look at each approach individually to find out whether its predictions are supported or falsified by the data.

According to Lexical Phonology, the stratum of the affix is decisive for gemination. The level 1 affixes *in-* and *dis-* are predicted to degeminate, and the level 2 affixes *un-* and *-ly* are predicted to geminate. Except for the fact that the level 2 affix *un-* geminates, all other predictions are wrong. The level 1 affixes *in-* and *dis-* geminate, and the level 2 affix *-ly* degeminates. Thus, the stratal prediction is clearly falsified by the data.

The predictions made by Stratal OT are very similar to the predictions made by Lexical Phonology as they are also based on lexical strata. The level 2 affixes *un-* and *-ly* are predicted to geminate. The gemination of the dual-level affixes *in-* and *dis-* is predicted to depend on the type of base found in a pertinent word. Items with a bound root are predicted to degeminate, and items with words as bases are predicted to geminate. The degemination of *-ly*, and the fact that gemination of *in-* and *dis-* is independent from a derivative's type of base, falsify the predictions made by Stratal OT.

The Prosodic Word Approach predicts that all affixes forming independent prosodic words geminate. All affixes not forming independent prosodic words are predicted to degeminate. According to the approach, *un-* always forms an independent prosodic word and is thus predicted to always geminate, *-ly* never forms an independent prosodic word and is thus predicted to never geminate. Gemination with *in-* and *dis-* is predicted to depend on the derivative the prefix is found in, i.e. on the prosodic word status of the prefix in the derivative. According to Raffelsiefen (1999), the prosodic word status of prefixes is indicated by various features, such as the semantic transparency of the derivative, its type of base and its stress pattern. In this study, the two features semantic transparency and type of base were used as indicators of prosodic word status. All other listed features were not applicable (see discussion in §4.2.3).

The Prosodic Word Approach makes the correct predictions for *un-* and *-ly*. It also correctly predicts that gemination with *in-* and *dis-* is not as consistent as gemination with *un-*. However, the approach fails to predict the correct gemination pattern for *in-* and *dis-*. Neither the semantic transparency of a derivative nor its type of base significantly influences gemination with *in-* and *dis-*. One can thus state, if prosodic word status is defined by a derivative's semantic trans-

parency and its type of base, the Prosodic Word Approach is not supported by the data.

However, there is some evidence for the influence of prosodic structure on gemination. Gemination with *in-* in the experimental study depends on stress. Only derivatives with a stressed base-initial syllable geminate. Furthermore, gemination with *dis-* might also depend on stress. As discussed in section 7.3.8, the degemination of the word *dissolution* is probably caused by its unstressed base-initial syllable, i.e. there is some evidence that *dis-* only geminates when the base-initial syllable of a derivative is stressed.

Based on these results, one can speculate that using different criteria to determine prosodic word status, such as stress, might have led to better predictions of the Prosodic Word Approach for *in-* and *dis-*. However, using stress to determine prosodic word status is very problematic. According to Raffelsiefen (1999), prefixal stress is the crucial determiner for prosodic word status, i.e. not base-initial stress (see §4.2.3). Prefixal stress is, however, very difficult to determine and highly debated in the literature (see discussion on prefixal stress in sections 3 and 5.5.3). Therefore, it was not directly investigated in this study. Instead the effect of base-initial stress on gemination was tested, but the exact relation of base-initial stress and prefixal stress is yet unclear. Further studies which investigate the relation of base-initial stress, prefixal stress and prosodic word status are needed to clarify the matter. Only if valid criteria for prosodic word status are available, one can further investigate whether gemination is governed by prosodic word status.

Let us now turn to the affix-specific psycholinguistic approaches. Both of them are based on the two segmentability hierarchies postulated in Chapter 3 (see (1) and (2) for the two hierarchies). The two hierarchies differ in how they rank the lexical semantics of the affix in relation to productivity, semantic transparency and type of base. The Segmentability Approach is based on both hierarchies and does not specify the role of semantics for segmentability, while the Morphological Informativeness Approach is solely based on the Semantic Segmentability Hierarchy.

Gemination does not pattern according to the Non-Semantic Segmentability Hierarchy. According to that hierarchy, *-ly* is more segmentable than *in-* and *dis-*. As *in-* and *dis-* geminate, one should also find gemination with *-ly*. This is not the case. The suffix *-ly* does not geminate. In turn, there is no support for affix-specific psycholinguistic approaches which predict gemination to pattern according to the Non-Semantic Segmentability Hierarchy.

According to the Semantic Segmentability Hierarchy, *-ly* is the least segment-

able and the least informative affix. It is thus most likely to degeminate. This assumption is supported by the data, i.e. we find degemination with *-ly*. There is also a difference in segmentability and informativeness between the four prefixes: *un-* should geminate to a higher degree than negative *in-* and *dis-*, which in turn should geminate to a higher degree than locative *in-*. As shown in (3), this pattern is at least partly observed, *un-* geminates to a higher degree than both *in-*prefixes and *dis-*. There is, however, no difference in the degree of gemination between locative *in-* and negative *in-*, and gemination with *dis-* is a little weaker than gemination with *in-*.

Overall though, the gemination pattern supports the affix-specific psycholinguistic approaches which predict gemination to pattern according to the Semantic Segmentability Hierarchy. The most segmentable and most informative affix *un-* geminates, the least segmentable and informative affix *-ly* degeminates, and the other affixes pattern in between. Thus, the affix-specific Segmentability Approach (if based on the Semantic Hierarchy) and the affix-specific Morphological Informativeness Approach are supported by the data. The more informative and segmentable an affix is in terms of its semantics, the higher is its degree of gemination.

Turning away from gemination, the studies also revealed that decomposability affects the acoustic realization of complex words. In the corpus data, prefixal consonant durations for *un-* and *in-* reflect the segmentability of the affix. The experimental study revealed gradient effects of decomposability for *-ly*-suffixed words. Both findings go in the same direction: the more decomposable a unit is, the less reduction is found.

While these findings support dual route models of lexical access, in which decomposability affects whether a complex word is accessed as a whole or via its parts (see, for example, Frauenfelder & Schreuder 1992; Schreuder & Baayen 2015; deVaan et al. 2011; Caselli et al. 2016) , there are still questions which remain unanswered. The studies found effects of categorical affix-specific segmentability, as well as word-specific decomposability effects. It remains unclear how these two different types of decomposability effects interact. Furthermore, while there are effects of word-specific decomposability for *-ly*, there are no effects for the prefixes. This evokes the question whether there is a difference between the retrieval and the processing of suffixed words on the one hand, and prefixed words on the other. Further studies are needed to investigate this difference and shed light on the interplay between categorical and word-specific effects of decomposability. Only then can adequate models of word storage and retrieval be specified.

For theories of speech production, the result that locative *in-* and negative *in-*

Draft of May 27, 2019, 18:04

differ in the duration of their nasal in the corpus study has important implications. As a number of other studies before, this outcome shows that morphological structure is mirrored in phonetic detail, i.e. that acoustic realizations are not solely based on phonemic representations (see §4.4 for an overview of studies on the topic). Currently, models of speech production are unspecified with regard to the processing of complex words (see, for example, Dell 1986; Johnson 1997; Levelt et al. 1999; Bybee 2002; Pierrehumbert 2001; 2002). They must be revised, or specified, in order to account for the fact that the acoustic realization of complex words shows traces of morphological structure.

9 Where do we go from here?

The present book set out to investigate the interface of morphology, phonology and phonetics by investigating gemination in English affixation. One main aim of this book was to clarify the role of boundary strength on the acoustic realization of complex words. While previous studies have found evidence for effects of boundary strength on the phonetics of words, the specific nature of these effects is yet unclear (cf. Chapter 2). Clarifying the nature of these effects is, however, quite important, as a clarification is necessary to accurately model the morpho-phonological and the morpho-phonetic interface.

As shown in Chapter 4, different formal linguistic and psycholinguistic theories deviate in their conceptualization of boundary strength. While some theories assume a categorical difference between sets of affixes (e.g. Lexical Phonology, Stratal OT), others assume boundary strength to be a gradient, probabilistic word-specific concept (e.g. the Decomposability Approach, the Morphological Informativeness Approach). While some approaches define boundary strength by means of mainly lexical factors (e.g. Lexical Phonology, Stratal OT), others mainly focus on prosodic aspects (e.g. Prosodic Phonology), and others concentrate on semantics (e.g. Morphological Informativeness). The differences in the conceptualization of boundary strength mirror general differences in theoretical assumptions about the morpho-phonological interface. The different conceptualizations of boundary strength also lead to different predictions for the phonetic realization of complex words. Studies which test these predictions will, in turn, have important implications about the nature of the interface between morphology, phonology and phonetics.

To further investigate possible effects of boundary strength, and to investigate the nature of the morpho-phonological and the morpho-phonetic interface, I conducted a corpus study and an experimental study on morphological gemination in English. As morphological geminates always occur across morphological boundaries, they provide the perfect test case for investigating possible effects of boundary strength. In the studies, I investigated the five English affixes *un-*, negative *in-*, locative *in-*, *dis-* and *-ly*. I tested the predictions of various morpho-phonological and morpho-phonetic approaches. By finding out which approach

can account best for the variation in gemination with English affixes, important implications about the interface between morphology, phonology and phonetics can be drawn.

The studies revealed that while the prefix *un-* geminates in corpus and experimental speech, the prefixes locative *in-* and negative *in-* show differences in their gemination pattern depending on speech mode. In the corpus data, in which one can assume a deeper semantic processing than in the experimental data, the prefixes geminate to a similar degree as *un-*. In the experimental data, they show smaller durational differences between doubles and singletons than *un-*, and gemination depends on the prosody of the derivatives. The prefix *dis-* geminates to a weaker degree than *un-* in both studies, and the suffix *-ly* never geminates.

These results falsify common assumptions about gemination in English (cf. §2.4). They also falsify stratal approaches of the morpho-phonological interface (e.g. Kiparsky 1982; 1985; Mohanan 1986; Bermúdez-Otero 2012; Kiparsky 2015; Bermúdez-Otero 2017). The variation in gemination can best be accounted for by the morphological informativeness and semantic segmentability of the affixes. The more meaning an affix carries and the more informative it is, the stronger it geminates.

The results, furthermore, indicate that, at least in some cases, prosodic structure affects gemination. It is, however, yet to be investigated how the informativeness and segmentability of an affix relate to prosodic structure, and how this relation affects the acoustic realization of complex words. While there are already approaches which assume a close relation between the prosodic structure of a word and its segmentability, such as the Prosodic Word Approach proposed by Raffelsiefen (1999), these approaches are currently not specified enough to account for the findings of this book. More research on the interaction between the prosodic structure of a complex word, its segmentability and its informativeness is needed to devise a more adequate model of the morpho-phonological interface.

Apart from gemination, the studies also revealed general effects of decomposability on the acoustic realization of complex words. These effects have important implications for models of morphological processing and models of speech production. It was shown that both categorical affix-specific segmentability and gradient word-specific decomposability affect the acoustic realization of complex words. Theories of morpho-phonological processing must thus be devised to account for categorical and gradient effects of morphological structure on speech production. Furthermore, the corpus study revealed that the two homophonous prefixes negative *in-* and locative *in-* were realized with different nasal durations.

This result calls for speech production models which allow morphological structure, or some correlate thereof, to affect phonetic detail.

To conclude, the present book shows that current approaches of the morpho-phonological and morpho-phonetic interface are not able to account satisfactorily for the variation found in the phonetic realization of complex words. The studies conducted provide empirical evidence that there is a close connection between morphological informativeness, segmentability and prosody, and that these three factors are crucial in the production of complex words. Further research is needed to accurately model their effects on the phonetic realization of complex words. Furthermore, the results of this book call for a revision of speech production models. These models must specify the morpho-phonological-phonetic interface in such a way that it is able to account for categorical and gradient effects of morphological structure on the acoustic realization of complex words.

Appendices

Appendix A: Decomposability rating

Note that only the experimental version of the decomposability rating for the affixes *un-*, *in-* and *-ly* is printed here. There were only minor alternations in the other versions of the decomposability rating.

