Eberhard Karls University Tübingen
Faculty of Humanities
Department of General and Computational Linguistics

The Variability of German "eigentlich"
-
Investigating phonetic variations and the conditions under which they occur.

A THESIS SUBMITTED FOR THE DEGREE OF
M.A. GENERAL LINGUISTICS
Submission Date: March 31, 2023
First Examiner and Supervisor: Dr. habil. Fabian Tomaschek
Second Examiner: Dr. Ching-Chu Hendrix-Sun

Submitted by
Name: Laura Benthaus
Address: Jürgensenstr. 32, 72074 Tübingen
Number of Matriculation: 5505584
E-Mail: laura.benthaus@student.uni-tuebingen.de

Abstract

In 2018, Ernestus and Smith presented their findings on the variability of the Dutch *eigenlijk* (*actually*) and the context in which this variability occurs in conversational speech. This exceptional paper contributed to the knowledge of phonetic variation. By replicating their analysis for the German equivalent *eigentlich*, the present paper aims to contribute to that understanding as well.

The inspection of *eigentlich* shows a wide range of variability with almost 170 different variations ranging from mono- to trisyllabic forms. By means of mixed-effects modeling, the predictors of the word's variation were analyzed. The tokens appear shorter in duration and form with a higher speech rate, when being followed by highly predictable words, and when being preceded by pauses or words of low surprisal. Feature sharing and low paradigmatic uncertainty of the preceding word accompanied less reduced forms. Overall, the findings emphasize the importance of the context's syntagmatic and paradigmatic probabilistic information for variation processes, which supports the *Smooth Signal Redundancy Hypothesis* by Aylett and Turk (2004). Moreover, the findings contribute to Kuperman et al. (2007)'s *Paradigmatic Signal Enhancement Hypothesis*. The present study follows up on the results of Ernestus and Smith (2018)'s study and agrees that more qualitative and quantitative research on one word's variability is necessary to extend our understanding of the phenomenon and of what is stored in the mental lexicon.

Contents

List of Figures

List of Tables

1 Introduction

1.1 Overview

In their 2005 published paper, Keune et al. reported that Dutch words with the suffix *-lijk* tend to show strong variability in conversational Dutch. Based on such findings, Ernestus and Smith (2018) recently demonstrated the strong variability for the Dutch *eigenlijk* (*actually*). In 2018, the authors published their outstanding research on the qualitative and quantitative features of the word's variation. The present paper aims to conduct a similar analysis for the German equivalent *eigentlich* to extend the understanding of variation in speech production.

For that purpose, the necessary data from the *Karl Eberhard Corpus* (Arnold and Tomaschek, 2016) was collected and analyses of predictors based on (generalized) linear mixed-effects modeling were conducted. With regard to the analyzed predictors, information-theoretic measures (predictability, paradigm entropy, and surprisal), temporal and prosodic measures (speech rate and presence of pauses), and co-articulation measures (phonetic feature sharing) were chosen.

The analyses emphasize that the German discourse marker has a wide range of variability in spontaneously produced speech (almost 170 forms) paralleling its Dutch counterpart (Niebuhr and Kohler, 2011). Almost all of the tokens of the present data are either of disyllabic (ca. 66%) or monosyllabic nature (about 36%). The analyses reveal that tokens appear to be shorter in word duration and shorter in the amount of segments with a higher rate of speech and when being followed by highly predictable words. In addition, pauses and words of low surprisal preceded shorter tokens. Feature sharing and a low paradigmatic uncertainty of the preceding word accompanied less reduced tokens.

The present results lead to the conclusion that, among other factors, the probabilistic information of the context as well as the paradigmatic uncertainty within all contexts play a role in variation processes of speech production. The findings support information-theoretic frameworks, such as the *Smooth Signal Redundancy Hypothesis* by Aylett and Turk (2004) and the *Paradigmatic Signal Enhancement Hypothesis* by Kuperman et al. (2007). Moreover, the results are in agreement with

Ernestus and Smith (2018) and Milin et al. (2009a) regarding the discussion of what is stored in the mental lexicon.

In the next sections, I will discuss the literature background on speech variation theories and how they explain the phenomenon. Following, an overview of *eigentlich*'s phonetic variation will be presented in Chapter 2 as well as the analysis of the conditions in Chapter 3. Lastly, the results of the present study will be discussed and summarized.

1.2 Background

1.2.1 Variation in Conversational Speech

The extend of variation and reduction in conversational speech has previously been reported in Johnson (2004)'s study. He demonstrated that reduction is a highly frequent phenomenon in conversational speech of American English. More than just investigating the amount of variation that occurs, other studies considered how the reduced word forms appear in conversational speech. For Dutch, for instance, Keune et al. (2005) reported that high-frequency words that contain the suffix *-lijk* are highly reduced as a function of socio-geographic variation. The authors illustrated that the trisyllabic words *natuurlijk* (of course) or *mogelijk* (possible) have highly frequent monosyllabic forms, which can look like [mok] for *mogelijk* or [tyk] for *natuurlijk* (Keune et al., 2005).

Based on such findings, one exceptional paper on speech variation has been published in 2018 by Ernestus and Smith. The authors documented the many variations of *eigenlijk* (*actually*) in conversational Dutch. They report that the word's variability ranges from trisyllabic forms, such as [ɛɪxələk] or [ɛɪɣələk], to disyllabic and monosyllabic variations, like [ɛɪxlək], [ɛɪxək], [ɛɪx], or [ɛɪk] (Ernestus and Smith, 2018). Though the trisyllabic forms come closest to the word's surface form, they occurred less frequently compared to the heavily reduced forms (Ernestus and Smith, 2018). Ernestus and Smith (2018) observed that the reduced variations occurred mostly "at higher speech rates and in phrase-medial position" (Ernestus and Smith, 2018, p.158). Moreover, they observed reduced tokens when the tokens followed a higher frequency word and when the following word's first syllable was unstressed. The authors conclude that speakers decide on the reduction degree based on the rhythm of the utterance next to other factors, such as contextual predictability or temporal aspects.

The documentation of phonetic variation in conversational speech "is essential for developing accurate acoustic models required for robust speech recognition" (Greenberg, 1999, p.173). Like Ernestus and Smith (2018)'s study, a lot of the research has been additionally concerned with formulating and testing theories that aim to explain why and when reduction occurs (Aylett and Turk, 2004). Therefore, the following section gives an overview of the literature background with a focus on the predictability of variation in speech production.

1.2.2 What Predicts Phonetic Variation?

Starting with psycholinguistic theories, Levelt (1989)'s theory of *Lexical Access* gives a framework of speech production. It focuses on information that are interacting in the process before actually articulating an utterance. The model consists of two systems: The first system is called *lexical selection* and refers to the competitive process of selecting the respective lexical item from the mental lexicon. The second system is called *form encoding* and entails the formation of a token's phonological code including prosodic and syllabic information. Levelt (1989) states that reduction is a mean for the speaker to gain speed and to facilitate a higher rate of communication. According to the author, how and when reduction appears depends on the characteristics of an utterance's syllabic and prosodic patterns. This locates the reduction processes to the second system of his theory. Though Levelt (1989)'s framework acknowledges a process of competition in the first system, he does not consider probabilistic information of the token or its context. However, it has been shown to be a relevant predictor for variation by, e.g., Jurafsky et al. (2001), Bell et al. (2003), Pluymaekers et al. (2005b), and Ernestus and Smith (2018). For instance, research showed that word frequency impacts anticipatory co-articulation (e.g., Tomaschek et al., 2018; Saito et al., 2020; Tomaschek et al., 2021a). Anticipatory co-articulation refers to forward utterance planning (Mildner, 2018). Such findings demonstrate "that phonetic variation is hard to predict simply on the basis of phonological rules" (Tomaschek et al., 2021a, p.62). Moreover, they imply that probabilistic information of the context plays a role in variation (e.g., Mildner, 2018; Tomaschek et al., 2021a).

