Articulatory planning is continuous and sensitive to informational redundancy

Mark Pluymaekers[a] Mirjam Ernestus[b,a] R. Harald Baayen[a,b]

[a]Radboud University Nijmegen, The Netherlands, and

[b]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Running title: Informational redundancy and reduction

Address correspondence to:

Mark Pluymaekers

Radboud University Nijmegen

P.O. Box 310

6500 AH Nijmegen, The Netherlands

Telephone: (+31)243612160

Fax: (+31)243521213

E-mail: Mark.Pluymaekers@mpi.nl

Abstract

This study investigates the relationship between word repetition, predictability from neighbouring words, and articulatory reduction in Dutch. The focus was on the seven most frequent words ending in the adjectival suffix **-lijk**. For each of these words, 40 occurrences were randomly selected from a large database of face-to-face conversations. The selected tokens were manually transcribed and the durations of their stems and suffixes were measured. The results show that the degree of articulatory reduction (as measured by duration and number of realized segments) was indeed affected by repetition, predictability from the previous word, and predictability from the following word. Interestingly, not all of these effects were significant across morphemes and target words. The implications of these findings for models of speech production are discussed.

Introduction

We speak in order to be understood. Nevertheless, the cognitive demands of conversational interaction may force speakers to reduce articulatory effort on certain words, leading to a temporary decrease in intelligibility. Although reductions occur frequently in spontaneous speech and can be quite extreme [Ernestus, 2000; Johnson, 2004], there is little evidence that their presence actually hinders communication. This has been explained by hypothesizing that speakers only reduce articulatory effort on words that are predictable for the listener, either from the linguistic context or from the situation in which the interlocutors find themselves [e.g. Lindblom, 1990; Jurafsky et al., 2001]. With regard to linguistic context, two factors have received much attention in the literature: **word repetition** and the predictability of a word from its neighbouring words (henceforth, **contextual predictability**). These two factors have in common that they are both concerned with the **informational redundancy** of a word in its context. In this section, we review the relevant literature for both variables.

Repetition

Effects of word repetition on reduction were first reported by Fowler and Housum [1987]. They found that second mentions of words in monologues were shorter and less intelligible in isolation than first mentions. Bard et al. [2000] replicated this effect for dialogues, showing that it was present irrespective of whether the speaker or the listener had uttered the first token of the word. No repetition effects were found when subjects read words in lists [Fowler, 1988], or when two tokens in a monologue were divided by a major episode boundary [Fowler et al., 1997].

These last two findings seem to justify Hawkins and Warren's [1994] conclusion that it is not so much repetition that matters, but rather whether a word presents 'given' or 'new' information. In their view, durational differences between first and second mentions are due to de-accentuation of the second token, which usually refers to given information. This hypothesis works well for the studies discussed above, which were all concerned with first and second mentions of words describing concrete entities or actions.

Recent findings by Gregory et al. [1999] and Aylett and Turk [2004] challenge Hawkins and Warren's account, however. Both studies report effects of the **number** of previous mentions of a word on its duration. This shows that durational differences cannot only be observed between first and second mentions, but also between, for instance, fifth and tenth mentions. Since theories of givenness [e.g. Prince, 1981] generally assume that one mention of a word is sufficient for its referent to reach the highest degree of givenness, these reductions cannot be ascribed to de-accentuation alone. In other words, there seems to be more to repetition effects than just the mere difference between old and new information.

Contextual predictability

Ever since Lieberman [1963], the relationship between contextual predictability and acoustic realizations has captivated researchers in phonetics, linguistics, and psycholinguistics alike. To determine the predictability of their target words, authors have used Cloze tests [e.g. Hunnicutt, 1985] or, as the availability of large speech corpora increased, co-occurrence statistics based on frequency. These statistics can be computed for a wide variety of linguistic units, from syllables [e.g. Aylett and Turk, 2004] to complete syntactic structures [e.g. Gahl and Garnsey, 2004]. Most studies, however, focus on words [e.g. Gregory et al., 1999; Fosler-Lussier and Morgan, 1999; Jurafsky et al., 2001; Bush, 2001; Bell et al., 2003]. Two well-known measures of contextual probability are 'conditional probability' and 'mutual information', both of which capture the likelihood of a certain word occurring given one or more of its neighbouring words.

Previous studies on contextual predictability and reduction have produced results that are both consistent and inconsistent. They are consistent in that all significant effects go in the same direction: Words that are more likely to occur are more reduced. They are inconsistent, however, with regard to the relevance of the different measures. Some words are completely unaffected by predictability, while others show effects of two or three probabilistic measures at the same time [Fosler-Lussier and Morgan, 1999; Bell et al., 2003]. Furthermore, the results of the various studies are difficult to compare, since all studies used slightly different sets of dependent and independent variables. These methodological differences have directed attention away from other important issues, such as the cognitive and articulatory processes underlying the effects.

Our approach

It is clear that for both repetition and contextual predictability, several issues remain to be addressed. In this study, we focus on two questions. First, are effects of repetition solely due to the distinction between 'given' and 'new' information and second, what do effects of repetition and contextual predictability reveal about speech production processes? Like most of the previous studies, we use corpus data to investigate these issues. What is new in our approach, is that we focus on words that are morphologically complex. By studying words that have internal structure, we hope to learn more about the effects of repetition and predictability on different parts of the word.