A.1 Personal Detail

Before starting the second part of the experiment, please fill out some questions about yourself.

1. What is your age?

2. Are you male or female?

3. What is your first language? If English, please specify which variety of English (e.g. British, American).

4. Are you bilingual - if so, which two languages did you grow up with?

5. Where did you grow up?

6. Which other languages do you speak and which proficiency level do you have (basic, medium proficient, proficient)?

7. Do you have any knowledge of Latin? If so, please specify (i.e. how long did you study it and proficiency level).

8. What is your course of study and which year are you in?

9. At which university do you study?

10. Have you ever studied English linguistics? If so, how long?

11. Have you ever studied phonetics in university?

12. Have you ever studied phonology in university?

13. Have you ever studied morphology in university?

14. Have you ever studied semantics in university?

A.2 Instructions

This is an experiment about English words. Some English words can be broken into smaller, meaningful units. For example, the word *uncool* can be broken down into two units: *un-* and *cool*. The unit *un-* is a unit which occurs at the beginning of many English words. In the word *uncool* it has been added to the word *cool* to make a new word which has the meaning of 'not cool'. Thus, *uncool* can be separated into two meaningful units.

The unit *dis-* is another unit which is often attached to words to form new English words. One example is the word *disconnect*. It can be broken down into the two meaningful units *dis-* (which in this case has a negative or reversing meaning) and *connect*.

Another example of such a unit is *-ly*, for example in the word *possibly* which consists of the parts *possible* and *-ly*. The unit *-ly* has the meaning "in a certain manner" in English.

Other words in English cannot be divided into more than one meaningful unit. Here are some examples of these words: *uncle, family, discipline*. It is impossible to break down the word *uncle* into smaller units. *Uncle* is not dividable into *un-* and *cle*.

Some words are easier to divide into meaningful units than others. The word *uncool* is very easy to break up into two units (*un* + *cool*). The word *dissolve*, however, may not seem quite so easy to segment. Even if you think it is possible to break dissolve down into *dis-* and *solve*, it may seem much easier to segment *uncool*.

As you see, words can be easier or more difficult to divide into meaningful units. The word *uncle* cannot be divided at all, *dissolve* may seem pretty difficult to break into parts and *uncool* may seem very easy to divide into its parts.

In this experiment, you will be presented with a list of words. Your task is to rate how easy or how difficult it is to divide the words into two meaningful parts. All of the words in the list either start with *dis-* or *un-*, or they end in *-ly*. **Please rate on a scale from 1 to 4 how difficult you find it to divide the word meaningfully into two parts (*dis-/ un-* + rest of the word or first part of the word + *-ly*)**. If you think a word is very easy to break down into two parts (like uncool into un- and cool) rate it with 1. If you think it is impossible to divide

the word meaningfully into these two parts (like uncle into un- and cle) rate it with 4. Use the levels between 1 and 4 to rate the different degrees of difficulty to divide the words. The more difficult you think it is to break down a word into two meaningful parts, the higher you should rate it.

It is very important that you provide an answer for each word, even if you are not certain of your answer. There is no right or wrong solution for this task. Just follow your intuition and provide your best guess.

If you don't know the word, please indicate this by ticking the last box on the right ("**I don't know this word**").

Please work through the words one at a time in the order they are presented, marking x in the appropriate column for each word.

Appendix B: Summaries of variables in initial models of corpus study

Table B.1: Summary of dependent variable and covariates used in the initial model for *un-*

Dependent Variable	Mean	St. Dev.	Min	Max	N
AbsoluteConsonantDuration	60	28	16	137	158
Numerical Predictors	Mean	St. Dev.	Min	Max	N
logRelativeFrequency	−0.797	2.518	−8.495	7.098	158
LSAScore	0.362	0.145	0.030	0.810	89
LocalSpeechRate	13.190	2.990	6.136	20.570	158
PrecedingSegmentDuration	87	29	20	167	158
logWordFormFrequency	7.182	1.953	0.000	9.838	158
Categorical Predictors	Levels				N
Environment	n#nV: 23	n#C: 68	n#V: 67		158
BaseInitialStress	stressed: 102	unstressed: 56			158

Table B.2: Summary of dependent variable and covariates used in the initial model for *in-*

Dependent Variable	Mean	St. Dev.	Min	Max	N
ABSOLUTECONSONANTDURATION	76	27	20	170	156

Numerical Predictors	Mean	St. Dev.	Min	Max	N
SEMANTICTRANSPARENCYRATING	2.518	1.108	1	4	156
logRELATIVEFREQUENCY	3.599	4.742	−9.670	10.490	156
LSASCORE	0.255	0.163	−0.040	0.680	56
LOCALSPEECHRATE	14.290	3.666	5.279	24.320	156
PRECEDINGSEGMENTDURATION	60	26	0	127	156
logWORDFORMFREQUENCY	8.582	1.774	1.609	10.700	156

Categorical Predictors	Levels				N
ENVIRONMENT	m#mV: 89	m#C: 67			156
SEMANTICTRANSPARENCYBINARY	opaque: 105	transparent: 51			156
TYPEOFBASE	bound root: 124	word: 32			156
AFFIX	inLoc: 70	inNeg: 86			156
BASEINITIALSTRESS	stressed: 117	unstressed: 39			156

Table B.3: Summary of dependent variable and covariates used in the initial model for *dis-*

Dependent Variable	Mean	St. Dev.	Min	Max	N
ABSOLUTECONSONANTDURATION	103	30	49	205	128

Numerical Predictors	Mean	St. Dev.	Min	Max	N
SEMANTICTRANSPARENCYRATING	1.68	0.98	1	4	128
logRELATIVEFREQUENCY	0.044	2.938	−8.724	9.360	128
LSASCORE	0.271	0.202	−0.040	0.690	40
LOCALSPEECHRATE	13.025	3.164	5.400	22.440	128
PRECEDINGSEGMENTDURATION	58	19	27	134	128
logWORDFORMFREQUENCY	7.875	1.736	1.099	10.738	128

Categorical Predictors	Levels				N
ENVIRONMENT	s#sV : 24	s#C: 45	s#V: 59		128
SEMANTICTRANSPARENCYBINARY	opaque: 54	transparent: 74			128
TYPEOFBASE	bound root: 18	word: 110			128
BASEINITIALSTRESS	stressed: 88	unstressed: 40			128
VOICING	voiced: 24	voiceless: 104			128

Table B.4: Summary of dependent variable and covariates used in the initial model for *-ly*

Dependent Variable	Mean	St. Dev.	Min	Max	N
ABSOLUTECONSONANTDURATION	43	22	5	111	154

Numerical Predictors	Mean	St. Dev.	Min	Max	N
logRELATIVEFREQUENCY	−0.643	2.721	−6.294	8.656	154
LSASCORE	0.358	0.210	−0.030	0.870	118
LOCALSPEECHRATE	13.540	3.477	6.418	26.010	154
PRECEDINGSEGMENTDURATION	84	37	13	212	154
logWORDFORMFREQUENCY	7.277	1.943	0.693	12.580	154

Categorical Predictors	Levels				N
ENVIRONMENT	l#l : 33	syllabic l#l: 48	#l: 73		154
BASEFINALSTRESS	stressed: 35	unstressed: 119			154
PRECEDINGSEGMENT	consonant: 109	vowel: 145			154

Appendix C: Overview of tested interactions in corpus study

Table C.1: Overview of tested interactions in corpus models. ✓: Interaction was tested in the model; –: Interaction was not tested in the model.

	un-	in-	un- & in-	dis-	-ly
Affix x NumberOfConsonants	–	–	✓	–	–
Affix x FollowingSegment	–	–	✓	–	–
Affix x LocalSpeechRate	–	–	✓	–	–
Affix x BaseInitialStress	–	–	✓	–	–
Affix x Pre.Seg.Dur.	–	–	✓	–	–
Affix x logWordFormFrequency	–	–	✓	–	–
Environment x Affix	–	–	✓	–	–
Environment x BaseFinalStress	–	–	–	–	✓
Environment x BaseInitialStress	–	✓	✓	✓	
Environment x LSAScore	✓	✓	–	✓	✓
Environment x PC1	–	✓	–	✓	–
Environment x PC2	–	✓		✓	–
Environment x PC3	–	–	–	✓	–
Environment x logRelativeFrequency	✓	✓	–	✓	✓
Environment x SemanticTransparencyBinary	–	✓	–	✓	–
Environment x SemanticTransparencyRating	–	✓	–	✓	–
Environment x TypeOfBase	–	✓	–	✓	–

Appendix D: Summaries of additional linear models in corpus study

Table D.1: Summary of model for variables predicting the absolute duration of [m] in *im*-prefixed words with PC

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.330	0.012	28.582	0.000
Environment-m#mV	0.049	0.007	6.983	<0.001
LocalSpeechRate	−0.004	0.001	−4.528	<0.001
BaseInitialStress-unstr.	−0.038	0.007	−5.108	<0.001
PC2	0.007	0.003	2.255	0.026

Table D.2: Summary of model for variables predicting the relative duration of [s] in *dis*-prefixed words with PC

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	1.997	0.226	8.834	<0.001
Environment -s#C	−0.935	0.217	−4.304	<0.001
Environment -s#V	−0.930	0.185	−5.024	<0.001
Voicing-voiceless	0.899	0.223	4.025	<0.001
PC1	0.133	0.052	2.555	0.012

Appendix E: Stimuli of experimental study

Table E.1: Stimuli of experiment study

un	in	im	dis	ly
nailed	noxious	mitigable	senting	unacademical
navigable	nominate	mixture	sever	associational
negotiable	numerable	motile	satisfied	nutritional
neighbourly	innervate	maculate	save	therapeutical
nested	innocuous	material	service	agricultural
netted	innominate	mature	similar	aerobical
neutered	innumerable	measurable	simulate	mental
neutral	intumesce	mediate	social	natural
nuanced	inturned	memorial	symmetry	normal
nurtured	intransgressible■	migrant	dissatisfied	heartful
named	intransient	mobile	dissave	doomful
natural	intenacity	moderate	dissect	flavourful
necessary	intemporal	modest	dissemble	wishful
needed	intoxicate	moral	disseminated	hateful
nerve	intubate	mortal	dissenting	lustful
noted	intestate	movable	disservice	careful
noteworthy	intransitive	mutable	dissever	painful
noticed	intrude	immitigable	dissimilar	hopeful
nourishing	intake	immixture	dissimulate	whole
numbered	interior	immotile	dissocial	stale
unnailed	into	immaculate	dissociate	sterile
unnavigable	intemperate	immaterial	dissuade	tranquil
unnegotiable	intolerant	immature	dissymmetry	vile
unneighbourly■	intrepid	immeasurable	discern	pale
unnested	intangible	immediate	disoblige	hostile
unnetted	intractable	immemorial	disinform	civil
unneutered	inexistent	immerse	disincline	cool
unneutral	inexplicit	immigrant	disimprove	cruel

un	in	im	dis	ly
unnuanced	inappeasable	imminent	disidentify	real
unnurtured	inappreciable	immobile	disentwine	sole
unnamed	ineffaceable	immoderate	disencumber	futile
unnatural	inextensible	immodest	disattach	unacademically▉
unnecessary	inexpungible	immoral	disaffirm	associationally▉
unneeded	inexcusable	immortal	disaffiliate	nutritionally
unnerve	inexpressive	immovable	disorient	therapeutically▉
unnoted	inartistic	immutable	disinvest	agriculturally
unnoteworthy	ineliminable	impassion	disassemble	aerobically
unnoticed	inefficacious	impanel	disinterest	mentally
unnourishing	inexpedient	impartible	disenchanted	naturally
unnumbered	inapposite	imprescriptible▉	disown	normally
unknown	inessential	imperforable	disorganize	heartfully
unknit	inelegant	imparity	disestablish	doomfully
untutored	inelligble	imperceivable	disallow	flavourfully
untrue	inapplicable	imperfectible	disabuse	wishfully
untouchable	inelastic	impassible	disorders	hatefully
untrained	inexact	imponderable	disintegrating	lustfully
untold	inedible	improvident	disengage	carefully
untested	inefficient	imperishable	disarm	painfully
untangled	inexperienced	impolitic	disappear	hopefully
untalented	inobservable	imperturbable	disagree	wholely
untypical	inoculate	impalpable	disadvantage	stalely
unteachable	inundate	impracticable	disabled	sterilely
untempered	inact	imprison	disapprove	tranquilly
untrimmed		implode	disobedient	vilely
untaped		import	dissertation	palely
untoasted		implant	dissident	hostilely
untailored		imprint	dissipate	civilly
untalkative		implicit	dissolute	coolly
untactful		impractical	dissonance	cruelly
untweezed		impotence		really
untorn		impossible		solely
untempted		impermissible		futilely
untacky		implausible		snobbily
unassumable		imprecise		roomily
unenjoyed				filthily
unoppressed				windily
unattested				stuffily
unextended				lousily
unoxidized				chubbily

un	in	im	dis	ly
unindexed				trashily
unexampled				mellowly
unamplified				worthily
unaired				wittily
unauthentic				tipsily
unordered				thirstily
unamused				sexily
uninstall				naughtily
unacquainted				rawly
unoriginal				hollowly
uneaten				temporarily
unobserved				luckily
unopened				happily
unease				gaily
unarm				noisily
unequal				groggily
unavailable				crazily
unable				truly
unaided				slowly
uneven				thoroughly
				newly
				narrowly
				shallowly
				twilly
				trolly
				swilly
				skelly
				lolly
				belly
				jelly
				dilly
				silly
				bully
				lilly