Probabilistic information and its effect on variation has been researched by means of two types. The first one is syntagmatic probability, which focuses on the surrounding (present) context (Tomaschek et al., 2021b). For instance, Aylett and Turk (2004) based their *Smooth Signal Redundancy Hypothesis* on this type of predictability. In their information-theoretic framework, the authors focus on the notion of effective communication, which aims to achieve a proper information transfer. This entails that speakers generate utterances that exhibit a smooth signal profile, i.e., each item in the utterance has an even probability of being recognized by the interlocutor. Apart from achieving a proper transfer of information, the speaker aims to reduce speech effort (Aylett and Turk, 2004). A reduced speech effort can be achieved by producing reduced utterances, as their "lexical, syntactic, semantic, and pragmatic factors" (Aylett and Turk, 2006, p.3049) in the respective context already lead to a high predictability of being recognized.

Based on the *Smooth Signal Redundancy Hypothesis*, more studies on variation and reduction were published (Tomaschek and Ramscar, 2022). For example, Hume and Mailhot (2013) extended this theory by means of Shannon (1948)'s surprisal and entropy. The authors concluded that a speaker wants to preserve the appearance of high surprisal elements in an utterance so that the listener can follow. Other studies not only considered syntagmatic probabilisties, such as Jurafsky et al. (2001), Bell et al. (2003), Aylett and Turk (2004), or Pluymaekers et al. (2005b), but also paradigmatic probability. This second type of probability includes the present context as well as the other possible contexts that the token could occur in (Tomaschek et al., 2021b). Some studies indicate that a lower paradigmatic uncertainty can be connected to reduction processes. Tomaschek and Ramscar (2022) summarize that some of these studies report, e.g., higher deletion probabilities of segments or shorter segment duration. On the other hand, some studies report results for an opposite effect. Hence, a lower uncertainty was "associated with *enhancement*" (Tomaschek and Ramscar, 2022, p.3) of a word's characteristics. For example, Kuperman et al. (2007) found such an effect for Dutch in the duration of interfixes in compounds, Lõo et al. (2018) report longer duration for Estonian words, and Tomaschek et al. (2021b) report enhancement for "stem vowels of regular English inflected verb forms" (Tomaschek et al., 2021b, p.171). Based on their findings, Kuperman et al. (2007) formulated the *Paradigmatic Signal Enhancement Hypothesis* capturing the effect of enhancement.

The opposite findings of the effect of paradigmatic predictability on the word's appearance illustrate "that many aspects of the phenomenon of speech reduction are not yet well understood" (Ernestus and Smith, 2018, p.160). Though the direction of its effect is still discussed, the previously described research shows that paradigmatic uncertainty does have some effect on speech production. Other studies also confirm an effect for speech processing. For example, based on an earlier study by Moscoso del Prado et al. (2004), Milin et al. (2009a,b) included the information of Serbian verb stems and their inflectional paradigms into their research on lexical processing. They operationalized the inflectional paradigm of the verb by means of Shannon (1948)'s entropy. The authors could observe an effect of the paradigmatic uncertainty and competition within all possible contexts on lexical processing (Milin et al., 2009a).

In addition to psycholinguistic theories, articulatory theories can also contribute to the discussion of how to explain variation. They focus on the movement of the articulators. Browman and Goldstein (1992)'s *Articulatory Phonology* describes the phenomena of variation according to the organization and dynamics of articulatory gestures. Accordingly, variation in speech is a result of the gesture's decrease in magnitude, which leads to a gestural overlap (Browman and Goldstein, 1992). How gestures might overlap depends on the linguistic constraints of the respective language. For example, Wesener (2001) reported an overlap of word-final fricative [x] to word-initial nasals in German function words, which he partly attributed to "mechanisms of gestural overlap" (Wesener, 2001, p.22) referencing *Articulatory Phonology.* Though gestural aspects contribute to explaining phonetic variation, the focus of this chapter and the thesis at hand lies on probabilistic aspects that might predict variability. With this in mind, the overview on the literature background can be concluded by saying that variation is a complex phenomenon impacted by several aspects.

1.3 The Present Study

Zellers et al. (2018) and Ernestus and Smith (2018) state that qualitative and quantitative studies based on "large conversational speech corpora" (Zellers et al., 2018, p.7) provide important insights into "the conditions under which certain reductions are likely to occur" (Zellers et al., 2018, p.7). More of such research is needed in order to get a better understanding of variation in conversational speech (Ernestus and Smith, 2018). Therefore, the present paper aims to contribute to the understanding of variation in speech production by replicating the structure of Ernestus and Smith (2018)'s research for a study on German variation.

For that purpose, Chapter 2 documents the range of variation of the German *eigentlich* (*actually*) and its frequency of occurrence in the *Karl Eberhard Corpus* (KEC) (Arnold and Tomaschek, 2016). The word's variations serve as data for the analyses of predictors in Chapter 3. The previous section has shown that some of the previously investigated probabilistic conditions are more clear in their effect on variability than others. As Linke and Ramscar (2020) pointed out more recently: "speech is a highly structured system of nested communicative distributions" (Linke and Ramscar, 2020, p.21). Therefore, the present study will focus on probabilistic predictors of *eigentlich*'s variability in the KEC. Probabilistic measures, which are covered in the present study, are syntagmatic and paradigmatic predictability of the surrounding words. Moreover, temporal and prosodic measures will be analyzed next to co-articulation measures. The results of the analyses will be discussed in Chapter 4. Lastly, the overall findings of the present thesis will be concluded.

2 Phonetic Variation of *eigentlich*

2.1 Data Collection

Similar to Ernestus and Smith (2018)'s paper, the purpose of this chapter is to document the phonetic variation of *eigentlich* before investigating the conditions in which they occur. For that purpose, the data has been collected from the *Karl Eberhard Corpus*, which contains material "of spontaneously spoken southern German elicited in dialogues" (Arnold and Tomaschek, 2016, p.9). The corpus came in form of audio files as well as their corresponding *Praat* (Boersma and Weenink, 2020) TextGrid files. The TextGrid files are made up of four levels of linguistic information: The spoken words (word level), the way the words are pronounced (segment level), the canonical pronunciation (canonical level), and the type of word (part-of-speech level).

In a first step, more than 1080 tokens that contained *eigentlich* were taken from the corpus by means of the programming language *Python* (Van Rossum and Drake, 2009). They were produced by 78 male and female speakers with an age range of 19 to 33 years (Arnold and Tomaschek, 2016). The tokens were mainly adverbs, but six of them came in the form of quadrisyllabic adjectives. To resemble the data of Ernestus and Smith (2018) (see Figure 6 in the Appendix), the adjectives have been excluded from further analyses. The analyzed data then results in a total of 1077 tokens. These tokens have been manually annotated on the segment level.

In a second step, the word's two preceding and two following words have also been extracted for each token, which later enables the analysis in Chapter 3. Additionally, the duration and frequency of all tokens has been extracted from the corpus. Information on the syllable count of the *eigentlich*-tokens has been manually added.

The third step includes the calculation of probabilistic data (e.g., conditional probability or paradigmatic entropy) for the analyses in Chapter 3 by means of *Python*'s module *scipy.stats* (Virtanen et al., 2020).

The collected data can be viewed in file *DF2.csv* in the *GitHub* repository. The link to the repository containing the collected data and the programming scripts can be found in the appendix (A). The following sub-chapter presents an overview of the various findings of *eigentlich*'s variation in the KEC.