We concentrate on the seven most frequent words ending in the Dutch suffix **-lijk**. These words, which are listed in Table 1, are suitable targets for several reasons. First of all, they are hard to analyze in terms of 'given' versus 'new' information. This is especially true for 'eigenlijk', 'natuurlijk', and 'namelijk', which mainly serve as discourse markers. 'Duidelijk', 'waarschijnlijk', 'makkelijk', and 'moeilijk' can also function as predicates (and hence present new information), but because they are seldom the topic of subsequent utterances, they cannot be treated as 'given' in the sense intended by Hawkins and Warren [1994]. Furthermore, words ending in **-lijk** can be extremely reduced [Ernestus, 2000], and these reductions are at least partly predictable from probabilistic measures such as word frequency [Pluymaekers et al., submitted] and mutual information [Keune et al., in press]. By studying the durational properties of a number of highly frequent **-lijk** words, we hope to obtain answers to the questions formulated above.

INSERT TABLE 1 APPROXIMATELY HERE

Materials and method

The materials were taken from the subcorpus 'Spontaneous speech' of the Corpus of Spoken Dutch [Oostdijk, 2000]. This subcorpus contains 225 hours of face-to-face conversations, all of which have been orthographically transcribed. We restricted ourselves to speakers from the Netherlands, since they have been shown to use reduced forms more than speakers from Flanders [Keune et al., in press]. For each of the words in Table 1, a randomized list was made of all occurrences in the subcorpus that were not surrounded by pauses or disfluencies. From this randomized list, the first 40 tokens were selected for further analysis. If the recording quality of a selected token was too poor for acoustic measurements, it was replaced with the next token on the list. In total, 280 tokens were analyzed.

The dependent variables in this study were the durations of the stem and the suffix and the number of realized segments in these two morphemes. Durations were measured by hand using the software package PRAAT [Boersma, 2001]. Boundaries were placed between the previous word and the stem, between the stem and the suffix, and between the suffix and the following word. If a segment was ambiguous as to whether it belonged to the stem or the suffix (like the [ə] in the realization [naməkl] for 'namelijk' /namələk/), it was considered part of the suffix. Since most recordings contained at least some background noise, it was hard to establish a clear-cut segmentation strategy [see also Vorstermans et al., 1996]. Figure 1 shows the manual segmentations for two tokens of the word 'duidelijk', including parts of the previous and the following word. The top token was relatively easy to segment, since there was hardly any background noise or overlapping speech. The bottom token was much harder, mainly due to the presence of overlapping speech. In all cases, we placed boundaries where we could see visible changes in the waveform pattern supported by abrupt formant transitions in the spectrogram. To obtain information about the number of realized segments, broad phonetic transcriptions were made of the stem and the suffix. If there was neither visible nor audible acoustic evidence for the presence of a particular segment, that segment was considered absent.

INSERT FIGURE 1 APPROXIMATELY HERE

To assess the effects of repetition, we counted how often the target word had been uttered during the conversation before the selected token occurred. Given the results of Bard et al. [2000], we did not distinguish between tokens uttered by the same speaker and tokens uttered by other speakers. To reduce the effects of extreme counts, all values were logarithmically transformed.

As mentioned earlier, contextual predictability can be established in various ways. To avoid the problems associated with testing several probabilistic measures at the same time, we focused on just two variables: Mutual information with the previous word and mutual information with the following word. To compute mutual information, we used the following equation (X and Y denote either the previous word and the target word or the target word and the following word):

$$MI(X;Y) = log \frac{log\,(Frequency(XY))}{log\,(Frequency(X)) * log\,(Frequency(Y))}$$

The frequency estimates were logarithmically transformed before entering the equation. This was done to minimize the effects of very high frequencies on the outcome of the computation. All frequency estimates were obtained from the Corpus of Spoken Dutch.

Other variables known to affect reduction were taken into account as well. First of all, we controlled for the speaker characteristics Sex, Year of Birth, Education Level, and Region of Secondary Education. We also computed the Speech Rate (in syllables per second) in the largest chunk of speech containing the target word that did not contain an audible pause. The presence of Pitch Accent on the stem of the target word was monitored, and for each token we coded whether the following word started with a consonant or a vowel (henceforth, Following Segment). Finally, we included the factor Word as a control variable in the analyses. This enabled us to control for differences between the target words with respect to phonemic content, frequency of occurrence, and so on.

Results

Analysis

To investigate the effects of number of previous mentions and mutual information while controlling for other relevant variables, we used multiple regression analysis [e.g. Chatterjee et al., 2000]. In total, six regression models were fitted: three for the durations of the stem, the suffix, and the word as a whole, and three for the number of realized segments in the stem, the suffix, and the word as a whole. To find the best model in each case, we used a strict and straight-forward model selection procedure. First, we entered the control factors into the model, retaining only those variables that showed a significant effect. Then, the number of previous mentions (Mentions) was added, followed by mutual information with the previous word (MI Previous) and mutual information with the following word (MI Following). If any of these variables failed to show a significant effect, it was dropped from the equation. The resulting model was checked for interactions between the predictor variables, which were retained if they added to the predictive power of the model. Subsequently, diagnostic plots were used to identify data points that were outliers with regard to leverage or Cook's distance values. These outliers (usually three or four data points) were removed and the model was re-fitted to the remaining data. If a factor was no longer significant after the removal of outliers, it was dropped and the last two steps of the procedure were repeated. Finally, a bootstrap validation was performed to check for overfitting. Only those predictor variables that remained significant throughout this whole procedure are reported below.

