Gemination and degemination in English prefixation: Phonetic evidence for morphological organization

Sonia Ben Hedia
Ingo Plag
January 15, 2016

Abstract

This paper addresses the problem of morpho-phonological variability and the role of phonetic detail in morphologically complex words by investigating the gemination behavior of the English prefixes un- and in-. Traditionally, it is assumed that un- geminates while in- degemines, but empirical studies are rare and not conclusive. This paper presents the first study that uses data from natural speech (Switchboard Corpus). It is shown that both prefixes geminate, contra the received wisdom. Furthermore, there is a difference in duration between negative vs. locative in-. The results challenge widely-shared assumptions in morphological theory, lexical phonology and models of speech production.

1 Introduction

The pronunciation of morphologically complex words has traditionally been dealt with under the label of morpho-phonology. There is a long tradition of research in morphophonological alternations in which these alternations have been described as categorical in nature (and not as gradient or probabilistic), and in which words which do not undergo an expected alternation (or unexpectedly undergo one) are taken to be idiosyncratic lexicalized exceptions. Importantly, in this tradition morpho-phonology has been conceived of as being restricted to issues of phonological representations, to the effect that phonetic detail has been considered irrelevant for the study of allomorphy. However, recent research has identified two major problems with this traditional approach to the morphology-phonology interface (see also Plag 2014 for discussion).

The first problem is that morpho-phonological alternations are much more variable than previously thought. For example, Bauer et al. 2013, ch. 9, ch. 14 show that many polysyllabic forms in -able show stress shift hitherto unnoted in the literature. Such variability in stress shift is unexpected under traditional approaches and is not easily accommodated by theories of lexical phonology (e.g. Kiparsky 1982; Mohanan 1986). Similarly problematic facts have been discovered by Collie (2008), who finds a large number of derivatives that, according to pronunciation dictionaries, do not undergo the expected stress preservation, i.e. the preservation of a base main stress as a secondary stress in the derivative.
Another area where variation is more pervasive than previously thought concerns segmental processes across morphemic boundaries. For example, the assimilation involved in the prefixation of *in-* and *un-* is much less straightforward than usually assumed (see Plag 2003; Raffelsiefen 2007 for some standard analyses). In their recent overview, Bauer et al. (2013, 180) write that for the assimilation of *in-* with following labio-dentals, palato-alveolars and velars “the facts of variation are unfortunately unclear, except that it is clear that there is variation.” These authors also find that *un-* in contrast to what the literature usually says, can in fact assimilate to the following segment, but they state that “[w]hether the rates of assimilation and the factors influencing assimilation (formality, speech tempo, regional dialect, etc.) are the same for *inautious* and *unconscious* is not clear” (2013: 180).

The facts for English are also unclear concerning the pronunciation of two adjacent identical consonants across a morphological boundary, a phenomenon which we will call ‘morphological gemination’ or ‘gemination’ for short. Sometimes the two consonants are reduced in a process known as ‘degemination’, but even pronunciation dictionaries, which generally consider citation forms unaffected by context-specific or situation-specific influences, vary in how they treat, for example, word-final /l/ followed by adverbial *-ly*. According to the literature (e.g. Bauer 2001, 82, Bauer et al. 2013, 169), there is gemination with some words (*stately* and *vilely*), but not with others (*fully* and *really*), nor apparently after the suffix *-al* (e.g. *federally*, *globally*, *spiritually*). Degemination is said to be variable with yet other words, for example with *dully* and *wholly*.

As illustrated in section 2 below, it is widely believed that the prefix *in-* as in *innumerous*, is a classical case of degemination, while *un-* as in *unnamed*, is believed to geminate. There are only two phonetic studies available of morphologically-induced gemination in English, Oh and Redford (2012) and Kaye (2005). In these experimental studies, the authors found that both prefixes geminate, but the results are somewhat inconclusive due to methodological problems and a very small set of words being tested. Furthermore, all discussions of the degemination of *in-* in the literature ignore that *in-* actually represents two different prefixes, i.e. negative *in-* as in *incompetent*, and locative *in-* as in *infuse*. The two studies by Kaye and by Oh and Redford show, however, that the empirical situation is far from clear, although the presence or absence of gemination might have important theoretical implications.

In general, the theoretical literature is largely silent about what exactly might determine the above-mentioned variability in morpho-phonological alternations. A notable exception is Collie (2008), who shows that word frequency plays a role in the presence or absence of stress preservation (lower frequency of the derivative and higher frequency of the base favor stress preservation). Given that such frequency effects are indicative of morphological processing during speech production, and thus of morphological structure, variability in morpho-phonological alternations can provide significant evidence about the organization of complex words in the mental lexicon. In this particular case, the patterning of the data supports models in which complex words, even if completely regular, can be stored in the lexicon (e.g. de Vaan et al. 2007, 2011). This runs counter to some models of the mental lexicon which assume that only morphemes and irregular complex words are stored (e.g. Clahsen 1999; Marcus et al. 1995; Prasada and Pinker 1993; Pinker 1997).

The second major problem with traditional approaches to morpho-phonology is that
phonetic detail has been regarded as irrelevant. This seems to be problematic in view of the fact that recent research has shown that details of the acoustic signal may directly reflect morphological structure. Thus there is evidence that phonemically identical strings may systematically vary in their phonetic realization, depending on morphemic status, which is unpredicted by standard models of speech production and not expected to occur according to structuralist and formal theories of phonology-morphology interaction. For example, studies like Kemps et al. (2005a,b) and Blazej and Cohen-Goldberg (2015) found that phonologically segmentally identical free and bound variants of a base (e.g. help without a suffix as against help in helper) differ acoustically, and that listeners make use of such phonetic cues in speech perception.

Smith et al. (2012) discover systematic phonetic differences in the realization of the first three segments between prefixed and pseudo-prefixed words (such as mistime and mistake, respectively). In their experiments Sugahara and Turk (2004, 2009) find phonetic differences between the final segments of a monomorphemic stem as against the final segments of the same stem if followed a suffix. Stems followed by certain suffixes had slightly longer rhymes than their mono-morphemic counterparts. Plag et al. (2015) demonstrate that homophonous suffixes in English (plural, genitive, genitive-plural and 3rd person singular, as well as cliticized forms of has and is) display systematic durational differences. Hay (2007) shows that the phonetic properties of the English prefix un- depend on the segmentability of the word in question.

There is also articulatory evidence on the variability of intergestural timing in monomorphemic and complex words which points at incongruities in the representations of homophones. In an EPG study, Cho (2001) found that in Korean, timing of the gestures for [ti] and [ni] shows more variation when the sequence is heteromorphemic than when it is tautomorphemic, which indicates that morphological structure is reflected in the details of the articulatory gestures, with potential acoustic correlates in the speech signal.

There is no straightforward explanation for these morpho-phonetic findings. Some researchers interpret these facts as evidence for the presence of phonetic detail in lexical representations (e.g. Pierrehumbert 2001, 2002), others adduce prosodic or contextual properties of the words in question (e.g. Sugahara and Turk 2009) in order to make sense of the unexpected systematic relationship between morphological structure and phonetic patterning (see also Plag et al. 2015 for discussion).