Appendix F: Summaries of variables in initial models of experimental study

Table F.1: Summary of dependent variable and covariates used in the initial models for *un-*

Dependent variable	Mean	St. Dev.	Min	Max	N
AbsoluteConsonantDuration	103	49	12	299	2615

Numerical Predictors	Mean	St. Dev.	Min	Max	N
SemanticTransparencyRating	1.107	0.383	1	4	2039
logRelativeFrequency	−9.911	3.153	−9.911	5.966	2067
GlobalSpeechRate	2.035	0.769	0.206	4.610	2615
LocalSpeechRate	10.800	2.558	3.909	21.930	2615
Order	163	94	2	341	2615
PrecedingSegmentDuration	85	41	16	223	2067
logWordFormFrequency	3.276	2.835	0.000	9.79	2615

Categorical Predictors	Levels				N
Environment	n#nV: 966	n#C: 427			2615
	n#V: 674	#nV: 548			
Accentuation	accented: 1317	unaccented: 1298			2615
BaseInitialStress	stressed: 2196	unstressed: 419			2615
PostPause	no pause: 1642	pause: 973			2615
PrePause	no pause: 1230	pause: 1385			2615

Table F.2: Summary of dependent variable and covariates used in the initial models for *in-*

Dependent variable	Mean	St. Dev.	Min	Max	N
AbsoluteConsonantDuration	71	32	18	200	1232

Numerical Predictors	Mean	St. Dev.	Min	Max	N
SemanticTransparencyRating	1.699	1.073	1	4	1118
logRelativeFrequency	− 1.421	3.887	−9.855	8.165	1155
GlobalSpeechRate	2.284	0.584	0.501	4.705	1232
LocalSpeechRate	12.140	2.556	2.461	20.260	1232
Order	162	92	2	319	1232
PrecedingSegmentDuration	65	22	9	147	1155
logWordFormFrequency	3.084	2.529	0	11.970	1232

Categorical Predictors	Levels				N
Environment	n#nV:88	n#C:437			1232
	n#V: 630	#nV: 77			
Affix	inLoc: 270	inNeg: 885			1155
SemanticTransparencyBinary	opaque: 193	transparent: 962			1155
TypeOfBase	bound root: 138	word: 1017			1155
Accentuation	accented: 618	unaccented: 614			1232
BaseInitialStress	stressed: 608	unstressed: 624			1232
PostPause	no pause: 662	pause: 570			1232
PrePause	no pause: 473	pause: 759			1232

Table F.3: Summary of dependent variable and covariates used in the initial models for *im-*

Dependent variable	Mean	St. Dev.	Min	Max	N
ABSOLUTECONSONANTDURATION	95	31	14	245	1635

Numerical Predictors	Mean	St. Dev.	Min	Max	N
SEMANTICTRANSPARENCYRATING	1.784	1.096	1	4	1175
logRELATIVEFREQUENCY	−1.264	2.947	−9.664	6.683	1177
GLOBALSPEECHRATE	2.373	0.599	0.445	4.116	1635
LOCALSPEECHRATE	11.610	2.607	1.010	21.280	1635
ORDER	162	93	2	319	1635
PRECEDINGSEGMENTDURATION	62	22	16	159	1177
logWORDFORMFREQUENCY	4.796	2.620	0.000	9.497	1635

Categorical Predictors	Levels				N
ENVIRONMENT	n#nV: 966	n#C: 427			1635
	n#V: 674	#nV: 548			
AFFIX	inLoc: 315	inNeg: 862			1177
SEMANTICTRANSPARENCYBINARY	opaque:	transparent:			1177
	206	971			
TYPEOFBASE	bound root:	word:			1177
	109	1068			
ACCENTUATION	accented: 831	unaccented: 804			1635
BASEINITIALSTRESS	stressed: 1179	unstressed: 456			1635
POSTPAUSE	no pause: 915	pause: 720			1635
PREPAUSE	no pause: 803	pause: 832			1635

Table F.4: Summary of dependent variable and covariates used in the initial models for *dis-*

Dependent variable	Mean	St. Dev.	Min	Max	N
AbsoluteConsonantDuration	120	32	30	262	1114

Numerical Predictors	Mean	St. Dev.	Min	Max	N
SemanticTransparencyRating	1.495	0.887	1	4	829
logRelativeFrequency	−3.040	4.722	−10.710	6.651	829
GlobalSpeechRate	1.254	0.580	0.212	3.444	1114
LocalSpeechRate	11.980	2.519	3.826	19.230	1114
Order	171	97	6	340	1114
PrecedingSegmentDuration	55	16	15	122	923
logWordFormFrequency	3.724	2.899	0.000	10.640	1114

Categorical Predictors	Levels			N
Environment	s#sV-str.: 242	s#V-unstr.: 430	s#-str.: 157	1114
	sV-str.: 94	#sV-unstr.: 191		
SemanticTransparencyBinary	opaque:	transparent:		829
	162	667		
TypeOfBase	bound root:	word:		829
	70	759		
Accentuation	accented: 572	unaccented: 542		1114
PostPause	no pause: 696	pause: 418		1114
PrePause	no pause: 175	pause: 939		1114

Draft of May 27, 2019, 18:04

Table F.5: Summary of dependent variable and covariates used in the initial models for *-ly*

Dependent variable	Mean	St. Dev.	Min	Max	N
AbsoluteConsonantDuration	60	25	15	178	1645

Numerical Predictors	Mean	St. Dev.	Min	Max	N
SemanticTransparencyRating	1.636	0.921	1	4	1205
logRelativeFrequency	−2.657	2.285	−9.617	3.466	1205
GlobalSpeechRate	1.656	0.687	0.203	5.299	1645
LocalSpeechRate	10.670	2.624	3.593	22.660	1645
Order	176	96	6	340	1645
PrecedingSegmentDuration	97	64	10	364	1645
logWordFormFrequency	3.781	3.011	0.000	10.750	1645

Categorical Predictors	Levels			N
Environment	l#l-<ll>: 313	syll.l#l-<ll> : 132	l#l-<lel>: 151	1645
	#l-<l>: 609	l#-<l>: 115	syll.l#-<l>: 21	
	l#-<le>: 103	l-<ll> : 201		
Accentuation	accented: 820	unaccented: 825		1645
BaseFinalStress	stressed: 538	unstressed: 1107		1645
PostPause	no pause: 1642	pause: 973		1645
PrePause	no pause: 321	pause:1324		1645
TypeOfL	approx.: 1564	tap: 81		1645

Appendix G: Overview of tested interactions in experimental study

Table G.1: Overview of tested interactions in complex models in the experimental study. ✓: Interaction was tested in the model; –: Interaction was not tested in the model.

Two-way interactions	un-	in-	im-	un- & in-	dis-	-ly
ACCENTUATION x BASEINITIALSTRESS	✓	✓	✓	✓	–	–
ACCENTUATION x POSTPAUSE	–	–	–	–	–	✓
ACCENTUATION x PREPAUSE	✓	✓	✓	✓	✓	–
AFFIX x ACCENTUATION	–	–	–	✓	–	–
AFFIX x BASEINITIALSTRESS	–	–	–	✓	–	–
AFFIX x PREPAUSE	–	–	–	✓	–	–
BASEINITIALSTRESS x PREPAUSE	✓	✓	✓	✓	–	–
ENVIRONMENT x AFFIX	–	✓	✓	✓	–	–
ENVIRONMENT x ACCENTUATION	✓	✓	✓	✓	✓	✓
ENVIRONMENT x BASEFINALSTRESS	–	–	–	–	–	✓
ENVIRONMENT x BASEINITIALSTRESS	✓	✓	✓	✓	–	–
ENVIRONMENT x PC1	–	✓	✓	–	✓	–
ENVIRONMENT x PC2	–	✓	✓	–	✓	–
ENVIRONMENT x PC3	–	✓	✓	–	✓	–
ENVIRONMENT x PC4	–	✓	✓	–	–	–
ENVIRONMENT x PREPAUSE	–	–	–	✓	–	–
ENVIRONMENT x logRELATIVEFREQUENCY	✓	✓	✓	–	✓	✓
ENVIRONMENT x SEMANTICTRANS.BINARY	–	✓	✓	–	✓	–
ENVIRONMENT x SEMANTICTRANS.RATING	–	✓	✓	–	✓	✓
ENVIRONMENT x TYPEOFBASE	–	✓	✓	–	✓	–
Three-way interactions						
ACC. x PREPAUSE x BASEINITIALSTR.	✓	✓	✓	✓	–	–
AFFIX x ENVIRONMENT x BASEINITIALSTR.	–	–	–	✓	–	–
AFFIX x ENVIRONMENT x PREPAUSE	–	–	–	✓	–	–
AFFIX x ENVIRONMENT x ACC.	–	–	–	✓	–	–
ENVIRONMENT x ACC. x BASEINITIALSTR.	✓	✓	✓	✓	–	–

Table G.2: Overview of tested interactions in complete models in the experimental study ✓: Interaction was tested in the model; –: Interaction was not tested in the model.