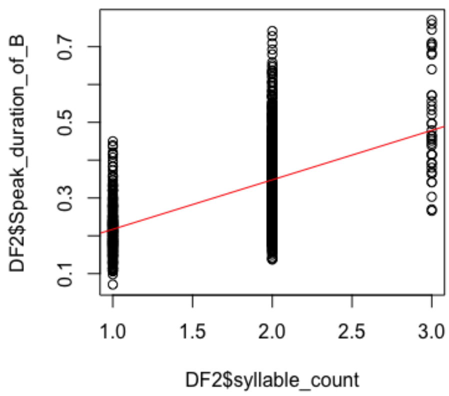
2.2 Variation of *eigentlich* in the KEC

In total, the KEC yielded almost 170 pronunciation variations of *eigentlich*. The following Table 1 only shows some of the most frequent variations as not all tokens can be included here. However, the presented variations make up more than half of the total occurrence of *eigentlich* in the KEC. The table is built similarly to the one in Ernestus and Smith (2018) (see Figure 6). Its focus lies on the token structure, which is divided into mono-, di-, and trisyllabic forms. For each structure, some of the most frequent variations are presented in addition to their number of occurrence.

Variations of *eigentlich*		
Token Structure	Number of Occurrence	Transcription
Trisyllabic		
(1) Vowel+ Plosive+ Nasal+ Plosive+ Lateral+ Vowel+ Fricative	9	[aignglɪç]
(2) Vowel+ Plosive+ Nasal+ Lateral+ Vowel+ Fricative	8	[aignlɪç]
(3) Vowel+ Plosive+ Nasal+ Plosive+ Lateral+ Vowel+ Fricative	4	[aignklɪç]
Disyllabic		
(1) Vowel+ Nasal+ Vowel+ Fricative	179	[aiŋɪç]
(2) Vowel+ Nasal+ Vowel+ Fricative	77	[ainɪç]
(3) Vowel+ Nasal+ Plosive+ Lateral+ Vowel+ Fricative	59	[ainglɪç]
Monosyllabic		
(1) Vowel+ Fricative	171	[aiç]
(2) Vowel	57	[ai]
(3) Vowel+ Nasal+ Fricative	14	[ainç]
(4) Vowel+ Fricative	12	[aiʃ]

Table 1: Examples of Variations in the KEC.

The canonical form of the word is transcribed as [aigɛntlɪx] in the KEC. Instead of [ɛ], the canonical form is sometimes also written with the mid central schwa as [aigəntlɪx], e.g., in Kohler (2001). Table 1 illustrates that trisyllabic tokens do not have the highest number of occurrence in the corpus, though the word's canonical form is trisyllabic, i.e., [ai] - [gɛnt] - [lɪx]. Only about three percent of the tokens consisted of three syllables, which makes the trisyllabic tokens the rarest forms in the KEC. The data confirms that the variations in the corpus are mostly of disyllabic nature (ca. 66%). About 36% of the tokens were monosyllabic. The most frequent variation is [aiŋɪç] with an occurrence count of 179. The second most frequent variation of *eigentlich* is the monosyllabic [aiç], which also occurred over 170 times. Regarding the tokens' duration, the variations with three syllables had a longer mean duration than variations with less syllables. On average, trisyllabic tokens had a duration of about 501 milliseconds (ms), disyllabic tokens had a mean duration of about 346ms, and tokens with one syllable had a mean duration of about 220ms. Similar to the results of Ernestus and Smith (2018), there is an overlap between the duration of mono-, di-, and trisyllabic variations, which can be seen in the first plot of Figure 1. The shortest duration was 71ms of a monosyllabic token, while on the other hand, the longest duration could be observed on a disyllabic token with 1.176 seconds. The second plot of Figure 1 observes a similar relationship for the token duration and the difference count of the token and the canonical form. It shows that a token with less segments usually has a shorter duration, though there are overlaps.



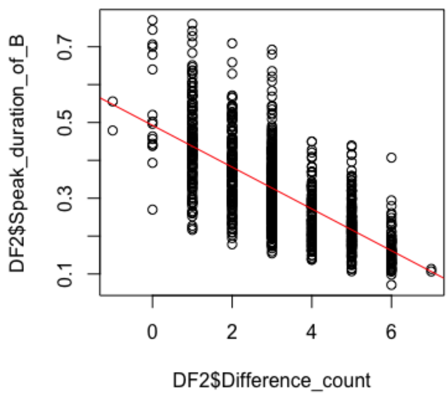


Figure 1: (1) Overlap Between Token Duration and Syllable Count. (2) Overlap Between Token Duration and Difference Count.

Generally, the findings of variation in the KEC match the findings of Kohler (2001) in the *Kiel Corpus of Spontaneous Speech*. Niebuhr and Kohler (2011) sum up Kohler (2001)'s findings by stating that for trisyllabic tokens, the first two syllables can be reduced to [aɪgŋ] and the last syllable might be [liç]. These syllables can either be connected by a [t] or [k] or no segment. Each part can occur in a more or less reduced variation. This can also be observed in Table 1 for the three most frequent trisyllabic tokens. For each of the three examples, the first two syllables appear as [aɪgŋ]. The final syllable generally occurs as [lɪç]. With regard to the "juncture between these [syllables]" (Niebuhr and Kohler, 2011, p.321), the present tokens have been found to be either connected by a [k], a [g], or by being absent.

The disyllabic tokens in the KEC appear to resemble this form. Their first syllable is usually [aiŋ] or [ain]. Hence, the [gn] shifted to the nasal [ŋ] or [n], which reduces the form to one initial syllable. The last syllable is usually reduced to [ɪç]. However, Table 1 contains one disyllabic variation that includes the full [lɪç] in addition to the plosive [g], which might be a remnant of Niebuhr and Kohler (2011)'s mentioned juncture between syllables that shifts to the final syllable. In the KEC, there are also 48 instances of the disyllabic form [ainklɪç] with [k] as well as four instances of disyllabic [aintlɪç] with [t] as former juncture.

Next to these forms, there are also the heavily reduced monosyllabic forms. As can be observed in Table 1, the initial diphthongal vowel [ai] of the syllable is always present in the most frequent forms. The syllable either ends with the fricative [ç] or [ʃ], or with no consonant. Only one of those presented monosyllabic forms includes a nasal in between ([ainç]).

When considering these heavily reduced forms, the question of the word's *phonetic essence*, which has been identified by Niebuhr and Kohler (2011), is also an important aspect of the research on variation. A phonetic essence contains "major phonetic characteristics of the "missing" segments" (Niebuhr and Kohler, 2011, p.328) that make the correct identification of the word possible for a listener. Niebuhr and Kohler (2011) state that the word's "palatality, nasality, and duration" (Niebuhr and Kohler, 2011, p.321) are essential parts, which are still present in reduced forms like [aiɪ]. The tokens in the KEC enable the identification of a general phonetic essence: It can appear heavily reduced as [ai(ŋ)(ɪ)(ç)]. After the vowel, the following segments, such as the consonants [ŋ] and [ç] and the vowel [ɪ] in between, appear to be optional. As can be seen in Table 1, the variation [aiç] is one of the most frequent variations of *eigentlich*, which only consists of a diphthongal vowel and a fricative. In form, this essence is similar to the one identified by Ernestus and Smith (2018), which includes a "diphthongal vowel and at least one [...] obstruent" (Ernestus and Smith, 2018, p.156).