The present paper tries to address the problem of morpho-phonological variability and the role of phonetic detail in morphologically complex words by investigating the gemination behavior of the English prefixes un- and in-. The first aim of this study is to shed more light on the behavior of these prefixes, as the conclusions that can be drawn from the two available experimental studies are rather limited. This paper presents the first study that uses data from natural speech. Furthermore, we will include a much larger number of different words than the two previous studies, whose data sets taken together only covered ten different pertinent word types. Using multiple linear regression as our statistical tool, we will compare the durations of the nasals in prefixed words that have a singleton underlying nasal (e.g. unkind, impossible) with those of words that have a double underlying nasal consonant (e.g. unnatural, immature), using data from natural telephone conversations extracted from the Switchboard Corpus (Godfrey and Holliman 1997). Furthermore we investigate factors which have received little or no attention in the previous literature on these prefixes, but which could potentially influence the duration
of the nasals in question, such as the type of affix (i.e. negative vs. locative in-) and the morphological segmentability of the word in question. Finally, we will discuss the implications of our results for theories of morpho-phonology and speech production.

2 Gemination and Degemination in English prefixes

In languages with phonological geminates, a geminate is taken to be a double consonant which is articulated with a particularly long duration (e.g. Hartmann and Stork 1972; Crystal 2008; Catford 1988; Trask 1996; Matthews 1997). In such languages there is usually a phonemic difference between a geminate and the corresponding single consonant, like, for example, in Italian *fatto* ‘done’ versus *fato* ‘fate’. The longer duration of geminates can manifest itself in absolute duration, or in duration relative to the preceding vowel (‘relative duration’, e.g. Ridouane 2010; Miller 1987; Oh and Redford 2012).

In English there is no such phonemic difference. However, two adjacent identical consonants may emerge through affixation (e.g. *unnatural* or compounding (e.g. *book case*). Spencer (1996, 25) suggests different terms for those geminates that are phonological in nature (‘true geminates’), and those that arise through morphology (‘fake geminates’), a terminological distinction picked up, for example, by Oh and Redford 2012. In this paper we prefer to call adjacent identical consonants that arise through morphological processes ‘morphological geminates’ (or ‘geminates’ for short).

What is at issue for English morphological geminates is that there are essentially two possibilities, preservation or reduction. If the two consonants are preserved, we speak of gemination, if the two consonants are reduced we speak of degemination. In case the two underlying consonants are preserved we should expect a significant durational difference between such double consonants and a single consonant, with the double consonant being longer (in either absolute or relative duration).

In the case of reduction, i.e. degemination, two options are possible. The first is categorical in nature: one of the two underlying nasals would be deleted, to the effect that there would be no durational difference between a singleton and the degeminated nasals. As we will see, this idea is the one favored in most of the literature on the degemination of English prefixes.

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 segmentability of the word in question, in addition to general phonetic and lexical factors like speech rate, word frequency etc.

The received wisdom concerning the behavior of the two prefixes *un-* and *in-* is summarized in table 1.

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Example</th>
<th>Phonetic Realization</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>un-</em></td>
<td><em>un</em> + <em>named</em></td>
<td>[ʌnəmæd]</td>
<td>gemination</td>
</tr>
<tr>
<td><em>in-</em></td>
<td><em>in</em> + <em>numerous</em></td>
<td>[ɪnmɔrəs]</td>
<td>degemination</td>
</tr>
</tbody>
</table>
What we consider the received wisdom emanates on the one hand from pronunciation dictionaries, on the other from well-cited works that are not based on systematic empirical investigation but rely on intuition and convenience examples.

Looking at pronunciation dictionaries (e.g. Kenyon and Knott 1953; Roach et al. 2011; Wells 2008) one finds a systematic difference between the representations of the prefix in- and the prefix un-. If an un-prefixed word is attached to a base starting in /n/, the word is transcribed with a long nasal (i.e. with [nː]). In contrast, if an in-prefixed word attaches to a base starting with /n/ the transcription only shows a short /n/ (i.e. [n]). The only exception we were able to find is the word innavigable in Roach et al. 2011, where the word is transcribed with two [n]s.

There is the complication that in- has three variants that may or may not involve gemination: im-, ir- and il-, as in immobile, irresponsible and illegal, respectively. In the dictionaries we consistently find 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.


The theoretical literature accounts for the alleged difference in gemination behavior between the two prefixes 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.), 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, while level 2 affixes form strong boundaries with their base and are phonologically less integrated. Similar in spirit is Harris’ (1994) account, in which the author distinguishes between root affixation (for in-) and word affixation (for un-). 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 suffix in- degeminarizes while the more productive un- geminates. Since productivity correlates with morphological segmentability (e.g. Hay and Baayen 2002; Hay and Plag 2004; Plag and Baayen 2009), this approach could be extended to make the explicit prediction that duration should vary across words depending on the segmentability of the word in question. For example, Hay (2007) showed for un-prefixed words that more decomposable words had a longer prefix duration than less decomposable words. As a measure of decomposability Hay used relative frequency, i.e. the ratio of the frequency of the derived word and the frequency of the base. In the majority of cases, derivatives are less frequent than their bases, but sometimes the opposite is the case. For example, government is more frequent than its base govern, which not only leads to a whole-word bias in lexical access, but also manifests itself in phonetic reduction in the derivative. Applied to gemination, less decomposable
words should display a shorter nasal duration than more decomposable words.

However, all previous attempts to understand the gemination behavior of in- and un-
appear somewhat futile because the actual facts are not clear. As already mentioned,
there are only two studies that have systematically collected empirical data to investi-
gate whether there is gemination with un- and in-. Oh and Redford (2012) and Kaye
(2005). Both studies investigated gemination in in-prefixes words by looking at words
that featured the allomorph \textit{in-}. The reason for this is that there are very few in-prefixes
words with a base starting in /n/. The OED (2013) lists only ten such types, of which
several share the same morphological family (\textit{innervate, innervation, innocuous, innocu-
ously, innocuousness, innominate, innumerable, innumerate}). For reasons
of consistency and simplicity in notation, we will continue to use ‘\textit{in-}’ as a representa-
tion of the morpheme, and thus as a proxy for any of the allomorphs of this morpheme. In
those cases where a particular allomorph was investigated, this will be mentioned, but
the allomorph will not necessarily be represented in the notation of the prefix itself.

Kaye (2005) investigated only two un-prefixes words (unknown, unnamed) and one
in-prefixed word (immature). In an elicitation task, ten speakers produced these words,
as well as the words’ bases in isolation. Kaye then compared the duration of the nasal 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- degeminates.

Oh and Redford’s (2012) study on the gemination of in- and un- compared the du-
ration 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 con-
taining a double nasal, while ammonia, immensely, immunity, immigrational, annex,
inute, annoyed, innerve were categorized as words containing a single nasal. Morpho-
logical structure was not taken into account in this categorization. The items were then
put into carrier sentences and read by eight participants in two different conditions (nor-
mal speech vs. careful speech). The analysis of the durations 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- fail to
show the expected degemination. Furthermore, there is variation in the gemination pat-
tern of in-, as not all words in the singleton set are morphologically simplex. The word
immigizational could be argued to be prefixed (compare migration, immigration), which
in turn would mean that in this word, in- degeminates while in the other prefixed words
it geminates.