Two-way interactions	un-	in-	im-	dis-	-ly
ACCENTUATION x BASEINITIALSTRESS	✓	✓	✓	–	–
ACCENTUATION x POSTPAUSE	–	–	–	–	✓
ACCENTUATION x PREPAUSE	✓	✓	✓	✓	–
BASEINITIALSTRESS x PREPAUSE	✓	✓	✓	–	–
ENVIRONMENT x ACCENTUATION	✓	✓	✓	✓	✓
ENVIRONMENT x BASEFINALSTRESS	–	–	–	–	✓
ENVIRONMENT x BASEINITIALSTRESS	✓	✓	✓	–	–
ENVIRONMENT x POSTPAUSE	–	–	–	–	✓
ENVIRONMENT x PREPAUSE	✓	✓	✓	–	–
Three-way interactions					
ACC. x PREPAUSE x BASEINITIALSTR.	✓	✓	✓	–	–
ENVIRONMENT x ACC.x BASEINITIALSTR.	✓	✓	✓	–	–
ENVIRONMENT x ACC. x POSTPAUSE	–	–	–	–	✓
ENVIRONMENT x PREPAUSE x BASEINITIALSTR.	–	✓	✓	–	–

Appendix H: Model summaries experiment

H.1 un-

Table H.1: Summary of model for variables predicting the duration of [n] in *un*-prefixed words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.861	0.003	932.530	264.576	<0.001
Environment-n#C	-0.036	0.002	111.594	-19.082	<0.001
Environment-n#V	-0.085	0.002	104.565	-48.059	<0.001
Accentuation- unaccented	-0.009	0.001	1975.654	-8.875	<0.001
LocalSpeechRate	-0.004	0.000	1507.685	-16.563	<0.001
PrePause-pause	0.004	0.001	1991.853	5.330	<0.001
PrecedingSegmentDuration	-0.039	0.015	1980.905	-2.533	0.011
Environment-n#C: Accentuation-unaccented	0.007	0.002	1901.505	4.632	<0.001
Environment-n#V: Accentuation-unaccented	0.013	0.001	1912.465	9.199	<0.001

H.2 in-

H.3 im-

Table H.2: Summary of model for variables predicting the duration of [n] in complete *un*-data set

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.760	0.004	485.232	193.333	<0.001
ENVIRONMENT-#nV	-0.028	0.004	168.155	-7.985	<0.001
ENVIRONMENT-n#C	-0.050	0.004	250.439	-12.865	<0.001
ENVIRONMENT-n#V	-0.113	0.003	171.455	-29.779	<0.001
ACCENTUATION-unaccented	-0.012	0.002	2479.817	-7.377	<0.001
LOCALSPEECHRATE	-0.006	<0.001	2057.861	-20.223	<0.001
PREPAUSE-pause	0.010	0.002	2465.922	6.234	<0.001
BASEINITIALSTRESS-unstressed	-0.007	0.003	81.164	-2.064	0.042
ENVIRONMENT-#nV: ACCENTUATION-unaccented	-0.011	0.003	2429.655	-4.229	<0.001
ENVIRONMENT-n#C: ACCENTUATION-unaccented	0.011	0.003	2419.592	3.943	<0.001
ENVIRONMENT-n#V: ACCENTUATION-unaccented	0.018	0.002	2421.693	7.761	<0.001
ENVIRONMENT#nV: PREPAUSE-pause	-0.024	0.003	2462.937	-8.171	<0.001
ENVIRONMENT-n#C: PREPAUSE-pause	-0.005	0.003	2479.329	-1.731	0.084
ENVIRONMENT-n#V: PREPAUSE-pause	-0.012	0.002	2439.388	-4.798	<0.001

Table H.3: Summary of model for variables predicting the duration of
[n] in *in*-prefixed words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.887	0.004	201.306	223.837	<0.001
Environment-n#C	0.001	0.003	60.047	0.357	0.722
Environment-n#V	-0.025	0.003	57.660	-7.595	<0.001
BaseInitialStress-unstr.	-0.021	0.005	46.606	-4.139	<0.001
Accentuation-unaccented	-0.006	0.002	1060.285	-2.437	0.015
LocalSpeechRate	-0.002	0.000	975.702	-9.702	0.000
PrecedingSegmentDuration	-0.046	0.017	1104.274	-2.621	0.009
Environment-n#C: BaseInitialStress-unstr.	0.027	0.006	45.721	4.887	<0.001
Environment-n#V: BaseInitialStress-unstr.	0.014	0.005	45.994	2.557	0.014
Environment-n#C: Accentuation-unaccented	0.004	0.002	1054.188	1.428	0.154
Environment-n#V: Accentuation-unaccented	0.007	0.002	1054.065	2.902	0.004

Table H.4: Summary of model for variables predicting the duration of
[n] in complete *in*-data set

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.960	0.001	206.499	684.709	<0.001
Environment-#nV	0.006	0.002	122.467	3.469	0.001
Environment-n#C	-0.001	0.001	119.097	-0.769	0.443
Environment-n#V	-0.010	0.001	105.191	-7.056	<0.001
BaseInitialStress-unstr.	-0.008	0.002	51.192	-4.077	<0.001
PrePause-pause	-0.001	0.001	1135.839	-0.985	0.325
Accentuation-unaccented	-0.002	0.001	1134.794	-2.417	0.016
LocalSpeechRate	-0.001	0.000	1001.695	-10.385	<0.001
Environment-n#C: BaseInitialStress-unstr.	0.010	0.002	49.730	4.721	<0.001
Environment-n#V: BaseInitialStress-unstr.	0.005	0.002	50.217	2.454	0.018
Environment-#nV: PrePause-pause	-0.003	0.002	1146.138	-1.994	0.046
Environment-n#C: PrePause-pause	0.002	0.001	1134.157	2.245	0.025
Environment-n#V: PrePause-pause	0.001	0.001	1129.964	0.676	0.499
Environment-#nV: Accentuation-unaccented	-0.003	0.001	1133.429	-2.083	0.038
Environment-n#C: Accentuation-unaccented	0.002	0.001	1127.432	1.985	0.047
Environment-n#V: Accentuation-unaccented	0.003	0.001	1126.725	2.859	0.004

Table H.5: Summary of model for variables predicting the duration of [m] in *im*-prefixed words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.844	0.004	343.329	204.592	<0.001
Environment-m#C	-0.010	0.002	42.325	-4.116	<0.001
BaseInitialStress-unstr.	-0.022	0.004	47.371	-6.057	<0.001
Accentuation-unaccented	0.002	0.002	1038.350	1.408	0.160
LocalSpeechRate	-0.003	<0.001	882.227	-10.600	<0.001
GlobalSpeechRate	-0.008	0.002	874.702	-5.220	<0.001
Environment-m#C: BaseInitialStress-unstr.	0.021	0.004	41.617	4.608	<0.001
BaseInitialStress-unstr.: Accentuation-unaccented	0.004	0.002	1083.980	2.344	0.019

Table H.6: Summary of model for variables predicting the duration of [m] in *im*-words with PC

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.847	0.004	356.512	205.413	<0.001
Environment-m#C	-0.012	0.002	41.352	-5.289	<0.001
BaseInitialStress-unstr.	-0.024	0.003	46.956	-7.119	<0.001
Accentuation-unaccented	0.002	0.002	1040.399	1.649	0.100
LocalSpeechRate	-0.003	0.000	777.506	-10.737	<0.001
GlobalSpeechRate	-0.009	0.002	864.789	-5.595	<0.001
PC4	-0.004	0.001	42.831	-3.243	0.002
Environment-m#C: BaseInitialStress-unstr.	0.022	0.004	40.275	5.372	<0.001
BaseInitialStress-unstr.: Accentuation-unaccented	0.004	0.002	1081.325	2.338	0.020

Table H.7: Summary of model for variables predicting the duration of [m] in *im*-words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.554	0.007	415.113	75.487	<0.001
Environment-#mV	0.013	0.006	74.556	2.409	0.018
Environment-m#C	-0.015	0.005	85.518	-2.714	0.008
BaseInitialStress-unstr.	-0.021	0.009	122.699	-2.486	0.014
PrePause-pause	0.005	0.004	1533.247	1.299	0.194
LocalSpeechRate	-0.006	0.000	1129.270	-12.922	<0.001
GlobalSpeechRate	-0.013	0.002	1484.573	-5.794	<0.001
Environment-#mV: BaseInitialStress-unstr.	-0.025	0.012	85.950	-2.012	0.047
Environment-m#C: BaseInitialStress-unstr.	0.019	0.011	116.749	1.755	0.082
Environment-#mV: PrePause-pause	-0.023	0.005	1524.321	-4.680	<0.001
Environment-m#C: PrePause-pause	-0.007	0.004	1530.068	-1.604	0.109
BaseInitialStress-unstr.: PrePause-pause	-0.024	0.007	1521.988	-3.301	0.001
Environment-#mV: BaseInitialStress-unstr.:: PrePause-pause	0.002	0.011	1513.528	0.152	0.879
Environment-m#C:: BaseInitialStress-unstr.: PrePause-pause	0.030	0.009	1525.853	3.424	0.001

H.4 un- and in-

H.5 dis-

H.6 -ly

Table H.8: Summary of model for variables predicting nasal duration in *un*- and *in*-prefixed words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.873	0.007	677.706	126.594	<0.001
Environment-n#C	0.012	0.007	592.421	1.774	0.077
Environment-n#V	-0.015	0.007	524.169	-1.996	0.046
Affix-inNeg	0.012	0.007	508.387	1.636	0.103
Affix-un	0.035	0.007	613.851	5.070	<0.001
PrePause-pause	-0.012	0.006	2991.736	-2.005	0.045
Accentuation-unaccented	-0.006	0.001	3064.055	-8.498	<0.001
LocalSpeechRate	-0.002	0.000	2863.181	-18.116	<0.001
Environment-n#C: Affix-inNeg	-0.012	0.008	462.937	-1.621	0.106
Environment-n#V: Affix-inNeg	-0.016	0.008	438.960	-1.946	0.052
Environment-n#C: Affix-un	-0.035	0.007	551.818	-4.903	<0.001
Environment-n#V: Affix-un	-0.038	0.008	497.743	-5.067	<0.001
Environment-n#C: PrePause-pause	0.015	0.006	2995.821	2.393	0.017
Environment-n#V: PrePause-pause	0.016	0.007	2995.415	2.349	0.019
Affix-inNeg: PrePause-pause	0.011	0.007	2991.839	1.713	0.087
Affix-un: PrePause-pause	0.017	0.006	2991.773	2.785	0.005
Environment-n#C: Accentuation-unaccented	0.005	0.001	3003.215	5.037	<0.001
Environment-n#V: Accentuation-unaccented	0.007	0.001	3006.199	8.736	<0.001
Environment-n#C:Affix-inNeg: PrePause-pause	-0.010	0.007	2996.499	-1.509	0.131
Environment-n#V:Affix-inNeg: PrePause-pause	-0.016	0.007	2996.286	-2.226	0.026
Environment-n#C:Affix-un: PrePause-pause	-0.017	0.006	2998.649	-2.620	0.009
Environment-n#V:Affix-un: PrePause-pause	-0.021	0.007	2996.462	-3.125	0.002

Table H.9: Summary of model for variables predicting the duration of [s] in complex *dis*-words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.641	0.006	251.773	105.233	0.000
Environment-s#V-str.	-0.032	0.005	55.027	-6.472	<0.001
Environment-s#V-unstr.	-0.039	0.004	57.731	-10.250	<0.001
Accentuation-unaccented	-0.010	0.003	765.176	-3.898	<0.001
LocalSpeechRate	-0.004	<0.001	766.957	-10.153	<0.001
Environment-s#V-str. : Accentuation-unaccented	0.006	0.004	748.006	1.514	0.131
Environment-s#V-unstr. : Accentuation-unaccented	0.013	0.003	749.505	4.294	<0.001

Table H.10: Summary of model for variables predicting the duration of [s] in *dis*-words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.560	0.008	382.437	73.865	<0.001
Environment-s#V-unstr.	-0.044	0.005	68.824	-8.743	<0.001
Environment-s#V-str.	-0.035	0.007	66.238	-5.434	<0.001
Environment-#sV-str.	0.020	0.007	85.063	2.968	0.004
Environment-sV-unstr.	-0.030	0.008	67.547	-3.923	<0.001
Accentuation-unaccented	-0.011	0.003	1024.793	-3.509	<0.001
PrePause-pause	0.009	0.003	1021.470	2.946	0.003
LocalSpeechRate	-0.006	<0.001	1046.887	-12.804	<0.001
Environment-s#V-unstressed: Accentuation-unaccented	0.016	0.004	1009.498	4.466	<0.001
Environment-s#V-stressed: Accentuation-unaccented	0.008	0.005	1007.838	1.700	0.089
Environment-#sV-stressed: Accentuation-unaccented	-0.011	0.004	1005.840	-2.576	0.010
Environment-sV-unstressed: Accentuation-unaccented	0.016	0.005	1005.572	2.911	0.004