Furthermore, the KEC yields 65 tokens with the fricative esh [ʃ] instead of [ç] as final segment. For example, there are 13 instances of the disyllabic form [aiŋɪʃ] and 12 monosyllabic tokens of [aiʃ] (see Table 1). Figure 2 on the following page shows an example taken from the KEC with a token being followed by the word *schon* (*already*). There are 14 similar cases in which the last segment of the token and the first segment of the following word is a postalveolar [ʃ], which might be attributed to a process of assimilation at word boundaries. Moreover, 38 tokens could be found in which the last segment is not the palatal fricative [ç] and the following first segment is an esh. For instance, there are 10 tokens of [ai], which are being followed by a word that starts with esh. This shows how assimilation of the final segment of the target token and the initial segment of the following one has an impact on the token's appearance.

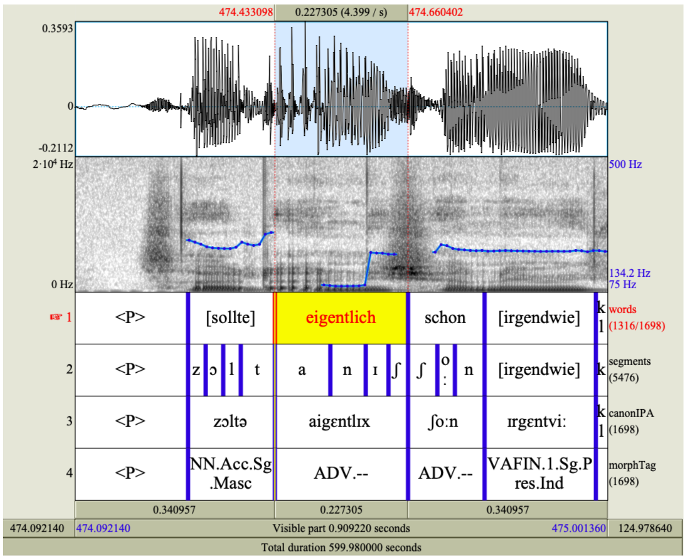


Figure 2: Example of Assimilation at Word Boundaries.

The next step now is to take a look at the conditions under which the here presented variations of the German word occur in. Therefore, the following chapter presents the analyses and their results.

3 Analysis of Predictors

3.1 Predictors

In order to investigate the conditions under which phonetic variability of the German token *eigentlich* occurs, the word's properties were taken to investigate whether they are affected by the chosen predictors.

The following two predictors have been reported to have an effect on phonetic variation by Ernestus and Smith (2018). The first one is the temporal measure of speech rate, which is defined as the phrase's syllable count divided by its duration. Here, a phrase consists of a fivegram with *eigentlich* in the center. Previous research found that a higher speech rate is more likely to result in more reduction (e.g., Bell et al., 2003; Pluymaekers et al., 2005a; Ernestus and Smith, 2018). A second factor is the (syntagmatic) predictability of the preceding and following word in the context of *eigentlich*. In the present analysis, the values were transformed logarithmically. For instance, Bell et al. (2003), Pluymaekers et al. (2005a), and Ernestus and Smith (2018) were able to show that the predictability of the surrounding words has an affect on a word's degree of reduction. While Bell et al. (2003) and Pluymaekers et al. (2005a) report the preceding and following word to affect variation, Ernestus and Smith (2018) report an effect mainly based on the preceding word.

The following two predictors are also part of the information-theoretic perspective. The first one is the amount of surprisal of the preceding and following word. Hume and Mailhot (2013) and Milin et al. (2009b) define surprisal as the "negative logarithm of [the token's] probability" (Hume and Mailhot, 2013, p.32) with a binary logarithmic base. It has been noted that a low surprisal results in more variability (Hume and Mailhot, 2013). The second one entails the paradigmatic uncertainty of the preceding and following word of the target token. The paradigmatic probability of the surrounding words, which is defined as all the conditioned contexts in which either the preceding or the following word occurs in the KEC, will be operationalized with Shannon (1948)'s entropy. The entropy formula is shown in Figure 3. It is defined as multiplying the token's probability by its surprisal (Hume and Mailhot, 2013). With this measure, the analysis might show if and how the paradigmatic uncertainty affects the token's reduction during speech production.

$$\begin{aligned} H_c(x_i) &= P(X = x_i) \cdot S(x_i) \\ &= -P(X = x_i) \cdot \log_2 P(X = x_i) \end{aligned}$$

Figure 3: Entropy Formula taken from Hume and Mailhot (2013, p.32)

As a prosodic measure, the analysis contains the fact whether the preceding or following element of *eigentlich* is a pause or not. A pause in the present data is defined as being tagged as "<P>" in the KEC, as can be observed in Figure 2 above. Bell et al. (2003) reported that surrounding disfluencies can impact a word's variability. Moreover, pauses extend the duration of a phrase and, therefore, can also result in a lower speech rate.

The last predictor is a measure of co-articulation. The aspect of phonetic feature sharing will be included in the analysis. This is defined as whether the preceding or following segment is present in the variation or not. In the present data, it is coded as Boolean variable. In addition, these predictors are operationalized more specifically by identifying three categories: Phonation (voiced, voiceless), horizontal position (front, central, or back), and the segment type (plosive, nasal, fricative, lateral). As Tomaschek et al. (2021a) reported, the aspect of feature sharing can play a role in phonetic variation, which will be further investigated in the present analysis.

3.2 Statistical Analyses

The analyses are focused on two continuous and two Boolean dependent variables, which illustrate properties of the word's variability. The first part of the analyses is concerned with the continuous dependent properties of *difference count*, which is defined as the number of segments of the canonical form minus the number of segments of the variation, and of *word duration*. These properties were analyzed by means of linear mixed-effects modeling. The second part of the analyses is concerned with two Boolean dependent properties of co-articulation. They are defined as whether the last segment of the preceding or the first segment of the following word shares a feature with the variation. They are analyzed with generalized linear mixed-effects modeling. The linear mixed-effects and the generalized mixed-effects models are part of the package *lme4* (Bates et al., 2015) in the programming language *R* (R Core Team, 2022).

In the following section, only those models will be presented that show a statistic significance in the analyses. Their significance is defined as the absolute t- or z-value being greater than 2.0. Speaker is a random effect in all present analyses. However, the focus of the present study does not lie on individual speaker differences.

3.3 Results

3.3.1 Models of Reduction Indicators

The present section starts with the results of the linear mixed-effects models. The first analysis is concerned with the predictors of difference count. The following Table 2 only shows the fixed effects that has a statistical significance. For the analysis of difference count, it is important to consider that a higher difference count means more reduction and a low difference count equals a less reduced variation. The mean difference count in the data indicates that tokens are generally reduced by three segments.

As predicted, speech rate is a significant predictor. The higher the speech rate, the higher the count of segment differences between the canonical form of the word and its variation.

With regard to context, the analysis shows that the probability of *eigentlich* given the following word affects the difference count. Hence, the higher the probability of the following word given the variation, the higher the difference count. The paradigmatic probability of the following word emerges as a significant predictor for this indicator of reduction as well. A high difference count is followed by a high paradigmatic entropy value. Neither for the conditional probability nor for the paradigm entropy does the preceding word have a significant effect on the word's variability. However, the surprisal of the preceding word shows an effect in the analysis. A low amount of surprisal of the preceding word is accompanied by a more reduced variation of *eigentlich*.

For the prosodic measure, only following pauses indicate a significant effect in the analysis. If a pause occurs after the variation, the difference count is lower and, therefore, the token is less reduced. This is an expected effect as pauses also contribute to a lower speech rate, which can result in lower degrees of reduction in a token. About 15% of the tokens are followed by a pause. Pauses preceding the variation (about 9%), however, do not show a statistically significant effect on the difference count.

Regarding co-articulation predictors, both the last segment of the preceding word and the first segment of the following word sharing features with the token are found to have a statistically significant effect on difference count.