To summarize, previous research on the gemination behavior of in- and un- suffers
from a number of problems. First, there is very little empirical evidence available, which means that the facts essentially are unclear. The little evidence that there is calls into question the received wisdom in the literature that \textit{un-} geminates and \textit{in-} degemminates. Second, existing empirical studies are extremely limited in their data sets and only consider words spoken in production experiments, either in isolation or in carrier sentences. What is lacking is data from natural speech. We simply know nothing about whether speakers geminate or do not geminate in their normal speech. Third, existing studies have never looked at the different kinds of \textit{in-} prefixes, locative versus negative, and potential differences between them. Fourth, alleged differences in gemination behavior have been explained in categorical terms, positing two kinds of boundary strength. It is unclear whether probabilistic approaches to morphology (making reference to gradient morphological segmentability, for example) can provide a more adequate explanation of whatever the facts may turn out to be.

This paper will address all four problems by presenting a corpus-based study of more than 300 tokens of \textit{un-} and \textit{in-} prefixed words. Our aim is to test previous assumptions about gemination and degemination in English prefixed words and to investigate which factors have an influence on the duration of the boundary-adjacent nasal. Our study will focus on a number of related research questions.

First, we want to clarify the facts, i.e. find out whether \textit{un-} and \textit{in-} geminate. To that end we will compare the durations of morphological geminates with the duration of singletons in \textit{un-} prefixed and \textit{in-} prefixed words. If the [n] in \textit{un-} words with morphological geminates (e.g. \textit{unnatural}) is longer than the [n] in \textit{un-} words with a singleton (e.g. \textit{uneven}), we can say that \textit{un-} geminates. If the [n] in \textit{in-} prefixed words with morphological geminates (e.g. \textit{innumerable}) is longer than in words with a singleton (e.g. \textit{independent}) we can say that \textit{in-} geminates.

Second, we want to find out which factors have an influence on the duration of morphological geminates. Up until now, no one has looked at the phonetic difference between locative and negative \textit{in-}. This leads us to our second research question: Is there a difference in nasal durations between locative and negative \textit{in-}? If so, this would constitute another case in which morphological structure is mirrored in the details of phonetic realization. This in turn would challenge the widely-held assumption that phonetic implementation has no access to morphological information.

Finally, we want to find out whether a word’s decomposability has an influence on nasal duration. We will measure morphological decomposability using semantic and frequent information (i.e. semantic transparency and relative frequency). If we found that a word’s decomposability has an effect on nasal duration, this would present additional evidence that morphological information has an influence on phonetic implementation. Furthermore, this finding would speak in favor of gradient approaches to morphological structure.

### 3 Methodology

#### 3.1 Data

In this study we investigate two different data sets. One data set contains \textit{un-} prefixed words, the other \textit{in-} prefixed words. All words in the data sets were taken from the
Switchboard Corpus (Godfrey and Holliman 1997), which consists of about 2400 two-sided phone conversations among North American speakers of English. It comprises over 3 million word tokens.

To compile our data sets, we first searched the corpus for words starting with the grapheme structures <un>, <in>, <imm>, <imp> and <imb>. We used the speech corpus management system LaBB-CAT (Fromont and Hay 2008, 2012) for the search and extracted all 21502 tokens in the corpus with the pertinent structures. Since we only wanted to include prefixed words we checked all of the extracted words for their morphological status by using established criteria (cf. e.g. Plag 1999, Chapter 5, Bauer et al. 2013, Chapter 3.2.2, Schulte 2015, Chapter 6). All words whose base was attested outside the derivative with a similar meaning counted as prefixed. Note that it did not matter whether the base of a word is a free morpheme (e.g. natural in unnatural) or a bound morpheme (e.g. -plicit in implicit and explicit). Furthermore, we also included words whose base exhibits some allomorphy (such as the vowel change in migrant - immigrant). Of all the extracted tokens 6243 were categorized as prefixed.

The prefixed words were then coded for the number of identical underlying segments they featured at their morphological boundary. After this categorization into double versus single nasals, it turned out that the corpus only contained 17 /m/-prefixed tokens with two underlying /n/s at the morphological boundary. We find only five different types with these tokens (innocous, innovated, innovation, innovative, innovativeness). Furthermore, out of these five types four share the same root. As discussed above this is not an unexpected outcome since /m/-prefixed words with a double consonant at the morphological boundary are extremely rare. Because of the low frequency of double nasals with the allmorph /In/ we decided to focus on the allomorph /Im/, for which 260 tokens were found.

As mentioned above, 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. They consider all allomorphs of in- to undergo degemination (e.g. Borowsky 1986; Cruttenden and Gimson 2014). The pronunciation dictionaries also show degemination for all allomorphs. It is therefore justified to take the gemination pattern of the allomorph /Im/ as representative of the gemination pattern of the morpheme in-. Investigating the allomorph /Im/ also has the advantage of giving us the possibility to directly link our results to the two previous studies on gemination which also analyzed /m/ instead of /n/.

We sampled the different types and found that there were only 22 un-prefixed tokens with an orthographic double nasal attested in the corpus. In order to keep the data sets as large as possible (to ascertain sufficient statistical power), but also to keep the differences in size between the several subsets relatively small, we decided to sample randomly between 70 and 100 tokens from the other token subsets. 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 too low, or because they were pronounced in very unnaturally-sounding ways, or with long hesitation breaks. The final data sets were of comparable size (158 tokens for un-, 157 for in-).

Phonetic studies have shown that the duration of a nasal consonant heavily depends on the neighboring segment. While following vowels lead to shorter nasal durations, following (non-nasal) consonants increase it (Umeda 1977, 854). We are thus faced with
three relevant environments: the suffixal nasal is followed either by a base-initial nasal (e.g. un-natural), or by a base-initial vowel (e.g. un-even), or by a base-initial non-nasal consonant (e.g. un-predictable). The allomorph /Im/, however, only has two relevant environments, a double nasal, or a nasal followed by a non-nasal consonant. This is due to the nature of the assimilatory allomorphy of in-. The prefix in- only takes the form /im/ when it is followed by homorganic consonants, i.e. by the bilabials /m/, /b/ or /p/. Thus, for the in-prefixed words containing a single nasal, only the sequences /mb/ and /mp/ exist. There is no attested sequence /mV/.

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. We therefore decided to include only one token of a specific type from one given speaker. Different tokens of the same type thus generally come from different speakers. However, due to the small number of items with double nasals in the un-data set we made an exception for two speakers, each of whom provided two tokens of one pertinent type.

After compiling the data sets, we extracted the sound files containing the pertinent tokens and transcribed them phonetically. We coded the tokens for their phonetic environment as described above.

The final data sets contained 158 complex words with the affix un- and 157 in-prefixed words (with the allomorph /Im/). Tables 2 and 3 summarize the distribution of the two final data sets by showing the type and token numbers for each phonetic environment described above.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Example</th>
<th>Number of Types</th>
<th>Number of Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>n#nV</td>
<td>unnecessary</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>n#C</td>
<td>unfit</td>
<td>53</td>
<td>68</td>
</tr>
<tr>
<td>n#V</td>
<td>unable</td>
<td>42</td>
<td>67</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>101</td>
<td>158</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Environment</th>
<th>Example</th>
<th>Number of Types</th>
<th>Number of Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>m#mV</td>
<td>immemorial</td>
<td>17</td>
<td>90</td>
</tr>
<tr>
<td>m#C</td>
<td>impossible</td>
<td>67</td>
<td>67</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>84</td>
<td>157</td>
</tr>
</tbody>
</table>

### 3.2 Acoustic measurements

After extracting the sound files from the Switchboard Corpus, we manually segmented the data and transcribed them phonetically using the software Praat (Boersma and Weenink
For each item a textgrid was generated which was made up of five tiers which were coded for context, item ID, item, segments and syllable structure. Figure 1 shows the waveform, the spectrogram and the textgrid of the item *unnecessary*.