Regression results

INSERT TABLE 2 APPROXIMATELY HERE

The results of the six regression models are summarized in Table 2. It shows for each model which of the predictor variables were significant. The beta coefficients indicating the direction and size of the effects are given in the main text below, as are the corresponding p-values. The factor Word was significant in all analyses, reflecting differences between the target words with respect to phonemic content and, possibly, word frequency. Since such differences are not the main interest of this study, these effects are not further addressed here. First, we discuss the results for the duration of the stem, followed by the results for the number of realized segments in the stem. These two steps are then repeated for the suffix and the word as a whole.

The stem was longer if it carried Pitch Accent ($\hat{\beta} = 51.0, t(260) = 4.61, p < 0.0001$) and shorter at higher Speech Rates ($\hat{\beta} = -15.3, t(260) = -6.30, p < 0.0001$). There was also an interaction between MI Previous and Word ($F(7, 260) = 6.63, p < 0.0001$), which is shown in Figure 2. MI Previous was significant for two of the seven target words: 'natuurlijk' ($\hat{\beta} = -108.3, t(260) = -5.31, p < 0.0001$) and 'eigenlijk' ($\hat{\beta} = -86.2, t(260) = -3.70, p < 0.0005$). In both cases, a higher value for MI Previous correlated with shorter realizations of the stem.

INSERT FIGURE 2 APPROXIMATELY HERE

A similar interaction was observed for the number of realized segments in the stem ($F(7, 262) = 4.32, p < 0.0005$). Again, the shortening effect of MI Previous was limited to the words 'natuurlijk' ($\hat{\beta} = -1.3, t(262) = -4.66, p < 0.0001$) and 'eigenlijk' ($\hat{\beta} = -0.8, t(262) = -2.50, p < 0.05$). There was also a main effect of MI Following ($\hat{\beta} = -0.2, t(262) = -2.14, p < 0.05$), indicating that words that were more predictable from their following words were realized with fewer segments in the stem.

Ten of the 280 tokens in the data set contained no visible trace of a suffix and were therefore excluded from the analyses for the suffix. The duration of the suffix was predicted by Sex ($\hat{\beta} = 15.3, t(255) = 2.92, p < 0.005$), Speech Rate ($\hat{\beta} = -15.7, t(255) = -7.39, p < 0.0001$) and Following Segment ($\hat{\beta} = 13.3, t(255) = 2.53, p < 0.05$). Suffixes were longer if they were produced by women, longer if they were followed by a vowel, and shorter at higher Speech Rates. A higher value for MI Following also correlated with shorter realizations ($\hat{\beta} = -11.9, t(255) = -2.19, p < 0.05$). Finally, we found an effect of Mentions ($\hat{\beta} = -9.0, t(255) = -2.07, p < 0.05$): The more the target word had been mentioned in the preceding discourse, the shorter the suffix. The number of realized segments in the suffix was only predicted by the factor Word.

The duration of the word as a whole was predicted by Year of Birth ($\hat{\beta} = -0.9, t(259) = -3.03, p < 0.005$), Speech Rate ($\hat{\beta} = -30.7, t(259) = -8.28, p < 0.0001$), and the presence of Pitch Accent on the stem ($\hat{\beta} = 59.7, t(259) = 3.66, p < 0.0005$). Older speakers produced longer words, words were shorter at higher Speech Rates, and an accented stem led to longer realizations of the word. Again, there was a significant interaction between MI Previous and Word ($F(7, 259) = 6.15, p < 0.0001$), which was very similar to the two interactions mentioned above for the stem. The main difference was that apart from 'natuurlijk' ($\hat{\beta} = -147.3, t(259) = -4.90, p < 0.0001$) and 'eigenlijk' ($\hat{\beta} = -126.1, t(259) = -3.67, p < 0.0005$), 'namelijk' was also significantly shorter if the mutual information with the previous word was higher ($\hat{\beta} = -51.9, t(259) = -2.04, p < 0.05$). This interaction is shown in Figure 3.

INSERT FIGURE 3 APPROXIMATELY HERE

Words were produced with more segments if the stem carried Pitch Accent ($\hat{\beta} = 0.6, t(260) = 2.25, p < 0.05$). Furthermore, words with high MI Following values contained fewer segments ($\hat{\beta} = -0.3, t(260) = -2.31, p < 0.05$). The interaction between MI Previous and Word was once more significant ($F(7, 260) = 5.47, p < 0.0001$), and again the effect was limited to 'natuurlijk' ($\hat{\beta} = -2.2, t(260) = -4.87, p < 0.0001$) and 'eigenlijk' ($\hat{\beta} = -1.6, t(260) = -3.01, p < 0.005$).

Summary and discussion

In this study, we have shown that the acoustic realizations of frequent words ending in the Dutch suffix **-lijk** are affected by word repetition, predictability from the previous word, and predictability from the following word. This section outlines the most important findings and discusses their implications for models of speech production. In addition, we point to directions for future research.