Fixed effects	Difference Count
	t-value
Intercept	12.233
Speech rate	11.000
Prob of eigentlich given following word (log)	3.785
Entropy of following word (paradigm)	2.726
Surprisal of preceding word	-2.196
Following Pause	-4.179
Last Segment of preceding word in eigentlich	-6.419
First Segment of following word in eigentlich	-2.059
Random effects	Variance | SD
Speaker	0.5107 | 0.7146

Table 2: Results of the Analysis of Difference Count with Linear Mixed-effects Modeling.

The data reveals that the probability of the last segment sharing a feature with the variation is almost 25%, and the probability for the first segment is about 21%. The model illustrates the effect of feature sharing in a way that a low difference count occurs if the last segment of the preceding word shares a feature with the variation. Hence, feature sharing accompanies less reduction in the token. The same can be observed for the first segment of the following word sharing features with the variation. With a t-value of -2.059, the first segment predictor just barely passes the bench mark for being statistically significant. The ANOVA analysis of the model with and without this predictor shows that the model with the first segment predictor slightly improves the model (Chisq: 4.2592, Df: 1, p<0.05, AIC: 3566 vs. 3564).

In addition, these co-articulation measures have been specified by analyzing which types of features also occur in the token. It turns out that for the last segment of the preceding word, the categories *voiced*, *front vowel*, *plosive*, *nasal*, and *lateral* have a statistically significant effect on the outcome. This effect can, however, only be reported, if the categories are put into the model individually. If taken together, only the categories *voiced* (t-value: -6.747) and *plosive* (t-value: -2.566) show a significant impact on difference count. For none of the categories of the following first segment could be observed a significant effect in the model.

The model that includes the two significant categories of the last segment of the preceding word gives more detailed information about which segment type affects the difference count. This results in a model that is significantly improved compared to the one in Table 2 (Chisq: 9.6616, Df: 1, p<0.01, AIC: 3564 vs. 3557). The plot of the model, which can be observed in Figure 4, illustrates a normal distribution.

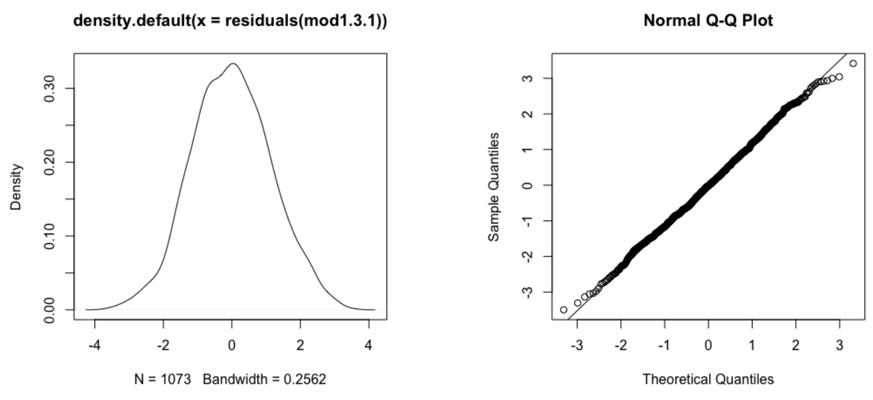


Figure 4: Normal Data Distribution of the Model with Specified Segment Predictors.

The second analysis of continuous outcomes investigates the predictors of word duration. The model is illustrated in Table 3 on the following page. The mean duration of the word's pronunciation is 312ms, which is a mean duration that is closest to the mean duration of the disyllabic tokens, as seen in Chapter 2. It almost matches the mean duration of tokens with a difference count of three (317ms) as well.

Similar to the first analysis, speech rate is a predictor for the present reduction indicator and, as expected, the analysis shows that the word's duration is shorter with a higher speech rate.

The pronunciation duration of the word is affected by the probability of the following word conditioned by the token. *Eigentlich* tends to be shorter if the predictability is higher. Neither paradigmatic probability nor surprisal of the following word show any significant effects on word duration.

The prosodic measure of the presence or absence of pauses surrounding the token indicates a statistically significant effect, though it differs for the preceding and the following pause. If a pause precedes *eigentlich*, the token has a shorter duration, while its duration is likely to be longer when followed by a pause.

In Table 3, it can be observed that the paradigmatic probabilistic information and surprisal of the preceding word do not affect the token's duration. A comparison of models shows that the surprisal of the preceding word only has a statistically significant effect on the outcome (t-value: 2.578), if the predictor of preceding pauses is excluded from the model. The comparison of both models indicates, however, that the prosodic measure makes the model slightly better with a lower Akaike Information Criterion value (AIC: -2058 vs. -2063).

Fixed effects	Word Duration
	t-value
Intercept	30.051
Speech rate	-16.011
Prob of eigentlich given following word (log)	-2.582
Preceding Pause	-3.508
Following Pause	9.549
Last Segment of preceding word in eigentlich	4.327
Random effects	Variance | SD
Speaker	0.00096 | 0.031

Table 3: Results of the Analysis of Word Duration with Linear Mixed-effects Modeling.

For co-articulation, the analysis only results in a statistically significant effect of the last segment on *eigentlich*'s duration. The tokens tend to be longer in duration when the preceding segment shares a feature with the target token. However, when specifying which type of feature sharing has an effect on duration, the model shows that the first segment of the following word sharing features with the variation does have an impact on duration as well. For both predictors, this will be further described in the following: When specifying the category for the preceding segment, each category by itself showed a significant effect on word duration: *front vowel* (t-value: 2.067), *voiced* (t-value: 3.816), *nasal* (t-value: 2.744), and *lateral* (t-value: 2.182). Including them together into the model reduced the significance of each predictor, though the predictor of *voiced* segments always remained significant. A comparison of the model in Table 3 with the model that includes the *voiced* segment could not identify an improvement of the model and emerges with a higher Akaike Information Criterion value (AIC: -2004 vs. -2008). For the first segment of the following word, its more detailed category of *plosives* can be observed to affect word duration (t-value: 2.514). Including this predictor into the model does not lead to a better fit of the model as the model of Table 3 has a lower Akaike Information Criterion (AIC: -2004 vs. -2008).

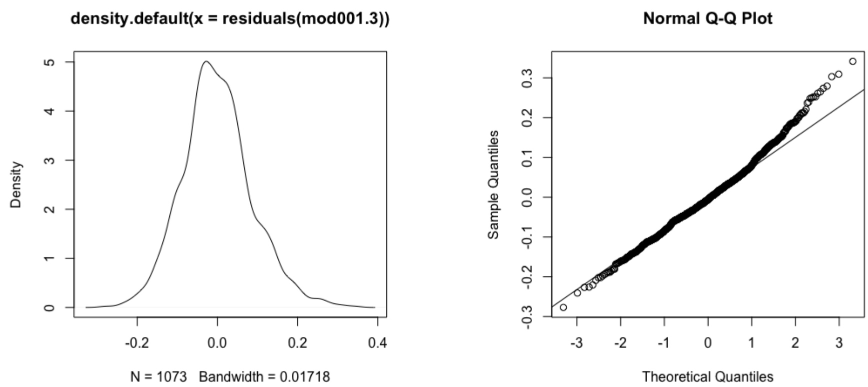


Figure 5: Slightly Right-skewed Data of the Model.

The plot of the analysis (in Table 3) appears to be slightly skewed on the right. Hence, it is a positively skewed distribution with more outliers on the right side of the distribution. This can be observed in the plots of Figure 5.

3.3.2 Models of Co-articulation

This section is concerned with the binomial analysis of co-articulation as outcomes. Hence, the analysis takes a closer look at what predicts whether the preceding segment or the following segment shares a feature with the variation.