Table H.11: Summary of model for variables predicting the duration of
[l] in *ly*-suffixed words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.558	0.010	363.717	56.204	<0.001
Environment-l#l-<lel>	0.013	0.012	59.281	1.047	0.300
Environment-l#l-<ll>	-0.020	0.007	71.391	-2.929	0.005
Environment-syll.l#l-<ll>	-0.006	0.007	100.898	-0.861	0.391
logRelFreq	<0.001	0.001	50.887	-0.319	0.751
SemanticTransRat	-0.002	0.001	1078.139	-2.109	0.035
LocalSpeechRate	-0.007	0.001	1127.018	-14.078	<0.001
TypeOfL-tap	-0.015	0.004	1127.694	-3.965	<0.001
PostPause-pause	0.008	0.003	1130.745	2.645	0.008
Accentuation-unaccented	0.003	0.002	1106.120	1.134	0.257
PrecSegDur	-0.146	0.023	866.828	-6.315	<0.001
Environment-l#l-<lel>: logRelFreq	-0.002	0.002	51.227	-0.942	0.350
Environment-l#l-<ll> : logRelFreq	-0.004	0.002	65.320	-1.879	0.065
Environment-syll.l#l-<ll> : logRelFreq	-0.006	0.002	80.837	-2.579	0.012
Environment-l#l-<lel> : Accentuation-unaccented	-0.018	0.005	1093.421	-3.386	0.001
Environment-l#l-<ll> : Accentuation-unaccented	<0.001	0.004	1095.772	-0.097	0.922
Environment-syll.l#l-<ll> : Accentuation-unaccented	0.003	0.006	1103.543	0.451	0.652

Table H.12: Summary of model for variables predicting the duration of [l] in *ly*-words

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	0.875	0.004	894.748	219.785	<0.001
Environment-syll.l#l-<ll>	0.006	0.004	1542.984	1.517	0.130
Environment-l#l-<lel>	0.011	0.005	871.672	2.178	0.030
Environment-#l-<l>	0.003	0.003	797.578	0.799	0.425
Environment-l#-<l>	0.017	0.004	850.401	4.349	<0.001
Environment-syll.l#-<l>	0.039	0.013	1569.673	2.962	0.003
Environment-l#-<le>	0.024	0.004	542.286	5.597	<0.001
Environment-l-<ll>	0.004	0.004	759.135	0.840	0.401
Accentuation-unaccented	0.001	0.001	1505.165	1.031	0.303
PostPause-pause	0.002	0.002	1503.609	0.834	0.405
logWordFormFreq	<0.001	<0.001	96.145	-1.833	0.070
LocalSpeechRate	-0.003	<0.001	1486.438	-17.471	<0.001
TypeOfL-tap	-0.007	0.002	1525.809	-4.477	<0.001
PrecedingSegmentDuration	-0.072	0.009	1297.435	-8.006	<0.001
Environment-syll.l#l-<ll>: Accentuation-unaccented	0.001	0.003	1500.938	0.517	0.605
Environment-l#l-<lel>: Accentuation-unaccented	-0.007	0.003	1488.434	-2.802	0.005
Environment-#l-<l>: Accentuation-unaccented	<0.001	0.002	1485.507	0.139	0.889
Environment-l#-<l>: Accentuation-unaccented	0.001	0.003	1550.320	0.271	0.786
Environment-syll.l#-<l>: Accentuation-unaccented	-0.007	0.013	1565.536	-0.511	0.609
Environment-l#-<le>: Accentuation-unaccented	-0.008	0.003	1527.444	-2.393	0.017
Environment-l-<ll>: Accentuation-unaccented	<0.001	0.002	1487.369	-0.136	0.892
Environment-syll.l#l-<ll>: PostPause-pause	-0.001	0.004	1497.828	-0.337	0.736
Environment-l#l-<lel>: PostPause-pause	0.005	0.004	1509.643	1.116	0.264
Environment-#l-<l>: PostPause-pause	0.003	0.003	1493.376	1.113	0.266
Environment-l#-<l>: PostPause-pause	0.024	0.004	1548.419	6.652	<0.001
Environment-syll.l#-<l>: PostPause-pause	0.009	0.015	1568.636	0.585	0.559
Environment-l#-<le>: PostPause-pause	0.014	0.004	1510.953	3.780	<0.001
Environment-l-<ll>: PostPause-pause	-0.002	0.004	1499.451	-0.466	0.641

Appendix I: Predicted durations experiment

Table I.1: Overview of predicted consonant durations for prefixed words in the experimental studies

Predicted durations

	Phonol. Doubles	Phonol. Singletons (consonant-adjacent)	Phonol. Singletons (vowel-adjacent)
un-[1]	148 ms	95 ms	51 ms
un-[2]	133 ms	94 ms	54 ms
in-[3]	86 ms	87 ms	53 ms
in-[4]	57 ms	98 ms	46 ms
im-[3]	98 ms	86 ms	NA
im-[4]	74 ms	84 ms	NA
dis-[1]	133 ms	NA	108 ms (103 ms)
dis-[2]	125 ms	NA	105 ms (105 ms)

Durational difference **Singleton–Geminate Ratio**

	Double-Singleton (consonant-adjacent)	Double-Singleton (vowel-adjacent)	Singleton-Double (consonant-adjacent)	Singleton-Double (vowel-adjacent)
un-[1]	53 ms	97 ms	1: 1.6	1: 3.0
un-[2]	39 ms	80 ms	1: 1.4	1: 2.5
in-[3]	-1 ms	33 ms	1: 1.0	1: 1.6
in-[4]	-41 ms	11 ms	1: 0.6	1: 1.2
im-[3]	12 ms	NA	1: 1.1	NA
im-[4]	-10 ms	NA	1: 0.9	NA
dis-[1]	NA	25 ms (30 ms)	NA	1: 1.2 (1: 1.3)
dis-[2]	NA	20 ms (20 ms)	NA	1: 1.2 (1: 1.2)

Note that for *dis-*, the predicted durations and the predicted singleton-geminate ratios for the singleton fricative in derivatives with an unstressed base-initial singleton is put in parentheses after the predicted values for the singletons with a stressed base-initial syllable.

[1] in accented position

[2] in unaccented position

[3] with stressed base-initial syllable

[4] with unstressed base-initial syllable

Draft of May 27, 2019, 18:04

References

Abu-Salim, Issam M. 1980. Epenthesis and geminate consonants in Palestinian Arabic. *Studies in the Linguistic Sciences* 10(2). 1–11.

Allen, Margaret Reece. 1978. The morphology of negative prefixes. In *North Eastern Linguistics Society - papers from annual meeting*, 1–11.

Aoyama, Katsura & Lawrence A. Reid. 2006. Cross-linguistic tendencies and durational contrasts in geminate consonants: An examination of Guinaang Bontok geminates. *Journal of the International Phonetic Association* 36(2). 145–157. DOI:10.1017/S0025100306002520

Aronoff, Mark & Shikaripur N. Sridhar. 1983. Morphological levels in English and Kannada, or atarizing Reagan. In John Richardson, Marks Mitchell & Amy Chukerman (eds.), *Papers from the parasession on the interplay of phonology, morphology, and syntax*, 3–16. Chicago, IL, USA: Chicago Linguistic Society.

Arppe, Antti & Juhani Järvikivi. 2007. Every method counts: Combining corpusbased and experimental evidence in the study of synonymy. *Corpus Linguistics and Linguistic Theory* 3(2). 331–159. DOI:10.1515/CLLT.2007.009

Aylett, Matthew & Alice Turk. 2004. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47(1). 31–56. DOI:10.1177/00238309040470010201

Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge, UK: Cambridge University Press.

Baayen, R. Harald, Richard Piepenbrock & Leon Guilkers. 1995. *The CELEX Lexical Database (CD-ROM).* Philadelphia, PA, USA.

Babyak, Michael. 2004. What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine* 66(3). 411–421.

Balota, David A., Melvin J. Yap, Keith A. Hutchison, Michael J. Cortese, Brett Kessler, Bjorn Loftis, James H. Neely, Douglas L. Nelson, Greg B. Simpson & Rebecca Treiman. 2007. The English Lexicon Project. *Behavior Research Methods* 39(3). 445–459. DOI:10.3758/BF03193014

References

Barry, Martin. 2000. A phonetic and phonological investigation of English clear and dark syllabic /l/. *Bulletin de la Communication Parlée* 5. 77–87.

Barth, Danielle & Vsevolod Kapatsinski. 2014. A multimodel inference approach to categorical variant choice: Construction, priming and frequency effects on the choice between full and contracted forms of *am*, *are* and *is*. *Corpus Linguistics and Linguistic Theory* 13(2). 1613–7035. DOI:10.1515/cllt-2014-0022

Bartko, John J. 1966. The intraclass correlation coefficient as a measure of reliability. *Psychological Reports* 19(1). 3–11. DOI:10.2466/pr0.1966.19.1.3

Bartoń, Kamil. 2016. *MuMIn: Multi-model inference: Version 1.15.6.* https://CRAN.R-project.org/package=MuMIn.

Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker. 2014. *Lme4: Linear mixed-effects models using Eigen and S4.* http://CRAN.R-project.org/package=lme4.

Bauer, Laurie. 2001. *Morphological productivity* (Cambridge Studies in Linguistics 95). Cambridge, UK: Cambridge University Press. DOI:10.1017/CBO9780511486210

Bauer, Laurie & Rodney D. Huddleston. 2002. Lexical word-formation. In Rodney D. Huddleston (ed.), *The Cambridge grammar of the English language*, 1621–1722. Cambridge, UK: Cambridge University Press.

Bauer, Laurie, Rochelle Lieber & Ingo Plag. 2013. *The Oxford reference guide to English morphology* (Oxford Linguistics). Oxford, UK: Oxford University Press.

Beckman, Mary. 1986. *Stress and non-stress accent.* Dordrecht, Netherlands: Foris Publications.

Ben Hedia, Sonia & Ingo Plag. 2017. Gemination and degemination in English prefixation: Phonetic evidence for morphological organization. *Journal of Phonetics* 62. 34–49. DOI:10.1016/j.wocn.2017.02.002

Benjamini, Yoav. 1988. Opening the box of a boxplot. *The American Statistician* 42(4). 257–262. DOI:10.2307/2685133

Berg, Kristian. 2016. Double consonants in English: Graphemic, morphological, prosodic and etymological determinants. *Reading and Writing* 29(3). 453–474. DOI:10.1007/s11145-015-9610-z

Bergmann, Pia. 2014. Reduction and deletion of glottal stops and geminates at pword-boundaries in German – effects of frequency and accentuation. In Javier Caro Reina & Renata Szczepaniak (eds.), *Syllable and word languages*, vol. 40 (Linguae et litterae series), 251–278. Berlin, Germany: Mouton de Gruyter.

Draft of May 27, 2019, 18:04

Bergmann, Pia. 2017. *Morphologisch komplexe Wörter im Deutschen – prosodische Struktur und phonetische realisierung* (Studies in Laboratory Phonology). Berlin, Germany: Language Science Press.

Berkovits, Rochelle. 1993. Progressive utterance-final lengthening in syllables with final fricatives. *Language and Speech* 36. 89–98.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labor in exponence. In Jochen Trommer (ed.), *The morphology and phonology of exponence*, vol. 41 (Oxford Studies in Theoretical Linguistics), 8–83. Oxford, UK: Oxford University Press. DOI:10.1093/acprof:oso/9780199573721.003.0002

Bermúdez-Otero, Ricardo. 2013. Amphichronic explanation and the life cycle of phonological processes. In Patrick Honeybone & Joseph Salmons (eds.), *The Oxford handbook of historical phonology* (Oxford Handbooks in Linguistics), 374–399. Oxford, UK: Oxford University Press.

Bermúdez-Otero, Ricardo. 2017. Stratal phonology. In S. J. Hannahs & Anna Bosch (eds.), *Routledge handbook of phonological theory*, 1–45. London, UK: Routledge.

Blazej, Laura J. & Ariel M. Cohen-Goldberg. 2015. Can we hear morphological complexity before words are complex? *Journal of Experimental Psychology. Human Perception and Performance* 41(1). 50–68. DOI:10.1037/a0038509

Boersma, Paul & David Weenink. 2014. *Praat: Doing phonetics by computer.* http://www.praat.org/.