The role of repetition was restricted to a significant effect of the number of previous mentions of the word on the duration of the suffix. It should be noted, though, that this variable approached significance for the durations of the stem and the entire word as well (p-values of .09 and .08, respectively). It is easy to see why the suffix should be affected most by repetition: It is by far the least informative part of the word. What is perhaps more surprising, is the fact that even a crude measure like number of previous mentions, which largely ignores syntactic, prosodic, and discourse structure, successfully predicts articulatory durations in spontaneous speech [see also Gregory et al., 1999; Aylett and Turk, 2004].

Furthermore, our results indicate that previously reported effects of repetition on duration cannot solely be ascribed to de-accentuation of words referring to 'given' information [cf. Hawkins and Warren, 1994]. This is true for the following reasons. First, we controlled for the presence of Pitch Accent on the target words. Second, we focused on words that were hard, if not impossible, to analyze in terms of 'given' versus 'new'. Discourse markers such as 'eigenlijk', 'natuurlijk', and 'namelijk' cannot be classified as given or new, while words like 'moeilijk', 'makkelijk', 'waarschijnlijk', and 'duidelijk' can present new information, but are hardly ever used to refer to given entities. This choice of materials made the present study crucially different from earlier studies investigating repetition effects, including those which reported effects of the number of previous mentions [Gregory et al., 1999; Aylett and Turk, 2004]. Finally, we selected random tokens of the words, regardless of their informational contribution to the preceding discourse.

The fact that we still found an effect of the number of previous mentions on the duration of the suffix provides firm evidence that there is more to repetition effects than just the difference between given and new information.

How can this result be accounted for, then? A possible explanation is offered by Pickering and Garrod [2004], who propose a model of dialogue in which the semantic, syntactic, and phonetic representations of interlocutors become aligned with each other by means of a priming mechanism. As a concomitant result of this priming, the activation of a word at all representational levels increases with each occurrence of that word. This allows speakers to save articulatory effort on words that have been used repeatedly during a conversation, as listeners (whose representations for those words are equally highly activated) require less phonetic evidence to identify them correctly.

In addition to the effect of number of previous mentions, we found several effects of contextual predictability. In this respect, our study adds to the available evidence for the relationship between probability of occurrence and articulatory reduction [e.g. Gregory et al., 1999; Fosler-Lussier and Morgan, 1999; Jurafsky et al., 2001; Bush, 2001; Bell et al., 2003]. This is not our only contribution, however. Because we focused on morphologically complex words, we were able to obtain information about the effects of contextual predictability on different parts of our target words. More specifically, our materials allowed us to check whether there were differences between the previous and the following context with respect to the range and the strength of their effects. The picture that emerges from our results is that effects of contextual predictability operate in a way that is not all that simple and straight-forward.

Consider the stem, for example. Its duration was only affected by mutual information with the previous word, and this effect was limited to just two of the seven target words: 'natuurlijk' and 'eigenlijk'. Similar interactions were observed for the duration of the word as a whole and the number of segments in the stem and the word. By themselves, these findings are not too difficult to explain. 'Natuurlijk' and 'eigenlijk' have far higher frequencies than the other words and, being discourse markers, their semantic contribution to an utterance is relatively small. This makes them highly suitable targets for reduction, especially when their contextual predictability is also high. Further support for this claim comes from Fosler-Lussier and Morgan's [1999] study, in which effects of predictability were also limited to high-frequency words.

This raises the question, why the effects of mutual information with the **following** word were not restricted to only the most frequent target words. All effects observed for this variable were main effects, unmediated by (the characteristics of) the particular target word. Furthermore, unlike effects of previous context, effects of following context operated on both the stem and the suffix. What do these observations tell us about the cognitive processes underlying predictability effects?

First of all, our results cannot be accounted for by simply postulating ready-made motor programs spanning two or more words [e.g. Bybee, 2001; Bush, 2001]. Although the 'chunking' of frequently occurring word combinations into multi-word units is cognitively very plausible, such an account fails to explain why effects of previous context were limited to high-frequency words while effects of following context were not. A similar problem was encountered by Gahl and Garnsey [2004], who found correlations between the probability of occurrence of a certain syntactic structure and the durations of words within that structure, regardless of the particular words used. Since it is very unlikely that all different word combinations used in their study were stored as units in the speaker's lexicon, there must be some other explanation for their (and our) findings.

One possibility is that articulation proceeds on a unit-by-unit basis, allowing articulatory effort to be adjusted for each unit on the basis of the informational redundancy of the unit itself (e.g. stem vs. suffix), the word it belongs to (predictable vs. unpredictable), and the syntactic structure it is part of (probable vs. improbable continuations). In fact, most theories of speech production assume that there is a single basic unit of articulation. There has been some debate, however, about which unit is most appropriate for this role.

Given our results, the morpheme does not seem a very suitable candidate, as some of our effects operated across morpheme boundaries while others did not. The syllable, which has been proposed by many researchers [e.g. Levelt and Wheeldon, 1994; Cholin, 2004], faces a similar problem: Some effects were limited to specific syllables, while others affected two or more 'units' at the same time. Furthermore, Pluymaekers et al. [submitted] have shown that the individual segments in a syllable are all subject to their own specific forces, further challenging the assumed unitary status of the syllable. Segments, on the other hand, have the disadvantage that their corresponding speech gestures often overlap considerably in time. These considerations suggest that the main problem may not lie in our inability to identify the basic unit of articulation, but rather in the assumption that there is one such unit.