As one can see in figure 1, we annotated the segments of the affix in question and the segments of the syllable immediately following the prefix under investigation. We used the SAMPA System to transcribe our data (Wells 1998). In the case of *unnecessary*, the annotation included [a], [n] and [æ]. Double nasals were treated as one segment in the annotation since no boundaries between two identical nasals were discernible. Furthermore, syllable boundaries were set according to the Maximum-Onset-Principle, observing sonority (see, for example, Blevins 1995).

The segmentation of the data was carried out according to a number of criteria that relied on the visual inspection of the waveforms and spectrograms of the items. These criteria were based on the features of specific sounds as described in the phonetic literature (cf. Ladefoged 2003), and refer to the properties of the soundwave (e.g. changes in amplitude or wavelength), the spectrum (e.g. changes in the formant structure). Nasals, for example, have a regular waveform which has a lower amplitude than the waveform of vowels. This can be seen in the uppermost panel of figure 1. In the spectrogram, given in the panel below the waveform, one can also see that nasal formants are quite faint in comparison to the ones of vowels. Boundaries were set at the nearest zero crossing of the waveform.

The reliability of the segmentation criteria was verified by a set of trial segmentations. In these trials, the four annotators used the criteria to segment the same 30 items. If there was any larger discrepancy in the placement of the boundaries, i.e. if any boundary differed from another by more than 0.01 seconds, the annotators discussed the discrepancy and refined the criteria in order to reduce the amount of intra-annotator variation. These trial segmentations were repeated two times until all boundaries were reliably placed with only small variations. For the final measurement, each annotator worked on a disjunct set of items. After the segmentation process was completed, a script was used to measure and extract word duration, the duration of the nasal in question, as well as the duration of its preceding and following segments in seconds.
Tables 4 and 5 show the distributions of the duration of the nasals in the *un*- data set (table 4) and in the *in*- data set (5) for each environment. In both data sets one can see that the mean and the median for the double nasal (n#nV and m#mV) is higher than the one of the single nasals (n#C, n#V and m#C). Generally, nasal durations vary a great deal by phonological environment. Overall, however, the durations of the nasals in our 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 word-internal singleton /n/ in monomorphemes is 38 ms long on average if followed by a vowel in stressed position, and 34 ms if followed by a vowel in unstressed position. Word-initial singleton /n/ is 71 ms long on average in that study, while double /n/s across a word boundary have a duration of 100 ms.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Example</th>
<th>mean</th>
<th>median</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>n#nV</td>
<td>unnecessary</td>
<td>0.099</td>
<td>0.101</td>
<td>0.021</td>
</tr>
<tr>
<td>n#C</td>
<td>unfit</td>
<td>0.064</td>
<td>0.060</td>
<td>0.024</td>
</tr>
<tr>
<td>n#V</td>
<td>unable</td>
<td>0.045</td>
<td>0.040</td>
<td>0.0202</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>0.060</td>
<td>0.054</td>
<td>0.029</td>
</tr>
</tbody>
</table>

Table 5: Summary duration of nasal in seconds in the *in*- data set

<table>
<thead>
<tr>
<th>Environment</th>
<th>Example</th>
<th>mean</th>
<th>median</th>
<th>standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>m#mV</td>
<td>immemorial</td>
<td>0.087</td>
<td>0.081</td>
<td>0.027</td>
</tr>
<tr>
<td>m#C</td>
<td>impossible</td>
<td>0.061</td>
<td>0.061</td>
<td>0.019</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>0.076</td>
<td>0.074</td>
<td>0.027</td>
</tr>
</tbody>
</table>

Figure 2 depicts these differences using boxplots. The distribution for *un*- is shown in the left panel, the distribution for *in*- in the right panel. The y-axis of each plot displays the duration of the nasal in seconds. The leftmost boxplot in each panel represents the distribution of measurements for the items which have a double nasal (n#nV as in unnatural and m#mV as in immature). The other boxplots show the distributions for items with a single nasal (n#C as in unfit, n#V as in uneven, m#C as in impossible, respectively). Each box represents 50% of items with this factor level. The black dot in each box marks the median nasal duration for each category, i.e. the 50% of all items of a particular category are longer and 50% are shorter than the duration marked by this dot. For example, for the *un*- prefixed items with a double nasal the graph shows that half of the items have a nasal duration that is longer than 101 milliseconds.\(^1\) Furthermore, we can see that 50% of the nasals in this category are between 83 and 115 milliseconds long.

\(^1\)Throughout this article, measurements reported in milliseconds are rounded.
Figure 2: Distribution of nasal duration in the \textit{un}-data set and in the \textit{in}-data set.

In both panels the distribution of the measurements shows rather clear differences between the different environments. For \textit{un}- an ANOVA shows a significant effect of environment ($F=51.442, p<2.2e–16$). The pair-wise comparison of the means using Tukey contrasts yields significant contrasts for all three pairs, as shown in table 6.$^2$

Table 6: Multiple comparison of means of nasal duration for \textit{un}-prefixed words (Tukey contrasts)

|       | Estimate | Std. Error | t-value | Pr(>|t|)   |
|-------|----------|------------|---------|------------|
| $n\#C - n\#nV$ | -0.035   | 0.005      | -6.462  | < 1e-06    |
| $n\#V - n\#nV$ | -0.055   | 0.005      | -10.004 | < 1e-06    |
| $n\#C - n\#V$  | -0.019   | 0.004      | -5.066  | 3.18e-06   |

For \textit{in}-, a Welch t-test shows a significant difference between the two environments ($t(154.22)=-7.0492, p=5.648e-11$).

3.3 Predictor variables: Selection

The duration of segments in natural speech is subject to a variety of different influences, and in order to address our research questions these influences need to be controlled.

$^2$For the statistical analyses presented in this paper, we used R (R Development Core Team 2014. For the pairwise comparisons the package \texttt{multcomp} (Hothorn et al. 2008) was used.
for. This can be done by coding the pertinent variables and using them as independent variables in a multiple regression model.

We can distinguish variables of interest and noise variables. In our case, the variables of interest are the number of nasals, the affix, semantic transparency and frequency measures. In addition to the variables of interest there are of course many other factors that might influence the duration of segments in speech production, such as speech rate, the following segment, or the position of the word in the utterance. The inclusion of all these variables raises a number of problems concerning the statistical analysis. Before describing in detail the variables that entered the analysis we will first discuss our general strategy of variable selection.

Multiple regression is an established and highly successful way to deal with the multitude of factors involved in predicting durational properties of morphemes. Such models will therefore also be used in the present paper. While regression analysis has the advantage of letting us look at the effect of one predictor in the presence of other, potentially intervening, predictors there are also some caveats. Two of them are especially pertinent for our analysis, collinearity and overfitting. Let us first discuss collinearity.