The first model includes the last segment of the preceding word as response. In about one fourth of the tokens, the last segment of the preceding word is present in the token. Table 4 below shows which predictors are statistically significant. Interestingly, the predictor speech rate by itself does not have a significant impact on the outcome (z-value: 1.608, $p>0.1$). However, when including the predictors of probability and information into the model, it indicates that speech rate does have an impact on the outcome (see Table 4). Then, a higher speech rate makes it more likely for the preceding segment to share features with the target token.

The syntagmatic probabilities of the surrounding words have an impact on the outcome as well, though they differ in the way they affect it. It is more likely for the preceding segment to be present in the token if the probability of the token given the preceding word is high. Hence, the token and the preceding word's co-occurrence can determine whether the segment shares a feature with the token. If the probability of the token given the following word is high, the preceding segment does not appear to be present in the token.

Fixed effects	Last Segment
	z-value | p
Intercept	4.828 | <0.001
Speech rate	3.037 | <0.01
Prob of eigentlich given preceding word (log)	3.114 | <0.05
Prob of eigentlich given following word (log)	-2.412 | <0.01
Surprisal of preceding word	-6.827 | <0.001
Entropy of preceding word	-5.772 | <0.001
Random effects	Variance | SD
Speaker	0.028 | 0.167

Table 4: Results of the Analysis: Whether the Last Segment of the Preceding Word Shares Features with *eigentlich* with Generalized Linear Mixed-effects Modeling.

In addition, the information-theoretic predictors of the preceding word have a high significance of predicting the outcome (both $p<0.001$). A low surprisal of the preceding word tends to result in feature sharing. The paradigmatic probability of the preceding word affects the outcome in a way that a low entropy value makes the presence in the token more likely.

The prosodic measures of preceding or following pauses do not have a statistically significant effect on the outcome. That preceding pauses cannot have an effect on whether the last segment of the preceding word is present in the token is obvious, as the preceding word cannot be a pause and a word containing segments at the same time. In comparison, the presence or absence of following pauses could have an impact with regard to speech rate. However, the analysis shows that it does not have a significant effect (z-value: -0.848, $p>0.1$).

The last analysis is concerned with the binomial analysis of the first segment of the following word as outcome. About one fifth of the following segments are present at least once in the variation. In comparison to the previous analysis, including only speech rate into the model shows a significant effect of the predictor (z-value: 4.281, $p<0.001$). When checking for other variables that may predict feature sharing with the first segment, Table 5 on the following page shows that the only significant predictors of the following segment sharing features with the variation are speech rate and the surprisal of the following word. Speech rate is a highly significant predictor in a way that a higher speech rate accompanies feature sharing between the first segment of the following word and the target token. This outcome is additionally predicted by the aspect of surprisal of the following word. If it is a word of low surprisal, i.e., a highly predictable word, the first segment tends to share features with the variation. Adding surprisal of the following word to the model has a significant effect on improving it compared to a model that only includes speech rate (Chisq: 6.2315, Df: 1, $p<0.05$, AIC: 1114 vs. 1110).

Fixed effects	First Segment
	z-value | p
Intercept	-6.615 | <0.001
Speech rate	4.698 | <0.001
Surprisal of following word	-2.459 | <0.05
Random effects	Variance | SD
Speaker	0.1887 | 0.4344

Table 5: Results of the Analysis: Whether the First Segment of the Following Word Shares Features with *eigentlich* with Generalized Linear Mixed-effects Modeling.

4 General Discussion

The goal of the present thesis is to contribute to "our knowledge about speech reduction" (Ernestus and Smith, 2018, p.160) by demonstrating the variability of German *eigentlich* and investigating the context in which it occurs. The study and its results, which have been presented in the previous chapters, will now be summarized and discussed in the present chapter.

Similar to the work of Ernestus and Smith (2018), the present paper started with an overview of *eigentlich*'s variations found in the *Karl Eberhard Corpus*. The data of the KEC confirms the results of previous findings on *eigentlich*, which indicate that it "[demonstrates] a wide range of variation in the production of the word" (Ernestus and Smith, 2018, p.129). This variability appears to be "parallel to the Dutch particle *eigenlijk*" (Niebuhr and Kohler, 2011, p.321).

The present findings regarding the distribution of forms are similar to the findings of the Dutch counterpart. Almost all of the variations of *eigentlich* are either disyllabic (ca. 66%) or monosyllabic (ca. 31%). The trisyllabic forms, which are closest to the word's canonical form, only occur about three percent of the times in the KEC, which is an even lesser occurrence compared to the Dutch trisyllabic tokens in the *Ernestus Corpus* with 13% (Ernestus and Smith, 2018). In agreement with Ernestus and Smith (2018), this observation "raises the question which form of such a word should be considered as canonical" (Ernestus and Smith, 2018, p.160).

Niebuhr and Kohler (2011) suggested a *phonetic essence* for the German discourse marker, which Ernestus and Smith (2018) replicated for Dutch. The present findings of *eigentlich* support both of theirs. The tokens in the KEC usually have a varying form of [ai(ŋ)(ɪ)(ç)], which can be observed based on the most frequent forms in Table 1. Thus, the initial diphthongal vowel [ai] can be followed by optional nasal sonorants ([ŋ] or [n]), a near-close front vowel [ɪ], and/or the voiceless fricative [ç]. When considering the most frequent variations of *eigentlich*, its phonetic essence in speech might be expressed as [aiç], similar to the Dutch one identified by Ernestus and Smith (2018). However, even more reduced forms, which occur frequently with only the initial diphthongal vowel [ai], might provide enough phonetic characteristics for its correct identification given the proper conditions.

These proper conditions can be connected to anticipatory assimilation at word boundaries. This could be observed for the token's final palatal fricative [ç] to a following word-initial postalveolar fricative [ʃ]. Figure 2 in Chapter 2 presented the assimilation example of *eigentlich schon*. Other tokens missed the final fricative completely but were followed by the fricative esh, which contributes to the notion of articulatory planning. Hence, a speaker might take into consideration that the following word starts with the fricative [ʃ]. The production of two following fricatives requires more air and, therefore, it might be less effort for the speaker to focus on the production of only one (Aylett and Turk, 2004). As the final fricative appears to be optional in the high-frequency variations seen in Chapter 2, it can either be left out in the token or the token shares the same fricative ([ʃ]) with the following word. This finding can also be explained by means of *Articulatory Phonology*, which would explain that the magnitude of the final fricative decreases and an overlap of word-final and word-initial gestures occurs. To my knowledge, the observation of overlap between word-final and word-initial fricative has not been documented with regard to variability. Wesener (2001), for instance, only reported an overlap of word-final fricative [x] with word-initial nasals for German utterances, such as *noch mal* (*again*) (Wesener, 2001).

The second part of the paper at hand presented the statistical analyses and their results of the conditions under which the variation of *eigentlich* occur in the KEC. By means of linear mixed-effects modeling, the variables that may predict word duration and difference count have been investigated.

The effect of temporal aspects on variation can be explained with Levelt (1989)'s assumption that a higher speech rate favors reduction processes in order to facilitate a higher rate of communication (Levelt, 1989). Though temporal aspects are an important factor, "the degree of reduction is [not] only determined by how much time a speaker needs to plan and produce [an utterance]" (Ernestus and Smith, 2018, p.159). Ernestus and Smith (2018) emphasized the role of the phrase rhythm. However, this rhythm can be interrupted by disfluencies, such as pauses, as they might indicate planning problems (Bell et al., 2003). The effect of pauses is different for preceding pauses compared to following ones: When a pause precedes *eigentlich*, it is accompanied by a more reduced token. On the other hand, the effect of following pauses indicates less reduction. These findings point toward forward planning of an utterance, as the target can appear more reduced with the knowledge of the following word. Thus, the preceding pause might provide more planning time.