As an alternative, we propose a model in which articulatory planning is continuous and not unit-based. To ensure a relatively constant information density, articulatory effort is adjusted throughout the production of the utterance. Parts of the speech stream that carry little information are realized with less articulatory effort than more informative parts. Informativeness is determined on the basis of different dimensions simultaneously: The frequency of the word, the predictability from neighbouring words, the number of times the word has been mentioned, the probability of the syntactic structure it occurs in, etcetera. Sometimes these dimensions of informational redundancy interact, while in other cases they exert their influence separately and additively. More research is needed to examine the circumstances and ways in which the different informational measures can interact.

Our results also suggest that there is an asymmetry between predictability effects that arise from planning processes prior to the uttering of a word and predictability effects linked to the preparation of the following context. If speakers are planning the articulation of a word (let us call it 'target') that is both highly predictable from the previous context and semantically rather meaningless, they may choose to pronounce it in a highly reduced way. In the meantime, however, the words following the target also need to be planned. During this planning, both the words preceding the target and the target are taken into account, and it is not inconceivable that the target, by virtue of being involved in subsequent planning, is again subject to articulatory reduction (if the mutual information between the target and the following context is high). These two temporally cascaded planning processes may lead to different degrees of reduction, with the more robust reduction apparently coming from the articulatory planning process in which the target itself is also involved. This "involvement-in-planning" account could explain the differences we observed between effects of previous and following context, although at present it is of course highly tentative and in need of further investigation.

Apart from the points already raised in this discussion, we feel a number of issues need to be addressed in future research. The first issue is the relationship between the activation level of a word and its acoustic realization. There are several indications that an increase in activation leads to acoustic reduction, but little is known about the exact details of this relationship. The second point concerns the balance between speaker-internal and listener-motivated processes in explaining reductions. It is very well possible that some reduction processes are purely due to cognitive processes on behalf of the speaker, while others occur (partly) because the speaker actively takes the listener's knowledge and needs into account. We are convinced that by tackling these issues, speech researchers can finally come to understand the roles of speaker, listener, and context in explaining the enormous phonetic variation inherent in conversational speech.

Acknowledgment

This research was supported by Netherlands Organization for Scientific Research (NWO) grant number 360-70-130 to the third author.

References

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

Bard, E.; Anderson, A.; Sotillo, C.; Aylett, M.; Doherty-Sneddon, G.; Newlands, A.: Controlling the intelligibility of referring expressions in dialogue. J. Memory Lang. *42:* 1–22 (2000).

Bell, A.; Jurafsky, D.; Fosler-Lussier, E.; Girand, C.; Gregory, M.; Gildea, D.: Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. J. acoust. Soc. Am. *113:* 1001–1024 (2003)

Boersma, P.: Praat, a system for doing phonetics by computer. Glot International *5:* 341–355 (2001).

Bush, N.: Frequency effects and word-boundary palatalization in English; in Bybee, Hopper, Frequency and the emergence of linguistic structure, pp. 255–280 (Benjamins, Amsterdam 2001).

Bybee, J. L.: Phonology and language use (Cambridge University Press, Cambridge 2001).

Chatterjee, S.; Hadi, A.; Price, B.: Regression analysis by example (John Wiley & Sons, New York 2000).

Cholin, J.: Syllables in speech production: effects of syllable preparation and syllable frequency; PhD thesis University of Nijmegen (2004).

Ernestus, M.: Voice assimilation and segment reduction in casual Dutch: a corpus-based study of the phonology-phonetics interface (LOT, Utrecht 2000).

Fosler-Lussier, E.; Morgan, N.: Effects of speaking rate and word frequency on pronunciations in conversational speech. Speech Commun. *29:* 137–158 (1999).

Fowler, C.: Differential shortening of repeated content words produced in various communicative contexts. Lang. Speech *31:* 307–317 (1988).

Fowler, C.; Housum, J.: Talkers' signalling of "new" and "old" words in speech and listeners' perception and use of the distinction. J. Memory Lang. *26:* 489–504 (1987).

Fowler, C.; Levy, E.; Brown, J.: Reductions of spoken words in certain discourse contexts. J. Memory Lang. *37:* 24–40 (1997).

Gahl, S. and Garnsey, S.: Knowledge of grammar, knowledge of usage: syntactic probabilities affect pronunciation variation. Language *80:* 748–774 (2004).

Gregory, M.; Raymond, W.; Bell, A.; Fosler-Lussier, E.; Jurafsky, D.: The effects of collocational strength and contextual predictability in lexical production. Chicago Ling. Soc. *35:* 151–166 (1999).

Hawkins, S.; Warren, P.: Phonetic influences on the intelligibility of conversational speech. J. Phonet. *22:* 493–511 (1994).

Hunnicutt, S.: Intelligibility versus redundancy - conditions of dependency. Lang. Speech *28:* 47–56 (1985).

Johnson, K.: Massive reduction in conversational American English. Spontaneous speech: data and analysis. Proc. of the 1st session of the 10th international symposium, The National International Institute for Japanese Language, Tokyo 2004.

Jurafsky, D.; Bell, A.; Gregory, M.; Raymond, W.: Probabilistic relations between words: evidence from reduction in lexical production; in Bybee, Hopper, Frequency and the emergence of linguistic structure, pp. 229–254 (John Benjamins, Amsterdam 2001).