There are a number of measurements that we would want to use in our analysis that are correlated with each other, which can lead to serious problems in regression models (‘multicollinearity’, e.g. Baayen 2008, chapter 6). In order to address these problems different strategies are possible. One is residualization, but this strategy has been shown to be inadequate (e.g. Wurm and Fisicaro 2014). Another 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. A third strategy can be applied if correlations concern only noise variables. For this particular case Wurm and Fisicaro (2014) show that correlations of noise variables may lead to unstable and uninterpretable estimates for these variables, but this effect is not harmful for the model itself, as it does not affect the estimates for the variables of interest and still controls for the effect of the noise variables. To address potential collinearity problems in this paper we observed the strategies just mentioned. The details will be described in the next subsection for each of the variables in question.

The other major problem is overfitting. In general, comparatively small data sets like the present ones require the number of parameters in the model to be severely restricted. We therefore tried to include only those noise variables (in addition to the variables of interest) that are well-known to have an influence on nasal duration. To reduce the number of parameters, we also tried to conflate two or more predictors into a single new one, if possible.

In the next section we will describe each variable that was initially considered for inclusion into the model, and how we treated it in our analysis.

### 3.4 Predictor variables: Description

**Number of Nasals and Following Segment.** The first of our research questions asks whether *un-* and *in-* prefixed words containing double nasals across the morphological boundary show gemination or degemination. Coding this variable seems straightforward, but there is the complication, already mentioned in section 3.1, that the duration of a nasal consonant heavily depends on the neighboring segment. We therefore coded the
variable **Environment** for *un*-prefixed words using three levels reflecting these three environments: n#nV, i.e. a double /n/ followed by vowel (as in *unnatural*), n#C, i.e. a single /n/ followed by a non-nasal consonant (as in *unfit*)³, and n#V, i.e. a single /n/ followed by a vowel (as in *uneasy*).

For the words prefixed with *in*- we only obtain two levels, since we are dealing with the allomorph /m/, which only occurs before bases that have bilabial consonants in initial position (i.e. /m/, /b/, or /p/). The two levels are thus m#mV (as in *immemorial*) and m#C (as in *impossible*).

**Affix.** Our second research question asks whether there is a difference in the duration of the nasal between negative *in-* (e.g. *impossible*) and locative *in-* (e.g. *immigrant*). The factor **Affix** captures this difference, using the two levels NegIn and LocIn.

**Semantic Transparency.** Our third research question asks whether the decomposability of a word influences the duration of the nasal. One way to operationalize decomposability is semantic transparency. This variable 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. We therefore created the variable **SemanticTransparency**, in which we coded whether the meaning of the derivative was transparent or opaque. If the meaning of the derivative was fully compositional, it was categorized as transparent. We coded as fully compositional words in which the meaning of the derived word is straightforwardly computed by combining the meaning of the affix with the meaning of the base. Examples of transparent words are *unnatural* and *impossible*, whose meaning can be paraphrased as combining the prefixal meaning ‘not’ with the meaning of the base. Words that did not meet this strict criterion were categorized as opaque, as, for example, *impression* or *imposed*. All *un*-prefixed words turned out to be transparent (all with a clearly negative meaning of *un*-*), the variable **SemanticTransparency** is therefore of interest only for the *in*-data set, in which we find some variation.

**Relative Frequency.** Another measure of decomposability is probabilistic in nature, relative frequency (Hay 2003). Relative frequency is defined as the ratio of the frequency of a derived word to the frequency of its base. The more frequent a derivative is in comparison to its base, the less decomposable the complex word is and the higher its relative frequency. We computed the variable **RelativeFrequency** by dividing a word’s lemma frequency by its base lemma frequency. Frequencies were extracted from the DVD version of the Corpus of Contemporary American English (COCA) (Davies 2008), using the query tool *Coquery* (Kunter 2015). We consider COCA an adequate source for the frequency counts because the data in this corpus come from the same variety of English as the speech data in the Switchboard Corpus, i.e. North American English. We log-transformed **RelativeFrequency** before it entered the models.

³There were no words in which /n/ was followed by /m/ or /ŋ/.
Let us now turn to the noise variables.

**Speech Rate.** The probably most obvious influence on segmental duration is 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 given amount of time, the shorter these units are. 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 e.g. Pluymaekers et al. 2005a,b; Plag et al. 2015). To compute this ratio, relatively long strings of uninterrupted speech produced by a given speaker are required. Due to a large amount of turn taking found in the Switchboard Corpus, this measure was not feasible for data from this corpus. We therefore decided to use a more local measure, the number of segments per second. We computed the values for the variable `SpeechRate` for each item by dividing the number of segments included in the word by the total word duration in seconds. It is expected that the more segments are produced per second, i.e. the higher the speech rate, and the shorter the duration of the nasal(s) in question will be. As a second measure of speech rate, the duration of the preceding vowel was measured. This measure is explained in more detail in the following paragraph.

**Duration of the Preceding Vowel.** There are two reasons for including the duration of the preceding vowel in the analyses. The first reason is that it can be used as a second, extremely local, measure of speech rate (as, for example, in Ernestus et al. 2006). The second reason is that gemination may manifest itself on the vowel preceding the geminated segment (‘relative duration’, e.g. Ridouane 2010; Miller 1987; Oh and Redford 2012). In order to test whether we find effects of relative duration it is necessary to include `precedingVowelDuration` in our models. The inclusion of this variable has the additional advantage that it helps us to tease apart degemination effects and other kinds of reduction effects. In her study of the phonetics of *un-* Hay (2007) finds that with declining decomposability, the whole prefix becomes shorter and not only the nasal. Including `precedingVowelDuration` as a covariate controls for this effect.

**Accentuation.** Previous research has revealed that words which bear sentence accent show less reduction Bergmann 2014, and a longer duration than words which are not accented (e.g. Sluijter and van Heuven 1996; Turk and White 1999). The effect manifests itself in the duration of the individual segments of the word. For our data this translates into the prediction that in words which are accented, the consonant in question will be pronounced with a longer duration than when it is part of an unaccented word. Unfortunately, due to the nature of the data, it was not possible to reliably determine sentence accentuation. However accent-based lengthening is an effect that would manifest itself in the duration of the word and perhaps also in the duration of the vowel that precedes the nasal. The variable `SpeechRate` (as defined above on the basis of word durations) and the variable `precedingVowelDuration` thus cover, at least to some extent, effects of accentuation as well.
Number of Syllables and Segments in the Word. Apart from word duration, there are two other measurements 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. (2015) have shown the same effect for word-final /s/ in English. To take these facts into account, we used 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 transcribers, and the number of segments in the word as coded by the transcribers. All three of these measurements are highly correlated, with the number of segments being the most fine-grained one. It can be assumed that the more segments are part of a word, the shorter the duration of each segment becomes, hence also of the nasal in question. This variable, however, is already represented as part of our SPEECHRATE variable, so that there is indeed a strong correlation between the number of segments and speech rate. For the reasons outlined in section 3.3 we therefore decided not to include NUMBEROFSEGMENTS (or any of the other two phonological length variables that strongly correlate with it) as a separate variable after all.

Utterance 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 (e.g. 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). We therefore coded our data for utterance position, but preliminary inspection revealed that, similar to Hay’s study, nasal duration did not increase when the segment in question was part of an utterance-final word. We therefore did not include this variable in our models.