Booij, Geert Evert. 1983. Principles and parameters in prosodic phonology. *Linguistics* 1(6). 249–280.

Booij, Geert Evert. 1985. The interaction of phonology and morphology in prosodic phonology. In Edmund Gussmann (ed.), *Phono-morphology: Studies in the interaction of phonology and morphology*, 23–34. Lublin, Poland: Redakcja wydawnictw Katolickiego uniwersytetu lubelskiego.

Borowsky, Toni J. 1986. *Topics in the lexical phonology of English.* Amherst, MA, USA: University of Massachusetts dissertation.

Box, George E. P. & David R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society, Series B* 26(2). 211–252.

Breheny, Patrick & Woodrow Burchett. 2015. *Visreg: Visualization of regression models.* http://myweb.uiowa.edu/pbreheny/publications/visreg.

Brewer, Jordan. 2008. *Phonetic reflexes of orthographic characteristics in lexical representation.* Tucson, AZ, USA: The University of Arizona PhD thesis. http://arizona.openrepository.com/arizona/bitstream/10150/195213/1/azu_etd_2636_sip1_m.pdf.

References

Bürki, Audrey, Mirjam Ernestus, Cédric Gendrot, Cécile Fougeron & Ulrich Hans Frauenfelder. 2011. What affects the presence versus absence of schwa and its duration: A corpus analysis of French connected speech. *The Journal of the Acoustical Society of America* 130(6). 3980–3991.   DOI:10.1121/1.3658386

Bybee, Joan L. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14(03). 261–290.   DOI:10.1017/S0954394502143018

Bybee, Joan L. & Paul J. Hopper (eds.). 2001. *Frequency and the emergence of linguistic structure* (Typological Studies in Language). Amsterdam, Netherlands: John Benjamins Publishing Company.   DOI:10.1075/tsl

Byrd, Dani, Jelena Krivokapić & Sungbok Lee. 2006. How far, how long: On the temporal scope of prosodic boundary effects. *The Journal of the Acoustical Society of America* 120(3). 1589–1599.   DOI:10.1121/1.2217135

Caselli, Naomi K., Michael K. Caselli & Ariel M. Cohen-Goldberg. 2016. Inflected words in production: Evidence for a morphologically rich lexicon. *Quarterly Journal of Experimental Psychology* 69(3). 432–454. DOI:10.1080/17470218.2015.1054847

Catford, John C. 1988. *A practical introduction to phonetics*. Oxford, UK & New York, NY, USA: Clarendon Press and Oxford University Press.

Cho, Taehong. 2001. Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica* 58(3). 129–162.   DOI:10.1159/000056196

Cho, Taehong & Patricia A. Keating. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29(2). 155–190. DOI:10.1006/jpho.2001.0131

Cho, Taehong, James M. McQueen & Ethan A. Cox. 2007. Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35(2). 210–243.   DOI:10.1016/j.wocn.2006.03.003

Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. Cambridge, MA, USA: MIT Press.

Clahsen, Harald. 1999. Lexical entries and rules of language: A multidisciplinary study of German inflection. *The Behavioral and Brain Sciences* 22(6). 991–1013. DOI:10.1017/S0140525X99002228

Clark, John Ellery & Colin Yallop. 1995. *An introduction to phonetics and phonology*. 2nd edn. (Blackwell Textbooks in Linguistics 9). Oxford, UK: Blackwell.

Cohen, Clara. 2014. Probabilistic reduction and probabilistic enhancement. *Morphology* 24(4). 291–323.   DOI:10.1007/s11525-014-9243-y

Draft of May 27, 2019, 18:04

Cohen-Goldberg, Ariel M. 2013. Towards a theory of multimorphemic word production: The heterogeneity of processing hypothesis. *Language and Cognitive Processes* 28(7). 1036–1064. DOI:10.1080/01690965.2012.759241

Cohn, Abigail C., William H. Ham & Robert J. Podesva. 1999. The phonetic realization of singleton-geminate contrasts in three languages of Indonesia. In John J. Ohala (ed.), *Proceedings of the 14th international congress of phonetic sciences*, 587–590. American Institute of Physics.

Collie, Sarah. 2008. English stress preservation: The case for 'fake cyclicity'. *English Language and Linguistics* 12(3). 505–532. DOI:10.1017/S1360674308002736

Crawley, Michael J. 2012. *The R book.* Hoboken, NJ, USA: Wiley. https://books.google.de/books?id=XYDl0mlH-moC

Cronbach, Lee J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika* 16(3). 297–334. DOI:10.1007/BF02310555

Cruttenden, Alan & Alfred Charles Gimson. 2014. *Gimson's pronunciation of English.* 8th edn. London, UK & New York, NY, USA: Routledge.

Crystal, David. 2008. *A dictionary of linguistics and phonetics.* 6th edn. (Language library). Boston, MA, USA, Malden, MA, USA & Oxford, UK: Credo Reference & Blackwell Publlishing. http://search.credoreference.com/content/title/bkdictling?institutionId=4361.

Davies, Mark. 2011-. *Corpus of American soap operas: 100 million words.*

Davies, Mark. 2008–2014. *The corpus of contemporary American English: 450 million words, 1990–present.*

Davies, Mark. 2007. *The British national corpus, version 3 (BNC XML Edition).* http://www.natcorp.ox.ac.uk/.

Davis, Stuart. 2011. Geminates. In Marc van Oostendorp, Colin Ewen & Elizabeth Hume (eds.), *The Blackwell companion to phonology: Phonological interfaces*, vol. 2, 837–859. Hoboken, NJ, USA: Wiley-Blackwell.

Davis, Stuart. 2017. Geminates and weight manipulating phonology in Chuukese (Trukese). In Haruo Kubozono (ed.), *The phonetics and phonology of geminate consonants*, vol. 2 (Oxford Studies in Phonology and Phonetics), 230–259. Oxford, UK: Oxford University Press.

Davis, Stuart & Marwa Ragheb. 2014. Geminate representation in Arabic. In Samira Farwaneh & Hamid Ouali (eds.), *Perspectives on Arabic linguistics XXIV–XXV*, vol. 1, 3–19. Amsterdam, Netherlands: John Benjamins Publishing Company.

Delattre, Pierre. 1969. *The general phonetic characteristics of languages.* Santa Barbara, CA, USA: University of California Final Report: 1967-1968.

References

Dell, Gary S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93(3). 283–321.

deVaan, Laura, Mirjam Ernestus & Robert Schreuder. 2011. The lifespan of lexical traces for novel morphologically complex words. *The Mental Lexicon* 6(3). 374–392. DOI:10.1075/ml.6.3.02dev

Dmitrieva, Olga. 2017. Production of geminate consonants in Russian: Implications for typology. In Haruo Kubozono (ed.), *The phonetics and phonology of geminate consonants*, vol. 2 (Oxford Studies in Phonology and Phonetics), 34–65. Oxford, UK: Oxford University Press.

Draper, Norman Richard & Harry Smith. 1998. *Applied regression analysis*. 3rd edn. (Wiley series in probability and statistics. Texts and references section). New York, NY, USA: Wiley.

Ehri, Linnea C. 1993. How English orthography influences phonological knowledge as children learn to read and spell. *Literacy and Language Analysis* 21. 21–43.

Eriksson, Anders & Mattias Heldner. 2016. The acoustics of word stress in English as a function of stress level and speaking style. In *16th Annual Conference of the International Speech Communication Association: Speech beyond speech: Towards a better understanding of the most important biosignal*, 41–45. Red Hook, NY, USA: Curran Associates Inc.

Ernestus, Mirjam, Mybeth Lahey, Femke Verhees & R. Harald Baayen. 2006. Lexical frequency and voice assimilation. *The Journal of the Acoustical Society of America* 120(2). 1040–1051. DOI:10.1121/1.2211548

Frauenfelder, Uli H. & Robert Schreuder. 1992. Constraining psycholinguistic models of morphological processing and representation: The role of productivity. In Geert Evert Booij & Jaap van Marle (eds.), *Yearbook of morphology 1991*, 165–183. Dordrecht, Netherlands: Springer.

Fromont, Robert. 2003-2015. *LaBBCat*. University of Canterbury. https://labbcat.canterbury.ac.nz/system/.

Fromont, Robert & Jennifer Hay. 2012. LaBB-CAT: An annotation store. In Paul Cook & Scott Nowson (eds.), *Proceedings of the Australasian Language Technology Association Workshop 2012*, 113–117. Dunedin, New Zealand.

Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America* 27(4). 765–768. DOI:10.1121/1.1908022

Fry, Dennis B. 1958. Experiments in the perception of stress. *Language and Speech* 1. 126–152.

Gagné, Christina L., Thomas L. Spalding & Kelly A. Nisbet. 2016. Processing English compounds: Investigating semantic transparency. *SKASE Journal of Theoretical Linguistics* 13(2). 2–21.

Gahl, Susanne. 2008. 'Time' and 'thyme' are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language* 84(3). 474–496. DOI:10.1353/lan.0.0035

Galea, Luke. 2016. *Syllable structure and gemination in Maltese*. Cologne, Germany: University of Cologne PhD thesis.

Galea, Luke, Martine Grice, Anne Hermes & Doris Mücke. 2014. Lexical and surface geminates in Maltese. In *Proceedings of the 10th international seminar on speech production*, 130–133.

Giegerich, Heinz Joachim. 1992. *English phonology: An introduction* (Cambridge Textbooks in Linguistics). Cambridge, UK & New York, NY, USA: Cambridge University Press.

Giegerich, Heinz Joachim. 1999. *Lexical strata in English: Morphological causes, phonological effects*. Cambridge, UK & New York, NY, USA: Cambridge University Press. https://books.google.de/books?id=G_u1KCnohBoC.

Giegerich, Heinz Joachim. 2012. The morphology of *-ly* and the categorial status of 'adverbs' in English. *English Language and Linguistics* 16(3). 341–359. http : / / journals . cambridge . org / article _ S1360674312000147. DOI:10.1017/S1360674312000147

Godfrey, John J. & Edward Holliman. 1997. *Switchboard-1 release 2*. Philadelphia, PA, USA.

Godfrey, Ross. 2016. *Morphologically conditioned durational effects in English: Phonetic implementation or lexical access?* Toronto, Canada: University of Toronto Generals paper.

Gussmann, Edmund. 2002. *Phonology: Analysis and theory*. Cambridge, UK & New York, NY, USA: Cambridge University Press.

Hall, Tracy Alan. 1999. The phonological word. In Tracy Allen Hall & Ursula Kleinhenz (eds.), *Studies on the phonological word*, vol. 174 (Current issues in linguistic theory), 1–22. Philadelphia, PA, USA: John Benjamins Publishing Company. DOI:10.1075/cilt.174.02hal

Hall, Tracy Alan. 2001. The distribution of superheavy syllables in modern English. *Folia Linguistica* 35(3–4). 399–442. DOI:10.1515/flin.2001.35.3-4.399

Ham, William Hallett. 2001. *Phonetic and phonological aspects of geminate timing* (Outstanding dissertations in linguistics). New York, NY, USA: Routledge.

Hanique, Iris & Mirjam Ernestus. 2012. The role of morphology in the acoustic reduction. *Lingue e Linguaggio* 2. 147–164.

References

Hanique, Iris, Mirjam Ernestus Ernestus & Barbara Schuppler. 2013. Informal speech processes can be categorical in nature, even if they affect many different words. *Journal of the Acoustical Society of America* 133. 1644–1655. http://scitation.aip.org/content/asa/journal/jasa/133/3/10.1121/1.4790352, accessed 2014-10-23.

Hanique, Iris, Barbara Schuppler & Mirjam Ernestus. 2011. Morphological and predictability effects on schwa reduction: The case of Dutch word-initial syllables. In *11th annual conference of the international speech communication association 2010*, 933–936. Red Hook, NY, USA: Curran Associates. http://www.mpi.nl/people/ernestus-mirjam/publications, accessed 2014-10-23.