Keune, K.; Ernestus, M.; Van Hout, R.; Baayen, R.H.: Social, geographical, and register variation in Dutch: from written 'mogelijk' to spoken 'mok'. To appear in Corpus Linguistics and Linguistic Theory.

Levelt, W. J. M.; Wheeldon, L.: Do speakers have access to a mental syllabary? Cognition *50:* 239–269 (1994).

Lieberman, P.: Some effects of semantic and grammatical context on the production and perception of speech. Lang. Speech *6:* 172–187 (1963).

Lindblom, B.: Explaining phonetic variation: a sketch of the H&H theory; in Hardcastle, Marchal, Speech production and speech modeling, pp. 403–440 (Kluwer, Dordrecht 1990).

Oostdijk, N.: The Spoken Dutch Corpus project. ELRA Newsletter *5:* 4–8 (2000).

Pickering, M.; Garrod, S.: Toward a mechanistic psychology of dialogue. Behav. Brain Sci. *27:* 169–226 (2004).

Pluymaekers, M.; Ernestus, M.; Baayen, R.H.: Lexical frequency and acoustic reduction in spoken Dutch. Under revision for J. acoust. Soc. Am.

Prince, E.: Toward a taxonomy of given-new information; in Cole, Radical pragmatics, pp. 225–256 (Academic Press, New York 1981).

Vorstermans, A.; Martens, J.; Van Coile, B.: Automatic segmentation and labelling of multi-lingual speech data. Speech Commun. *19:* 271–293 (1996).

[Table 1: English translations, frequencies per million in the Corpus of Spoken Dutch, and citation forms of the seven words investigated in this study.]

[Table 2: Summary of the regression results for the six models fitted in this study. A star (*) indicates that the variable in question was a significant predictor. The horizontal line in the middle separates the variables of interest (above) from the control variables (below). The bottom row shows the amount of variance explained (R^2) by each model.]

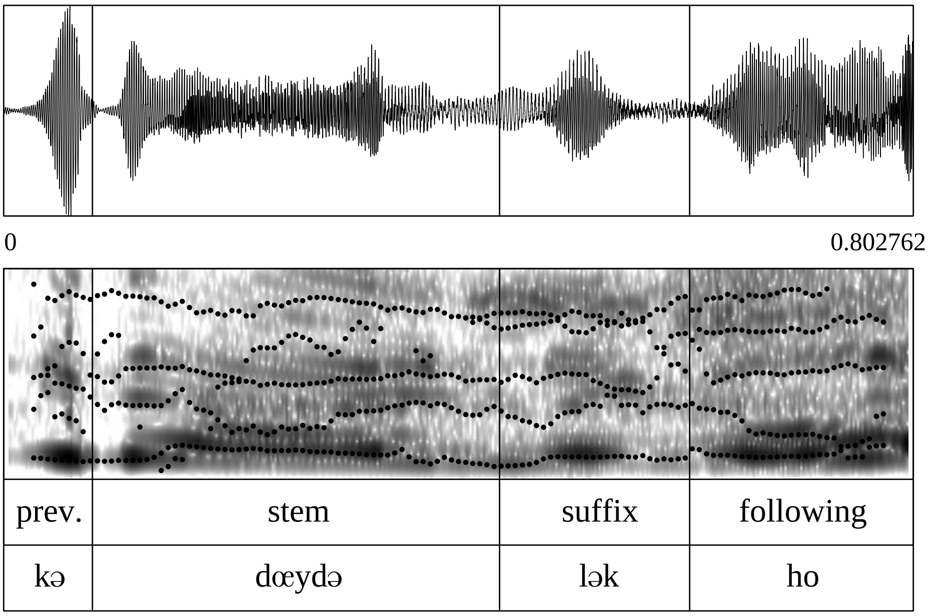
Word	Eng. translation	Frequency	Citation form
eigenlijk	actual/actually	1922	/ɛixələk/
natuurlijk	natural/naturally	1440	/natyrlək/
waarschijnlijk	probable/probably	335	/ʋarsxɛinlək/
moeilijk	difficult	320	/mujlək/
duidelijk	clear/clearly	272	/dœydələk/
namelijk	namely	135	/namələk/
makkelijk	easy/easily	96	/mɑkələk/

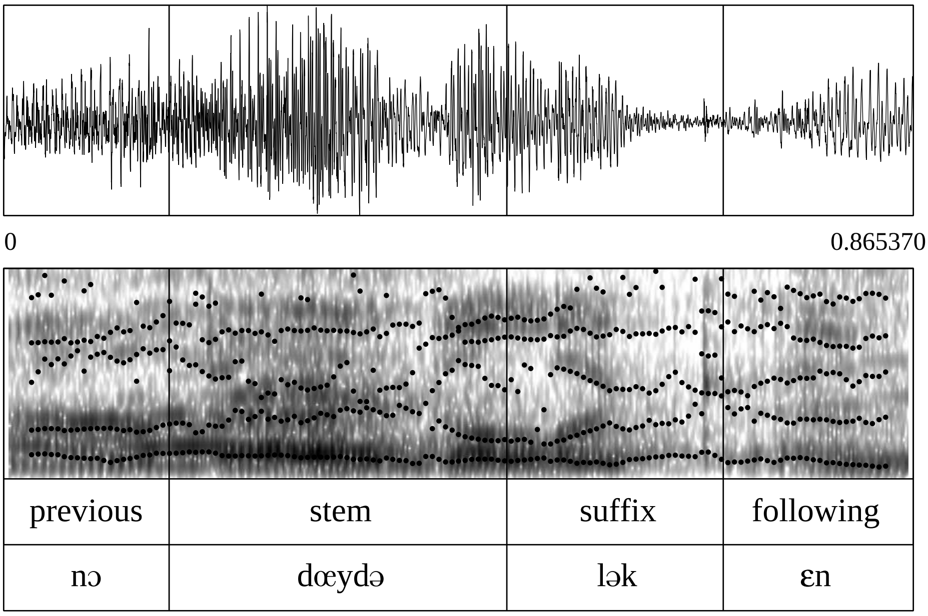
	Stem		Suffix		Entire word	
Predictor variable	Duration	Segments	Duration	Segments	Duration	Segments
Mentions			*			
MI Previous						
MI Previous * Word	*	*			*	*
MI Following		*	*			*
Following Segment			*			
Pitch Accent	*				*	*
Sex			*			
Speech Rate	*		*		*	
Word	*	*	*	*	*	*
Year of Birth					*	
Explained variance (R^2)	.58	.70	.38	.35	.51	.56