Stress. Stressed syllables tend to have a longer duration than unstressed syllables (e.g. Fry 1955, 1958; Lieberman 1960; Beckman 1986; Harrington et al. 1998, see also Laver 1994 for an overview). Thus, if /an/ or /um/ bear stress, the nasal in the prefix might be longer. It is uncontroversial that in- is normally unstressed, but that there are also some derivatives in which the prefixes carry main stress (e.g infinite, impotent). In our data set there were no words with main-stressed prefixes.

The prefix un- is taken to be unstressed, but pronunciation dictionaries such as Wells (2008) note not only unstressed un- (as in unfathomable), but also secondarily stressed un- (e.g. in undefined), and optionally stressed un- (e.g. in unleash). It is unclear on which basis Wells assigns the stress marks. A closer look reveals, however, that a secondary stress mark (or an optional stress mark) is assigned to un- when the prefix is followed by an unstressed syllable, and no stress mark is assigned when the prefix is followed by a main-stressed syllable (unless the base is monosyllabic). Irrespective of whether there is any reality to the distinction between secondarily and unstressed un- as given in Wells (2008), we coded whether the prefix was followed by a stressed syllable or by an unstressed syllable. This seemed to be potentially relevant as Umeda (1977) found that nasals before unstressed vowels are shorter. The variable STRESSPATTERN
was coded with two levels: beforeStressed and beforeUnstressed.

**Frequency.** Frequency has also been shown to affect the duration of a word. More frequent words tend to have shorter durations (see, for example, Aylett and Turk 2004, Gahl 2008). Frequency was therefore included as a covariate. We collected two different types of frequency (again from COCA, Davies 2008), word form frequency and lemma frequency. While word form frequency refers to all words in the corpus which have the same word form as the items of our data set, by searching for lemma frequency we obtained the number of all words belonging to the same lemma as the investigated item. Lemma frequency thus also takes words with different forms into account, such as for example a verb’s past tense form (e.g. *immigrate* and *immigrated*). Preliminary inspection of our data revealed, however, that the two frequency measurements highly correlated. Hence, we decided to only include one in our models, WordFormFrequency. We log-transformed this variable before it entered the model.

### 3.5 Statistical analysis

We fitted one linear regression model to the *un*-data set (‘un-model’) and one linear regression model to the *in*-data set (‘in-model’). In both models the absolute duration of the nasal in seconds (AbsoluteNasalDuration) was used as the dependent variable. We also fitted models with relative duration as the dependent variable, but these models proved to be much less powerful, and had fewer significant predictors than models that had absolute duration as the dependent variable and the duration of the preceding vowel as a covariate. We therefore restricted our study to absolute duration as the dependent variable.

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 these variables as random effects.

The following modeling strategy was adopted, following established practices in the field (e.g. Baayen 2008). First, we conducted an initial model incorporating all of the above mentioned variables. We then looked at the residuals of the model, which need to be normally distributed. If visual inspection revealed that the residuals had a non-normal distribution, we used transformations and the exclusion of outliers to obtain the desired pattern. In both models transformation of the dependent variable AbsoluteNasalDuration was necessary to alleviate problems of non-linearity. Following Plag et al. (2015), Following Plag et al. (2015), we used the R function `boxcox()` to determine the optimal parameter lambda for a Box-Cox power transformation of the response variable (Box and Cox 1964; Venables and Ripley 2002).

We then checked for collinearity in our models by looking at the correlations between potentially correlated variables, and followed the strategies described in section 3.3. We also tested for interactions between the variables of interest (i.e. number of nasals, the affix, semantic transparency and frequency measures). These interactions turned out to be statistically not significant.

The regression models were simplified by stepwise excluding insignificant predictors. A predictor was considered significant if its absolute t-value was higher than 2.0. Predictors were also excluded if the Akaike Information Criterion (AIC) of the model including the
factor 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.

The plots of the models were generated with the visreg package (Breheny and 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 better interpretability of the plots the response variable AbsoluteNasalDuration is back-transformed to seconds.

Tables 7 and 8 give an overview of the variables initially included in each model and summarize their distributions.

Table 7: Summary of dependent variable and covariates used in the initial model for un-, $N=158$

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbsoluteNasalDuration</td>
<td>0.060</td>
<td>0.028</td>
<td>0.016</td>
<td>0.137</td>
</tr>
<tr>
<td>Numerical Predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>logRelativeFrequency</td>
<td>-0.797</td>
<td>2.518</td>
<td>-8.495</td>
<td>7.098</td>
</tr>
<tr>
<td>SpeechRate</td>
<td>13.190</td>
<td>2.990</td>
<td>6.136</td>
<td>20.570</td>
</tr>
<tr>
<td>PrecedingVowelDuration</td>
<td>0.087</td>
<td>0.029</td>
<td>0.020</td>
<td>0.167</td>
</tr>
<tr>
<td>logWordFormFrequency</td>
<td>7.182</td>
<td>1.953</td>
<td>0.000</td>
<td>9.838</td>
</tr>
<tr>
<td>Categorical Predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environment</td>
<td>n#nV: 23</td>
<td>n#C: 68</td>
<td>n#V: 67</td>
<td></td>
</tr>
<tr>
<td>StressPattern</td>
<td>beforeStressed: 102</td>
<td>beforeUnstressed: 56</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Summary of dependent variable and covariates used in the initial model for in-, $N=157$

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>AbsoluteNasalDuration</td>
<td>-2.656</td>
<td>0.374</td>
<td>-3.902</td>
<td>-1.885</td>
</tr>
<tr>
<td>Numerical Predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SpeechRate</td>
<td>14.300</td>
<td>3.659</td>
<td>5.279</td>
<td>24.320</td>
</tr>
<tr>
<td>PrecedingVowelDuration</td>
<td>0.0608</td>
<td>0.0255</td>
<td>0.00</td>
<td>0.127</td>
</tr>
<tr>
<td>logWordFormFrequency</td>
<td>8.578</td>
<td>1.769</td>
<td>1.609</td>
<td>10.700</td>
</tr>
<tr>
<td>Categorical Predictors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Environment</td>
<td>m#C: 67</td>
<td>m#mV: 90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affix</td>
<td>inLoc: 71</td>
<td>inNeg: 86</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SemanticTransparency</td>
<td>opaque: 62</td>
<td>transparent: 103</td>
<td></td>
<td></td>
</tr>
<tr>
<td>StressPattern</td>
<td>beforeStressed: 117</td>
<td>beforeUnstressed: 40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

18
4 Results

4.1 The prefix un-

A model was fitted according to the procedure described above. The Box-Cox transformation parameter for un- was 0.3434343. The residuals showed a non-normal distribution. Following standard procedures (e.g. Crawley 2002; Baayen and Milin 2010), we removed outliers (defined as items with standardized residuals exceeding $-2.5$ or $+2.5$) and refitted the model (the removal of outliers resulted in the loss of 4 observations, i.e. 2.5% of the observations). This resulted in a satisfactory distribution of residuals. After model simplification only two significant predictors remained, Environment and SpeechRate. The model explains 59% of the variation found in the data (Adjusted R-squared: 0.597).4

Table 9 summarizes the ANOVA results for the final model, and table 10 documents the estimates for each predictor and their p-values as found in the final model. The reference level for the categorical predictor Environment is n#nV.