In line with the notion of forward planning is the present observation that a token is more likely to be reduced given its high predictability given the following word. While Ernestus and Smith (2018) found such an effect for the preceding word and Pluymaekers et al. (2005a) and Bell et al. (2003) report it for both neighboring words, the present analysis did not find an impact of the preceding one. However, these results generally confirm that the syntagmatic probability plays a role in reduction processes, which is in line with Aylett and Turk (2004)'s framework.

Moreover, the analyses could not identify an effect for the conditional predictability of the token's reduction degree given the preceding word. The surprisal of the preceding word, however, shows a significant effect and goes along with shorter tokens. Hume and Mailhot (2013) considered the amount of surprisal of the token itself. Here, the surprisal of the surrounding word shows an impact on the token's appearance. This confirms that a low surprisal preceding the token results in "patterns [that] are typically prone to [...] reduction processes" (Hume and Mailhot, 2013, p.31).

Regarding the amount of paradigmatic uncertainty in the token's context, the study shows that the paradigm of the following word affects the token. A low amount of paradigmatic uncertainty accompanies less reduced variations, which in the present data appears as variations with an amount of segments closer to the canonical form.

This finding points towards the *Paradigmatic Signal Enhancement Hypothesis* by Kuperman et al. (2007). More generally, it supports the notion that information and uncertainty take part in processes of speech production and variation (e.g., Greenberg, 1999; Hume and Mailhot, 2013).

Regarding the results of co-articulation and assimilation, feature sharing is observed to have an impact on variation. Feature sharing of the preceding and following segments is connected to fuller forms. Especially neighboring voiced segments as well as plosives share features with the variations of *eigentlich*. Though the present findings indicate that feature sharing of neighboring segments does not accompany reduction indicators, it still has an impact on a word's variation. This finding is in line with the presentation of findings by Tomaschek et al. (2021a), who also report that, among other factors, feature sharing affects a word's variability.

The second part of analyses took a closer look at the conditions under which *eigentlich* shared features with the neighboring segments. By means of generalized linear mixed-effects modeling, feature sharing has been investigated. For both, the preceding and the following segment, an effect of the word's syntagmatic and paradigmatic information was found, which favors the aspect of feature sharing. This is especially observable for the last segment of the preceding word. Apparently, co-articulation with preceding segments depends on the complete context of the utterance next to the paradigmatic uncertainty of the preceding word. These results suggest that the syntagmatic and paradigmatic probabilistic information are also taken into consideration for the planning of the utterance and the produced variability with regard to co-articulation (Bell et al., 2003; Tomaschek et al., 2021a). A higher speech rate is also correlated with assimilation, which is a predicted result by Browman and Goldstein (1992), who described the movement of articulators as timed events that can be affected by temporal aspects.

Overall, to put the discussed findings into a more general context, it can be summarized that they support Aylett and Turk (2004)' *Smooth Signal Redundancy Hypothesis*. The predictability of the token's respective context is required to provide sufficient information and less unexpectedness in order to enable reduction and co-articulation processes. These processes can decrease the speech effort for the speaker.

Moreover, the information that is considered in variation processes appears to not be limited to the present context, but extends to the paradigmatic information as well. This finding extends the effect of paradigmatic uncertainty on speech production, which follows up on Milin et al. (2009a)'s research, who reported an effect on processing.

Regarding the question of what is stored in the mental lexicon, the findings of the word's variations and their predictors contribute to the suggestion made by Ernestus and Smith (2018) as well as Milin et al. (2009a). As the present data made it possible to confirm a phonetic essence of *eigentlich*, it seems likely that the word's variability and the frequency of the variants might be included in the mental lexicon. In addition, this notion might be extended to the syntagmatic and paradigmatic information of the token's context being included as well. The observation of assimilation at word boundaries with regard to forward utterance planning indicates that the mental lexicon might include the "regularities" (Milin et al., 2009a, p.17) in which the variation can appear.

Further qualitative and quantitative studies based on large data sets are necessary for this research field. Variation studies based on large corpora containing spontaneous speech "can identify trends for certain speaking styles and allows for the calculation of sophisticated statistical models" (Zellers et al., 2018, p.7). This study did not intend to be a proper replication study of the one conducted by Ernestus and Smith (2018). It replicates their structure and analyses. Hence, some conditions that have been included by the authors have not been investigated in the present study, such as stress or phrase position. The focus of this study lies on the predictors of syntagmatic and paradigmatic probabilities of the surrounding words. Further investigation of *eigentlich*'s variability might, therefore, include the predictors that Ernestus and Smith (2018) analyzed. Such a study might enable a comparison between both findings. On the other hand, other contextual aspects, such as the surrounding parts of speech and their impact on the word's variability, should also be taken into consideration. Furthermore, the structure and/or the analyses might be replicated for a completely different word. For instance, Tomaschek et al. (2021a) found the German *irgendwie* (*somehow*) to occur in more than 100 different forms in the *Karl Eberhard Corpus*. A qualitative and quantitative analysis of the word's features might present new observations and findings regarding variability. Lastly, the present study has not taken into consideration individual speaker differences as Ernestus and Smith (2018) did. Thus, this might be an effect that can also be investigated more thoroughly in future studies of variation.

5 Conclusion

To summarize, the thesis at hand investigates *eigentlich*'s variability and the context in which it occurs. For that purpose, the conducted study by means of mixed-effects modeling includes measures of probability and information and of co-articulation, next to temporal and prosodic measures.

In conversational German, the word appears in a variety of forms. The most frequently produced forms of *eigentlich* are reduced by at least one syllable, which shows that the reduced forms are more common in conversational speech than the canonical form. The present findings support the notion that variability is a result of several factors in speech production. The focus of this study lies on the syntagmatic and paradigmatic information of the surrounding words, which have been shown to affect indicators of reduction and co-articulation. The effect of the syntagmatic and paradigmatic probabilistic information on *eigentlich*'s variability is in line with Aylett and Turk (2004)'s *Smooth Signal Redundancy Hypothesis*. Thus, it becomes evident that varied and reduced words are a mean to achieve the speaker's goal of effort reduction while providing a stable information transfer. Apart from the information of the present context, the analysis shows that all possible context of a word's occurrence are taken into consideration in the process of speech production as well, which points towards the *Paradigmatic Signal Enhancement Hypothesis* (Kuperman et al., 2007). Regarding the question of what might be stored in the mental lexicon, the study supports Ernestus and Smith (2018) and Milin et al. (2009a)'s conclusions that the possible variations, their frequency, and their contextual regularities may be involved.

This study as well as the one conducted by Ernestus and Smith (2018) make it clear that more research based on one word is necessary to understand the extend of variation in conversational speech.

References

Arnold, D. and Tomaschek, F. (2016). The karl eberhard corpus of spontaneously spoken southern german in dialogues - audio and articulatory recordings. In Draxler, C. and Kleber, F., editors, *Tagungsband der 12. Tagung Phonetik und Phonologie im deutschsprachigen Raum*, pages 9–11. München: Institut für Phonetik und Sprachverarbeitung.

Aylett, M. and Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and speech*, 47:31–56.

Aylett, M. and Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *The Journal of the Acoustical Society of America*, 119(5):3048–3058.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., and Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in english conversation. *The Journal of the Acoustical Society of America*, 113(2):1001–1024.

Boersma, P. and Weenink, D. (2020). Praat: doing phonetics by computer [computer program]. version 6.1.16. http://www.praat.org/.

Browman, C. P. and Goldstein, L. (1992). Articulatory phonology: An overview. *Haskins Laboratories Status Report on Speech Research*, (111/112):23–42.