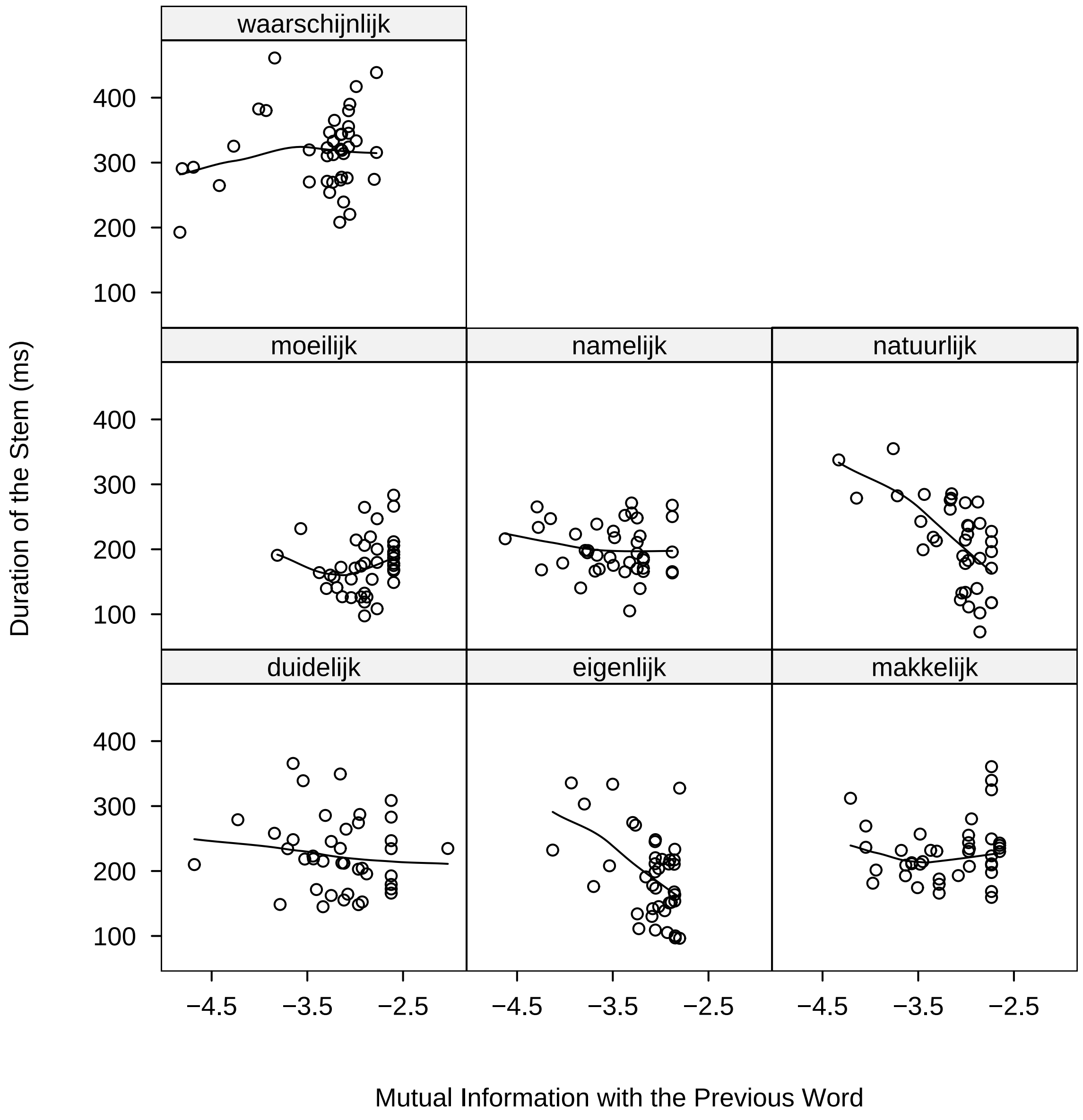
[Figure 1: Two segmentation examples of the word 'duidelijk'. The top token was produced without background noise or overlapping speech, resulting in a waveform that was relatively easy to segment. In the bottom token, it was much harder to determine segment boundaries. In both cases, we placed boundaries where we could see both visible changes in the waveform pattern supported by abrupt formant transitions in the spectrogram.]