Table 9: ANOVA of linear model for variables predicting the Box-Cox transformed duration of [n]

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>Sum Square</th>
<th>Mean Square</th>
<th>F value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environment</td>
<td>2</td>
<td>0.26</td>
<td>0.13</td>
<td>90.74</td>
<td>0.000</td>
</tr>
<tr>
<td>SpeechRate</td>
<td>1</td>
<td>0.08</td>
<td>0.08</td>
<td>52.46</td>
<td>0.000</td>
</tr>
<tr>
<td>Residuals</td>
<td>150</td>
<td>0.22</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10: Summary of linear model for variables predicting the Box-Cox transformed duration of [n]

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.539</td>
<td>0.015</td>
<td>37.106</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>Environment - n#C</td>
<td>-0.051</td>
<td>0.009</td>
<td>-5.372</td>
<td>2.91e-07</td>
</tr>
<tr>
<td>Environment - n#V</td>
<td>-0.101</td>
<td>0.010</td>
<td>-10.587</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>SpeechRate</td>
<td>-0.008</td>
<td>0.001</td>
<td>-7.243</td>
<td>2.14e-11</td>
</tr>
</tbody>
</table>

For SpeechRate table 10 shows that with increasing speech rate, the duration of the nasal in un-prefixed words decreases (Estimate=−0.007713). Figure 3 depicts this effect. The y-axis of the graph displays the duration of the nasal closure in seconds, the horizontal axis represents the speech rate. Each dot represents one observation of the data set. 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,

---

4The adjusted R-squared for a model with only SpeechRate is 0.272, a model with only Environment has an R-squared value of 0.466.
rate, i.e. the more segments are pronounced in a shorter amount of time, the shorter becomes the nasal. This is an expected effect.

Figure 3: Effect of SpeechRate on nasal duration in un-data set

Let us now turn to our variable of interest, Environment. The coefficients for $n\#C$ and $n\#V$ both go in the same direction and are highly significant. 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 4 illustrates this effect.

Figure 4: Effect of Environment on nasal duration in un-data set

The predicted mean duration for double nasals is 90 milliseconds. For words with the
environment the nasal is predicted to be 63 milliseconds long, and 42 milliseconds for words having the n#V environment. 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 milliseconds as against 2 times 42 milliseconds). This results 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 milliseconds, 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 phonological gemination has shown, however, that the durational differences between geminates and their corresponding singletons may vary substantially (see, for example Aoyama 2001; Aoyama and Reid 2006; Galea et al. 2014; Mattei and Di Benedetto 2000). For morphological geminates across word boundaries, Delattre (1968) has demonstrated that different languages show different singleton-to-geminate duration ratios. For English, he finds an increase from singleton to geminate /n/ of 50%. The geminates in that study straddled a word boundary and were surrounded by vowels (as in I've seen Nelly). Umeda (1977) finds in his data that a word-initial nasal is 71 ms long, and that a geminate straddling a word-boundary is 100 ms long (again, in exclusively vocalic environments). Oh and Redford (2012, 86, Figure 2) arrive at an estimated 82 ms for word-internal singletons, 110 milliseconds for un-geminates, and 131 for geminates across word boundaries (again, only vocalic environments were tested).

It thus seems that neither phonological gemination nor morphological gemination necessarily go together with a doubling of phonetic duration, but both types of gemination can result in much smaller durational increases. There is thus good reason to believe that even the smaller of the two contrasts in our data (i.e. the one between n#C vs. n#nV) can be interpreted as good evidence for gemination. The longer duration of the pre-consonantal (as against the pre-vocalic) singleton /n/, which results in a smaller durational difference to the geminate, can be attributed to the type of following segment (C vs. V).

4.2 The prefix in-

The model for the in-data set was fitted according to the same procedure as the previous model. The Box-Cox transformation of the dependent variable AbsoluteNasalDuration resulted in lambda=0.4646465. We removed outliers in the same way as before, losing two items. The final model showed a satisfactory distribution of residuals. The model explains about 52 % of the variation (Adjusted R-squared=0.5272). The ANOVA of the final model and an overview of the model coefficients are given tables 11 and 12.

---

5Oh and 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.
Table 11: ANOVA of linear model for variables predicting the Box-Cox transformed duration of [m]

<table>
<thead>
<tr>
<th></th>
<th>DF</th>
<th>Mean Square</th>
<th>F value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environment</td>
<td>1</td>
<td>0.086</td>
<td>75.795</td>
<td>5.284e-15</td>
</tr>
<tr>
<td>SpeechRate</td>
<td>1</td>
<td>0.043</td>
<td>38.126</td>
<td>5.921e-09</td>
</tr>
<tr>
<td>StressPattern</td>
<td>1</td>
<td>0.061</td>
<td>53.891</td>
<td>1.249e-11</td>
</tr>
<tr>
<td>Affix</td>
<td>1</td>
<td>0.009</td>
<td>7.898</td>
<td>0.006</td>
</tr>
<tr>
<td>Residuals</td>
<td>150</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 12: Summary of linear model for variables predicting the Box-Cox-transformed duration of [m]

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Error</th>
<th>t value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.361</td>
<td>0.014</td>
<td>25.923</td>
<td>&lt; 2e-16</td>
</tr>
<tr>
<td>ENVIRONMENT - m#C</td>
<td>-0.046</td>
<td>0.007</td>
<td>6.856</td>
<td>1.72e-10</td>
</tr>
<tr>
<td>SpeechRate</td>
<td>-0.003</td>
<td>0.001</td>
<td>-4.244</td>
<td>3.84e-05</td>
</tr>
<tr>
<td>StressPattern - beforeUnstressed</td>
<td>-0.036</td>
<td>0.007</td>
<td>-4.865</td>
<td>2.86e-06</td>
</tr>
<tr>
<td>Affix - inNeg</td>
<td>0.020</td>
<td>0.007</td>
<td>2.810</td>
<td>0.006</td>
</tr>
</tbody>
</table>

The final model includes four variables with a significant effect on nasal duration: Environment, SpeechRate, StressPattern and Affix. Figure 5 displays the effects of the two noise variables on nasal duration. The left panel shows the effect of SpeechRate. This effect is as expected: the higher the speech rate, the shorter the nasal. In the right panel, the estimated mean duration of nasals by stress pattern is plotted. With an estimated mean duration of 75 milliseconds, the nasal is 21 milliseconds shorter before an unstressed syllable than before a stressed syllable (95 milliseconds). This result is expected, too. Umeda (1977) also found that nasals before unstressed vowels are shorter.
Figure 5: Effects of speech rate and stress pattern on nasal duration for words prefixed with *in*-

Let us turn to the variables of interest. The left panel of figure 6 displays the effect of ENVIRONMENT. Double nasals are significantly longer than singles. The estimated mean duration for double nasals is 95 milliseconds, while it is 69 milliseconds for single nasals, a difference of 26 milliseconds.

The effect of AFFIX is displayed in the right panel of figure 6. The nasal in negative *in-* is significantly longer (by 12 milliseconds) 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.

Figure 6: Effects of environment and affix on nasal duration for words prefixed with *in*-

23
4.3 Summary

In both models the phonetic variable \textit{SpeechRate} had the expected effect on nasal duration. The higher the speech rate, the shorter the duration of the nasal. For \textit{in-} we could also observe an influence of \textit{StressPattern} on nasal duration, in the expected direction.