Hanote, Sylvie, Nicolas Videau, Franck Zumstein & Philippe Carré. 2010. Les préfixes anglais *un-* et *de-*: étude phonétique et acoustique. *Corela. Cognition, Représentation, Langage* HS-9. http://corela.revues.org/pdf/1081.

Harrell Jr, Frank E. 2017. *Hmisc: Harrell miscellaneous.* http://biostat.mc.vanderbilt.edu/Hmisc.

Harris, John. 1994. *English sound structure.* Oxford, UK & Cambridge, MA, USA: Blackwell.

Hartmann, Reinhard Rudolf Karl & Francis Colin Stork (eds.). 1972. *Dictionary of language and linguistics.* London, UK: Applied Science Publ.

Hay, Jennifer. 2001. Lexical frequency in morphology: Is everything relative? *Linguistics* 39(6). 1041–1070. DOI:10.1515/ling.2001.041

Hay, Jennifer. 2003. *Causes and consequences of word structure* (Outstanding dissertations in linguistics). New York, NY, USA: Routledge. http://www.loc.gov/catdir/enhancements/fy0646/2004297538-d.html.

Hay, Jennifer. 2007. The phonetics of 'un'. In Judith Munat (ed.), *Lexical creativity, texts and contexts*, vol. 58 (Studies in functional and structural linguistics), 39–57. Amsterdam, Netherlands & Philadelphia, PA, USA: J. Benjamins Pub. Co.

Hayes, Bruce. 1986. Inalterability in CV phonology. *Language* 62(2). 321. DOI:10.2307/414676

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20(2). 253–306. http://www.jstor.org/stable/4178626.

Jespersen, Otto. 1965. *A modern English grammar on historical principles: Morphology.* Vol. 6. London, UK: Allen & Unwin.

Johnson, Keith. 1997. Speech perception without speaker normalization: An exemplar model. In Keith Johnson & John W. Mullennix (eds.), *Talker variability in speech processing*, 145–167. San Diego, CA, USA: Academic Press.

Draft of May 27, 2019, 18:04

Jones, Daniel. 1959. The use of syllabic and non-syllabic *l* and *n* in derivatives of English words Ending in syllabic *l* and *n*. *STUF – Language Typology and Universals* 12(1–4). 136–144.　DOI:10.1524/stuf.1959.12.14.136

Kaye, Alan S. 2005. Gemination in English. *English Today* 21(2). 43–55. DOI:10.1017/S0266078405002063

Kemps, Rachel, Mirjam Ernestus, Robert Schreuder & R. Harald Baayen. 2005. Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory & Cognition* 33(3). 430–446.

Kenstowicz, Michael. 1994. *Phonology in generative grammar*. 1st edn. (Blackwell Textbooks in Linguistics 7). Oxford, UK: Blackwell.

Kenyon, John Samuel & Thomas A. Knott. 1953. *A pronouncing dictionary of American English*. Springfield, MA, USA: G. & C. Merriam Co.

Kiparsky, Paul. 1982. Lexical morphology and phonology. In Linguistic Society of Korea (ed.), *Linguistics in the morning calm*, 3–91. Seoul, Korea: Hanshin Pub. Co.

Kiparsky, Paul. 1985. Some consequences of Lexical Phonology. *Phonology* 2. 85–138.　DOI:10.1017/S0952675700000397

Kiparsky, Paul. 2015. Stratal OT: A synopsis and FAQs. In Yuchau E. Hsiao & Lian-Hee Wee (eds.), *Capturing phonological shades within and across languages*, 2–44. Newcastle upon Tyne, UK: Cambridge Scholars Publishing.

Kirchner, Robert Martin. 2001. *An effort based approach to consonant lenition* (Outstanding dissertations in linguistics). New York, NY, USA: Routledge.

Kisler, Thomas, Uwe Reichel, Florian Schiel, Christoph Draxler, Bernhard Jackl & Nina Pörner. 2016. BAS Speech Science Web Services – an update of current developments. In *Proceedings of the 10th International Conference on Language Resources and Evaluation*. Paris, France: European Language Resources Association (ELRA).

Kotzor, Sandra, Benjamin J. Molineaux, Elanor Banks & Aditi Lahiri. 2016. Fake gemination in suffixed words and compounds in English and German. *The Journal of the Acoustical Society of America* 140(1). 356–367.　DOI:10.1121/1.4955072

Kraehenmann, Astrid. 2001. Swiss German stops: Geminates all over the word. *Phonology* 18(1). 109–145. http : / / www . jstor . org / stable / 4420190. DOI:10.2307/4420190

Kreidler, Charles W. 1989. *The pronunciation of English: A course book in phonology*. Oxford, UK & New York, NY, USA: B. Blackwell.

Kubozono, Haruo. 2017. Introduction to the phonetics and phonology of geminate consonants. In Haruo Kubozono (ed.), *The phonetics and phonology of geminate*

consonants, vol. 2 (Oxford Studies in Phonology and Phonetics), 1–12. Oxford, UK: Oxford University Press.

Kunter, Gero. 2016. *Coquery: A free corpus query tool.* http://www.coquery.org.

Kunter, Gero. 2017. *Processing complexity and the alternation between analytic and synthetic forms in English.* Düsseldorf, Germany: Heinrich-Heine-Universität Düsseldorf Habilitationsschrift.

Kuperman, Victor, Mark Pluymaekers, Mirjam Ernestus & R. Harald Baayen. 2007. Morphological predictability and acoustic duration of interfixes in Dutch compounds. *The Journal of the Acoustical Society of America* 121(4). 2261–2271. DOI:10.1121/1.2537393

Ladefoged, Peter. 1993. *A course in phonetics.* 3rd edn. Fort Worth, TX, USA: Harcourt Brace Jovanovich College Publishers.

Ladefoged, Peter. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques.* 1st edn. Malden, MA, USA: Blackwell.

Ladefoged, Peter & Keith Johnson. 2011. *A course in phonetics.* 6th edn. Boston, MA, USA: Wadsworth Cengage Learning.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages* (Phonological Theory). Oxford, UK & Cambridge, MA, USA: Blackwell Publishers.

Lahiri, Aditi & Jorge Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics* 16. 327–338.

Landauer, Thomas K., Peter W. Foltz & Darrell Laham. 1998. An introduction to latent semantic analysis. *Discourse Processes* 25(2–3). 259–284.

Laver, John. 1994. *Principles of phonetics.* Cambridge, UK: Cambridge University Press.

Leben, William. 1980. A metrical analysis of length. *Linguistic Inquiry* 11(3). 497–509. http://www.jstor.org/stable/4178177.

Lee-Kim, Sang-Im, Lisa Davidson & Sangjin Hwang. 2013. Morphological effects on the darkness of English intervocalic /l/. *Laboratory Phonology* 4(2). 475–511. DOI:10.1515/lp-2013-0015

Levelt, Willem Johannes. 1999. Models of word production. *Trends in Cognitive Science* 3(6). 223–232.

Levelt, Willem Johannes. 2000. Producing spoken language: A blueprint of the speaker. In Colin M. Brown & Peter Hagoort (eds.), *The neurocognition of language*, 82–122. Oxford, UK: Oxford University Press. DOI:10.1093/acprof:oso/9780198507932.003.0004

Draft of May 27, 2019, 18:04

Levelt, Willem Johannes, Ardi Roelofs & Antje S. Meyer. 1999. A theory of lexical access in speech production. *The Behavioral and Brain Sciences* 22(1). 1–38, discussion 38–75.

Levin, Juliette. 1985. *A metrical theory of syllabicity*. Cambridge, MA, USA: Massachusetts Institute of Technology dissertation. http://dspace.mit.edu/bitstream/1721.1/15321/2/13915526-MIT.pdf.

Lieberman, Philip. 1960. Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America* 32(4). 451–454. DOI:10.1121/1.1908095

LimeSurvey Project Team & Carsten Schmitz. 2015. *LimeSurvey: An open source survey tool*. Hamburg, Germany. https://www.limesurvey.org/.

Lindblom, Björn. 1963. Spectrographic study of vowel reduction. *The Journal of the Acoustical Society of America* 35(11). 1773–1781.

Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. In William J. Hardcastle & Alain Marchal (eds.), *Speech production and speech modelling*, vol. 55 (NATO ASI Series, Series D), 403–439. Dordrecht, Netherlands: Springer. DOI:10.1007/978-94-009-2037-8{_}16

Machač, Pavel & Radek Skarnitzl. 2009. *Principles of phonetic segmentation*. 1st edn. (Erudica 14). Prague, Czech Republic: Epocha.

Maddieson, Ian. 1985. Phonetic cues to syllabification. In Victoria Fromkin (ed.), *Phonetic linguistics*, 203–222. Orlando, FL, USA & San Diego, CA, USA: Academic Press.

Marchand, Hans. 1969. *The categories and types of present-day English word-formation: A synchronic-diachronic approach*. 2nd edn. Munich, Germany: Beck.

Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese & Steven Pinker. 1995. German inflection: The exception that proves the rule. *Cognitive Psychology* 29(3). 189–256. http://www.sciencedirect.com/science/article/pii/S0010028585710158. DOI:10.1006/cogp.1995.1015

Marslen-Wilson, William D. 2009. Morphological processes in language comprehension. In M. Gareth Gaskell (ed.), *The Oxford handbook of psycholinguistics*, 175–193. Oxford, UK: Oxford University Press. DOI:10.1093/oxfordhb/9780198568971.013.0011

Matthews, Peter H. 1997. *The concise Oxford dictionary of linguistics* (Oxford paperback reference). Oxford, UK & New York, NY, USA: Oxford University Press.

References

Miller, Ann. 1987. Phonetic characteristics of Levantine Arabic geminates with differing morpheme and syllable structures. *Ohio State University. Department of Linguistics: Working Papers in Linguistics* 36. 120–140.

Mohanan, Karuvannur Puthanveettil. 1986. *The theory of lexical phonology* (Studies in Natural Language and Linguistic Theory 6). Dordrecht, Netherlands, Boston, MA, USA & Norwell, MA, USA: D. Reidel.

Nespor, Marina & Irene Vogel. 2007. *Prosodic phonology*. Berlin, Germany: Bod Third Party Titles. http://books.google.de/books?id=GAETNIP_H34C.

Nooteboom, Sieb G. 1972. *Production and perception of vowel duration: A study of the durational properties of vowels in Dutch*. Utrecht, Netherlands: University of Utrecht.

O'Connor, Joseph D. 1973. *Phonetics*. 1st edn. (A Pelican original). Harmondsworth, UK: Penguin Books.

OED (ed.). 2013. *The Oxford English dictionary online*. Oxford, UK: Oxford University Press.

Oh, Grace E. 2013. Effects of boundary strength on geminate duration in English. *Studies in Phonetics, Phonology and Morphology* 19(3). 457–478. DOI:10.1121/1.4783731

Oh, Grace E. & Melissa A. Redford. 2012. The production and phonetic representation of fake geminates in English. *Journal of Phonetics* 40(1). 82–91. DOI:10.1016/j.wocn.2011.08.003

Okada, Akira. 2013. Phonological features of *in*-, textitun- and textitnon-. *International Proceedings of Economics Development & Research* 68. 123–128.

Oller, D. Kimbrough. 1973. The effect of position in utterance on speech segment duration in English. *The Journal of the Acoustical Society of America* 54(5). 1235–1247.

Payne, Elinor M. 2005. Phonetic variation in Italian consonant gemination. *Journal of the International Phonetic Association* 35. 153–181.

Pierrehumbert, Janet B. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In Joan L. Bybee & Paul J. Hopper (eds.), *Frequency and the emergence of linguistic structure*, vol. 45 (Typological Studies in Language), 137–157. Amsterdam, Netherlands: John Benjamins Publishing Company. DOI:10.1075/tsl.45.08pie

Pierrehumbert, Janet B. 2002. Word-specific phonetics. *Laboratory Phonology* 7. 101–139.

Pinker, Steven & Michael Ullman. 2002. The past and future of the past tense. *Trends in Cognitive Science* 6(11). 456–463.

Draft of May 27, 2019, 18:04

Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation* (Topics in English Linguistics 28). Berlin, Germany: Mouton de Gruyter.