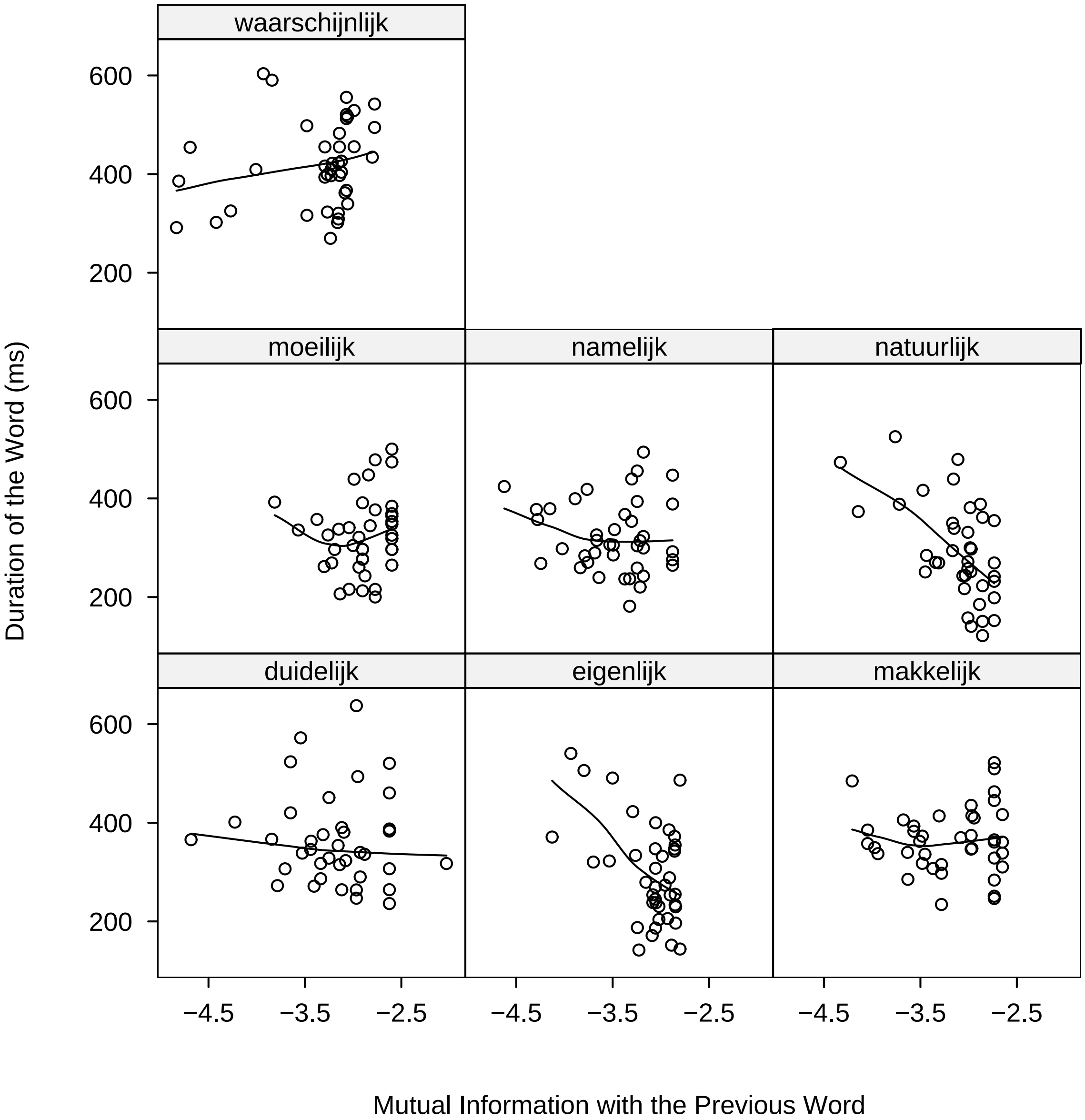
[Figure 2: Duration of the Stem plotted against Mutual Information with the Previous Word for each of the seven target words separately. As can be seen from the descending lines, a higher Mutual Information led to shorter realizations for 'natuurlijk' and 'eigenlijk'. The lines for 'namelijk' and 'duidelijk' also seem to fall somewhat, but these effects were not significant.]

[Figure 3: Duration of the Word plotted against Mutual Information with the Previous Word for each of the seven target words separately. As can be seen from the descending lines, a higher Mutual Information led to shorter realizations for 'natuurlijk', 'eigenlijk', and 'namelijk'. The line for 'duidelijk' also seems to fall somewhat, but this effect was not significant.]


0
0.802762
prev.
stem
suffix
following
kə
dœydə
lək
ho


0
0.865370
previous
stem
suffix
following
nɔ
dœydə
lək
ɛn


waarschijnlijk
moeilijk
namelijk
natuurlijk
duidelijk
eigenlijk
makkelijk
100
200
300
400
−4.5
−3.5
−2.5
Duration of the Stem (ms)
Mutual Information with the Previous Word


waarschijnlijk
moeilijk
namelijk
natuurlijk
duidelijk
eigenlijk
makkelijk
600
400
200
−4.5
−3.5
−2.5
Duration of the Word (ms)
Mutual Information with the Previous Word