Regarding our variables of interest, we can conclude that both models revealed that items containing a double consonant, i.e. a n#nV or a m#mV structure, showed significantly longer nasal durations than words with only a single underlying nasal, i.e. a n#C, n#V or a m#C structure, respectively. For those data for which the type of following segment was the same (i.e. a vowel), nasal duration more than doubled, which is a clear indication of gemination.

Keeping the environment constant across \textit{un-} and \textit{in-}, i.e. if we compare the double nasal duration with the duration of the nasal before consonant, the two suffixes display very similar kinds of duration differences. Using the same method of computation as for the partial effects plots, the models predict 27 milliseconds difference for \textit{un-}, 26 milliseconds for negative \textit{in-}, and 24 milliseconds for locative \textit{in-}. This means that all three prefixes geminate.

Furthermore, the duration of the nasal in \textit{in-} depends on which \textit{Affix} it represents. The nasals in items in which \textit{in-} had a negative meaning was longer than in the ones where it had a locative meaning. In both models, we could not find any effect of decomposability. Neither \textit{RelativeFrequency}, nor \textit{SemanticTransparency} had a significant influence on nasal duration in \textit{un-} and \textit{in-}prefixed words.

5 Discussion and Conclusion

In this section we will return to our research questions and discuss each of them with regard to the results yielded in our two corpus studies.

While the majority of the literature agrees on the assumption that \textit{un-} geminates and \textit{in-} degeminates, the empirical data investigated in this study show that both prefixes geminate. For both \textit{un-} and \textit{in-}prefixed words the morphological geminates displayed a significantly longer duration than their single counterparts. The present study thus replicates earlier experimental findings by Kaye (2005) and Oh and Redford (2012) with natural conversation data and does so for a much larger set of pertinent words.

The fact that the prefix \textit{in-} clearly geminates refutes widely-held assumptions about the gemination pattern of this prefix and challenges the alleged systematic difference between \textit{un-} and \textit{in-}. The empirical facts also invalidate the idea put forward by proponents of Lexical Phonology that \textit{in-} as a level 1 prefix should geminate. Our results tie in with other research that has found variable behavior of affixes with regard to the alleged level a given suffix should belong to. For example, Giegerich (1999, chapter 2) discusses a long list of affixes that show a morpho-phonological behavior that suggests double membership for these affixes. The prefix \textit{in-} is somewhat similar, in that it does not geminate, i.e. shows a level 2 behavior, but nevertheless shows assimilation, which is taken to be a level 1 characteristic (e.g. Borowsky 1986). This is further evidence that each affix comes with its own morpho-phonological restrictions, as demonstrated by Raffelsiefen (1999), who found that not even two of the many suffixes of English trigger exactly the same type
of morpho-phonological alternations. In other words, in order to really account for the morpho-phonology of English affixes, it seems that we would need as many strata as we have suffixes that trigger morpho-phonological alternations (see also Bauer et al. (2013, Chapter 27) for discussion).

The second research question was concerned with the difference between negative in- and locative in-, and whether the difference between these two prefixes is mirrored in their phonetic implementation. In our data we indeed found a significant difference in nasal duration between negative in- and locative in-, with the nasal in negative in- being longer than the one in locative in-. Thus, the two affixes display a difference in phonetic implementation.

This finding has serious implications for speech production models and for theories which take phonological representations as the sole base for articulation, such as Lexical Phonology and feed-forward speech production models (e.g. Levelt et al. 1999). In these models no morphological information is passed on to phonetics or articulation together with a particular phonemic string. Identical phonemic strings, such as the ones we investigated (im-), should thus display the same acoustic properties and there should not be any systematic, meaning-related or morphologically motivated difference.

That phonologically homophonous forms show systematic phonetic differences is, however, not a new finding. Gahl (2008) and Drager (2011) found durational and other differences for homophonous lexemes, Plag et al. (2015) found acoustic differences between homophonous S morphemes in English, and Smith et al. (2012) showed that the word-initial strings mis- and dis- display different acoustic properties depending on whether they represent prefixes or only pseudo-prefixes. The studies by Kemps et al. (2005a,b) and Blazej and Cohen-Goldberg (2015) demonstrated that these subtle differences in phonetic detail have an influence on lexical processing.

The present study thus contributes to the growing body of literature that provides empirical evidence that underlying representations cannot be the sole base for articulation. Identical phoneme strings are not always articulated in the same ways, and this variability cannot be explained away as general online phonetic influences (such as, for example, coarticulation or speech rate). Rather, and on top of the general phonetic influences, one can observe variation depending on word-specific information, including morphological structure.

The third research question was concerned with the role of morphological decomposability. We tested whether two measures of decomposability, semantic transparency and relative frequency, have an influence on the duration of the nasal in un- and in-prefixed words. Neither of the two measures yielded a significant effect, i.e. in both models a word’s decomposability did not affect the duration of the nasal.

Our data came from a speech corpus, and the limitations of this corpus did not allow us to investigate a very large number of types. This had the unfortunate consequence that the range of relative frequency was quite small. Not finding an effect of relative frequency may therefore be due to a lack of statistical power. In general, finding an effect of relative frequency seems not easy. There are other studies besides the present one that had difficulties in replicating the alleged effect of relative frequency on phonetic reduction.

Schuppler et al. (2012) and Hanique et al. (2013) find no general effect of relative frequency on the reduction or deletion of 3rd person present tense /t/ in Dutch, but it
may be the case that relative frequency is a measure that does not work well for inflection. Notably, Hay’s research into relative frequency effects dealt with derivational morphology. A puzzling result emerges from Pluymaekers et al. (2005a). These authors investigate some Dutch affixes and find an effect of relative frequency with one of the affixes (i.e. the prefix ge-). Surprisingly, the effect goes in the unexpected direction: the fricative of this prefix is longer in those words that should be less parsable (going by their relative frequency). In contrast, Hanique et al. (2010) also investigate Dutch ge- and do find an effect in the expected direction, but only at a lower speech rate.

Obviously, more research is needed to clarify the role of relative frequency in the phonetic implementation of complex words. With regard to the prefixes under discussion in this paper, larger data sets are necessary, with more different types.

Let us turn to another measure of decomposability, semantic transparency. In our data, we did not find an effect of semantic transparency on nasal duration. This failure may in part be due to the relatively conservative coding we employed for this variable. A binary coding that classifies everything as non-transparent that is not fully transparent is perhaps not fine-grained enough to detect more subtle, gradient effects of transparency. Note, however, that a more fine-grained rating of morphological complexity might not be very fruitful. Bürki et al. (2011) used such ratings but were not able to find a systematic effect of this variable on duration (of schwa in French) either. Again, more research is needed to clarify the role of semantic transparency in phonetic implementation.

Another issue we raised in the introduction to this paper is whether degemination is a gradient or a categorical phenomenon. Since we did not find any degemination effect, we cannot say anything about this question based on our data.

To summarize, this study has shown that traditional assumptions about the gemination and degemination of certain prefixes in English are not borne out by the data. Both negative and locative in- geminate. It also turned out that the two phonologically homophous in-prefixes can be distinguished by the acoustic duration of the nasal. None of the extant morphological theories or theories of speech production are able to accommodate the effects that we find, and more research is needed to clarify the role of phonetic detail in morphological organization and morphological processing.

References


R Development Core Team (2014). R: A Language and Environment for Statistical Computing.