Plag, Ingo. 2003. *Word-formation in English.* 1st edn. (Cambridge Textbooks in Linguistics). Cambridge, UK: Cambridge University Press.

Plag, Ingo. 2014. Phonological and phonetic variability in complex words: An uncharted territory. *Italian Journal of Linguistics* 26(2). 209–228.

Plag, Ingo, Julia Homann & Gero Kunter. 2017. Homophony and morphology: The acoustics of word-final S in English. *Journal of Linguistics* 53(1). 181–216. DOI:10.1017/S0022226715000183

Plag, Ingo, Gero Kunter & Mareile Schramm. 2011. Acoustic correlates of primary and secondary stress in North American English. *Journal of Phonetics* 39(3). 362–374. http://www.sciencedirect.com/science/article/pii/S0095447011000374. DOI:10.1016/j.wocn.2011.03.004

Pluymaekers, Mark, Mirjam Ernestus & R. Harald Baayen. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *The Journal of the Acoustical Society of America* 118(4). 2561–2569. DOI:10.1121/1.2011150

Pluymaekers, Mark, Mirjam Ernestus, R. Harald Baayen & Geert Evert Booij. 2010. Morphological effects on fine phonetic detail: The case of Dutch *-igheid.* In Cécile Fougeron (ed.), *Laboratory phonology 10* (Phonology and phonetics), 511–531. Berlin, Germany & New York, NY, USA: Mouton de Gruyter.

Prasada, Sandeep & Steven Pinker. 1993. Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes* 8(1). 1–56. DOI:10.1080/01690969308406948

Price, Patti J. 1981. Sonority and syllabicity: Acoustic correlates of perception. *Phonetica* 37(5–6). 327–343.

R Development Core Team. 2014. *R: A language and environment for statistical computing.* Vienna, Austria. http://www.r-project.org.

Raffelsiefen, Renate. 1999. Diagnostics for prosodic words revisited: The case of historically prefixed words in English. In Tracy Allen Hall & Ursula Kleinhenz (eds.), *Studies on the phonological word*, vol. 174 (Current Issues in Linguistic Theory), 133–201. Philadelphia, PA, USA: John Benjamins Publishing Company. DOI:10.1075/cilt.174.07raf

Revelle, William. 2017. *Psych: Procedures for psychological, psychometric, and personality research.* http://personality-project.org/r/psych.

Ridouane, Rachid. 2010. Geminates at the junction of phonetics and phonology. In Cécile Fougeron (ed.), *Laboratory phonology 10* (Phonology and Phonetics), 61–91. Berlin, Germany & New York, NY, USA: Mouton de Gruyter.

References

Ridouane, Rachid & Pierre Hallé. 2017. Word-initial geminates: From production to perception. In Haruo Kubozono (ed.), *The phonetics and phonology of geminate consonants*, vol. 2 (Oxford Studies in Phonology and Phonetics), 66–84. Oxford, UK: Oxford University Press.

Roach, Peter, Jane Setter, John H. Esling & Daniel Jones. 2011. *Cambridge English pronouncing dictionary*. 18th edn. Cambridge, UK & New York, NY, USA: Cambridge University Press.

Sanz, Javier. 2017. *Lexical storage vs. Phonological computation: Case study of stress variation in English -(at)ory derivatives, paper presented at Old World Conference of Phonology, Düsseldorf 20-22 February 2017*. Düsseldorf.

Schiel, Florian. 1999. Automatic phonetic transcription of non-prompted speech. In John J. Ohala (ed.), *Proceedings of the 14th international congress of phonetic sciences*, 607–610. American Institute of Physics.

Schreuder, Robert & R. Harald Baayen. 2015. Modeling morphological processing. In Laurie Beth Feldman (ed.), *Morphological aspects of language processing*, 131–154. New York, NY, USA: Psychology Press.

Schulte, Marion. 2015. *The semantics of derivational morphology: A synchronic and diachronic investigation of the suffixes -age and -ery in English: Zugl.: Bielefeld, univ., diss., 2014* (Language in Performance 49). Tübingen, Germany: Narr.

Schuppler, Barbara, Wim A. van Dommelen, Jacques Koreman & Mirjam Ernestus. 2012. How linguistic and probabilistic properties of a word affect the realization of its final /t/: Studies at the phonemic and sub-phonemic level. *Journal of Phonetics* 40(4). 595–607.

Sluijter, Agaath M. C & Vincent J. van Heuven. 1996. Acoustic correlates of linguistic stress and accent in Dutch and American English. In *Proceedings of the 4th International Conference on Spoken Language Processing*, vol. 3, 630–633.

Smith, Philip T. & Robert G. Baker. 1976. The influence of English spelling patterns on pronunciation. *Journal of Verbal Learning and Verbal Behavior* 15(3). 267–285. DOI:10.1016/0022-5371(76)90025-6

Smith, Rachel, Rachel Baker & Sarah Hawkins. 2012. Phonetic detail that distinguishes prefixed from pseudo-prefixed words. *Journal of Phonetics* 40(5). 689–705. http://www.sciencedirect.com/science/article/pii/S0095447012000356. DOI:10.1016/j.wocn.2012.04.002

Spencer, Andrew. 1996. *Phonology: Theory and description* (Introducing Linguistics 1). Oxford, UK & Malden, MA, USA: Blackwell Publishers.

Sproat, Richard. 1993. Looking into words. In Sharon Hargus (ed.), *Studies in lexical phonology*, 173–195. Cambridge, MA, USA: Academic Press.

Draft of May 27, 2019, 18:04

Sproat, Richard & Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21(3). 291–311.

Stemler, Steven E. & Jessica Tsai. 2008. Best practices in interrater reliability: Three common approaches. In Jason W. Osborne (ed.), *Best practices in quantitative methods*, 29–49. Los Angeles, CA, USA: Sage Publ.

Stockwell, Robert P. & Donka Minkova. 2001. *English words: History and structure.* Cambridge, UK: Cambridge University Press. http://www.loc.gov/catdir/samples/cam033/00052896.html.

Sugahara, Mariko & Alice Turk. 2009. Durational correlates of English sublexical constituent structure. *Phonology* 26(3). 477–524. DOI:10.1017/S0952675709990248

Szpyra, Jolanta. 1989. *The phonology-morphology interface: Cycles, levels and words* (Croom Helm linguistics series). London, UK: Routledge.

Toft, Zoë. 2013. The phonetics and phonology of some syllabic consonants in southern British English. In *Papers on phonetics and phonology: The articulation, acoustics and perception of consonants* (ZAS papers in linguistics 28), 111–144. Berlin, Germany: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung. http://www.zas.gwz-berlin.de/fileadmin/material/ZASPiL_Volltexte/zp28/zaspil28-toft.pdf.

Topintzi, Nina. 2008. On the existence of moraic onset geminates. *Natural Language & Linguistic Theory* 26(1). 147–184. http://www.jstor.org/stable/40270305.

Topintzi, Nina & Stuart Davis. 2017. On the weight of edge geminates. In Haruo Kubozono (ed.), *The phonetics and phonology of geminate consonants*, vol. 2 (Oxford Studies in Phonology and Phonetics), 260–282. Oxford, UK: Oxford University Press.

Trask, Robert L. 1996. *A dictionary of phonetics and phonology.* 1st edn. London, UK: Routledge.

Tucker, Benjamin V. & Mirjam Ernestus. 2016. Why we need to investigate casual speech to truly understand language production, processing and the mental lexicon. *Mental Lexicon* 11(3). 375–400. DOI:10.1075/ml.11.3.03tuc

Umeda, Noriko. 1977. Consonant duration in American English. *The Journal of the Acoustical Society of America* 61(3). 846–858. DOI:10.1121/1.381374

University of Colorado Boulder. 2017. *LSA @ CU Boulder.* http://lsa.colorado.edu/, accessed 2017-1-9.

Vago, Robert M. & Catherine O. Ringen. 2011. Geminates: Heavy or long? In Charles E. Cairns & Eric Raimy (eds.), *Handbook of the syllable*, vol. 1 (Brill's

handbooks in linguistics), 155–170. Leiden, Netherlands & Boston, MA, USA: Brill. DOI:10.1163/ej.9789004187405.i-464.47

Venables, William N. & Brian D. Ripley. 2011. *Modern applied statistics with S*. 4th edn. (Statistics and computing). New York, NY, USA, Berlin, Germany & Heidelberg, Germany: Springer.

Videau, Nicolas & Sylvie Hanote. 2015. Pronunciation of prefixed words in speech: The importance of semantic and intersubjective parameters. *Lexis. Journal in English Lexicology* 9. http://lexis.revues.org/pdf/982.

Wang, Hsueh-Cheng, Li-Chuan Hsu, Yi-Min Tien & Marc Pomplun. 2014. Predicting raters' transparency judgments of English and Chinese morphological constituents using latent semantic analysis. *Behavior Research Methods* 46(1). 284–306. DOI:10.3758/s13428-013-0360-z

Warner, Natasha, Erin Good, Allard Jongman & Joan Sereno. 2006. Orthographic vs. morphological incomplete neutralization effects. *Journal of Phonetics* 34(2). 285–293. DOI:10.1016/j.wocn.2004.11.003

Warner, Natasha, Allard Jongman, Joan Sereno & Rachel Kemps. 2004. Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics* 32(2). 251–276. DOI:10.1016/S0095-4470(03)00032-9

Wells, John Christopher. 2008. *Longman pronunciation dictionary: For upper intermediate - advanced learners*. 3rd edn. (Always Learning). Harlow, UK: Pearson Education Ltd.

Wennerstrom, Ann. 1993. Focus on the prefix: Evidence for word-internal prosodic words. *Phonology* 10(2). 309–324. DOI:10.1017/S0952675700000075

Wijk, Axel. 1966. *Rules of pronunciation for the English language: An account of the relationship between English spelling and pronunciation*. Oxford, UK: Oxford University Press.

Wurm, Lee H. & Sebastiano A. Fisicaro. 2014. What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language* 72. 37–48. http://www.sciencedirect.com/science/article/pii/S0749596X13001368. DOI:10.1016/j.jml.2013.12.003

Zirkel, Linda. 2010. Prefix combinations in English: Structural and processing factors. *Morphology* 20(1). 239–266.

Zumel, Nina & John Mount. 2014. *Practical data science with R*. Shelter Island NY, USA: Manning.

Zwicky, Arnold. 1995. Why English adverbial *-ly* is not inflectional. In Audra Dainora (ed.), *Papers from the 31th regional meeting of the Chicago linguistic*

society, vol. 31 (CLS publications in print / Chicago Linguistic Society), 523–535. Chicago, IL, USA: Chicago Linguistic Society.

Name index

Name index

Warner, Natasha, 91
Weenink, David, 94
Wells, John Christopher, 1, 13, 14, 16,
        25, 26, 28, 30, 31, 115, 119, 120█
Wennerstrom, Ann, 56, 57
Wijk, Axel, 1, 14, 15
Wurm, Lee H., 105

Yallop, Colin, 111

Zirkel, Linda, 141
Zumel, Nina, 130
Zwicky, Arnold, 32, 209

Subject index

Did you like this book?

This book was brought to you for free
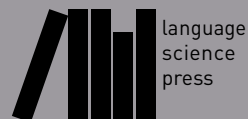
Gemination and degemination in English affixation

In English, phonological double consonants only occur across morphological boundaries, for example, in affixation (e.g. in *unnatural*, *innumerous*). There are two possibilities for the phonetic realization of these morphological geminates: Either the phonological double is realized with a longer duration than a phonological singleton (gemination), or it is of the same duration as a singleton consonant (degemination).

The present book provides the first large-scale empirical study on the gemination with the five English affixes *un-*, locative *in-*, negative *in-*, *dis-* and *−ly*. Using corpus and experimental data, the predictions of various approaches to the morpho-phonological and the morpho-phonetic interface are tested. By finding out which approach can account best for the gemination pattern of English affixed words, important implications about the interplay between morphology, phonology and phonetics are drawn.