Ernestus, M. and Smith, R. (2018). *5. Qualitative and quantitative aspects of phonetic variation in Dutch eigenlijk*, pages 129–163. De Gruyter Mouton, Berlin, Boston.

Greenberg, S. (1999). Speaking in shorthand – a syllable-centric perspective for understanding pronunciation variation. *Speech Communication*, 29(2):159–176.

Hume, E. and Mailhot, F. (2013). *The role of entropy and surprisal in phonolo-*

gization and language change, volume Origins of Sound Change: Approaches to Phonologization, chapter 2, pages 29–47.

Johnson, K. (2004). Massive reduction in conversational american english. *Proceedings Spontaneous Speech: Data and Analysis. Proceedings of the 1st Session of the 10th International Symposium*, pages 1–27.

Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. (2001). *Probabilistic relations between words: Evidence from reduction in lexical production*, pages 229–254. John Benjamins.

Keune, K., Ernestus, M., van Hout, R., and Baayen, R. H. (2005). Variation in dutch: From written mogelijk to spoken mok. *Corpus Linguistics and Linguistic Theory*, 1(2):183–223.

Kohler, K. J. (2001). Articulatory dynamics of vowels and consonants in speech communication. *Journal of the International Phonetic Association*, 31(1):1–16.

Kuperman, V., Pluymaekers, M., Ernestus, M., and Baayen, H. (2007). Morphological predictability and acoustic duration of interfixes in dutch compounds. *The Journal of the Acoustical Society of America*, 121:2261–71.

Levelt, W. J. M. (1989). *Speaking : From Intention to Articulation.* ACL-MIT Press Series in Natural-language Processing. A Bradford Book.

Linke, M. and Ramscar, M. (2020). How the probabilistic structure of grammatical context shapes speech. *Entropy*, 22(1):1–23.

Lõo, K., Järvikivi, J., Tomaschek, F., Tucker, B. V., and Baayen, R. H. (2018). Production of estonian case-inflected nouns shows whole-word frequency and paradigmatic effects. *Morphology*, 28(1):71–97.

Mildner, V. (2018). Aspects of coarticulation. In *CHALLENGES IN ANALYSIS AND PROCESSING OF SPONTANEOUS SPEECH*, pages 27–48, MTA Nyelvtudományi Intézet, Budapest.

Milin, P., Filipović Đurđević, D., and Moscoso del Prado Martín, F. (2009a). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from serbian. *Journal of Memory and Language*, 60(1):50–64.

Milin, P., Kuperman, V., Kostic, A., and Baayen, H. (2009b). *Words and paradigms bit by bit: An information-theoretic approach to the processing of inflection and derivation*, pages 214–252.

Moscoso del Prado, F., Kostic, A., and Baayen, H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition*, 94:1–18.

Niebuhr, O. and Kohler, K. J. (2011). Perception of phonetic detail in the identification of highly reduced words. *Journal of Phonetics*, 39(3):319–329. Speech Reduction.

Pluymaekers, M., Ernestus, M., and Baayen, H. (2005a). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica*, 62:146–59.

Pluymaekers, M., Ernestus, M., and Baayen, H. (2005b). Lexical frequency and acoustic reduction in spoken dutch. *The Journal of the Acoustical Society of America*, 118(4):2561–2569.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Saito, M., Tomaschek, F., and Baayen, H. (2020). An ultrasound study of frequency and co-articulation. In *12th International Seminar on Speech Production*, pages 206–209.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423.

Tomaschek, F., Philipsen, M., Zimmermann, M., and Linke, M. (2021a). Fifty shades of words – the relation between words spelled in a lexicon and actual articulations in german. In *Book of Abstracts - Zur Tagung "Phonetik & Phonologie 17"; https://drive.google.com/file/d/1ukXeIFWjJXm36S_JIivBrCMMKLOLIZxI/view*, pages 62–63.

Tomaschek, F. and Ramscar, M. (2022). Understanding the phonetic characteristics of speech under uncertainty—implications of the representation of linguistic knowledge in learning and processing. *Frontiers in Psychology*, 13.

Tomaschek, F., Tucker, B. V., Fasiolo, M., and Baayen, R. H. (2018). Practice makes perfect: the consequences of lexical proficiency for articulation. *Linguistics Vanguard*, 4(s2):20170018.

Tomaschek, F., Tucker, B. V., Ramscar, M., and Baayen, H. (2021b). Paradigmatic enhancement of stem vowels in regular english inflected verb forms. *Morphology*, 31(2):171–199.

Van Rossum, G. and Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272.

Wesener, T. (2001). Some non-sequential phenomena in german function words. *Journal of the International Phonetic Association*, 31(1):17–27.

Zellers, M., Schuppler, B., and Clayards, M. (2018). *1. Introduction, or: why rethink reduction?*, pages 1–24. De Gruyter Mouton, Berlin, Boston.

A Supplementary Material

The link to the collected data and the programming scripts:

`https://github.com/laurabnths/Variability_of_eigentlich`

Table 5.2: Tokens of *eigenlijk* found in the corpus. Phonemic and allophonic transcriptions are shown. The allophonic transcriptions specify the voicing of /k/ and /x/, but no other detail. Note that [ɛɪxl] was once perceived as disyllabic and once as monosyllabic.

Token structure	N	Transcription	
		Phonemic	**Allophonic**
Trisyllabic	**20**		
Vowel + fricative + schwa + lateral + schwa + stop	20	/ɛɪxələk/ 20	[ɛɪxələk] 17, [ɛɪxələg] 2, [ɛɪɣələk] 1
Disyllabic	**82**		
Vowel + fricative + lateral + schwa + stop	60	/ɛɪxlək/ 57	[ɛɪxlək] 42, [ɛɪxləg] 7, [ɛɪɣlək] 8, [ɛɪɣləg] 1
		/ɛɪxləŋ/ 1	[ɛɪxləŋ] 1
		/ɛxlək/ 2	[ɛxlək] 1, [ɛxləg] 1
Vowel + fricative + schwa + stop	15	/ɛɪxək/ 14	[ɛɪxək] 6, [ɛɪxəg] 2, [ɛɪɣək] 5, [ɛɪɣəg] 1
		/ɛxək/ 1	[ɛxəg] 1
Vowel + fricative + lateral	1	/ɛɪxl/ 1	[ɛɪxl] 1
Vowel + fricative + schwa + lateral	1	/ɛɪxəl/ 1	[ɛɪxəl] 1
Vowel + lateral + schwa + stop	4	/ɛɪlək/ 4	[ɛɪlək] 3, [ɛɪləg] 1
Monosyllabic	**57**		
Vowel + fricative	16	/ɛɪx/ 13	[ɛɪx] 11, [ɛɪɣ] 2
		/ɛx/ 3	[ɛx] 2, [ɛɣ] 1
Vowel + stop	22	/ɛɪk/ 22	[ɛɪk] 13, [ɛɪg] 9
Vowel + fricative + stop	18	/ɛɪxk/ 16	[ɛɪxk] 15, [ɛɪɣk] 1
		/ɛxk/ 2	[ɛxk] 2
Vowel + fricative + lateral	1	/ɛɪxl/ 1	[ɛɪxl] 1
Total	**159**		

Figure 6: Tokens of *eigenlijk* in Ernestus and Smith (2018, p.140)

B Anti-plagiarism Declaration

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst habe, dass ich keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe, dass ich alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet habe, dass die Arbeit weder vollständig noch in wesentlichen Teilen Gegenstand eines anderen Prüfungsverfahrens gewesen ist, dass ich die Arbeit weder vollständig noch in wesentlichen Teilen bereits veröffentlicht habe, dass das in Dateiform eingereichte Exemplar mit dem eingereichten gebundenen Exemplar übereinstimmt.

Tübingen, den 31.03.2023

Laura Benthaus