Articulatory planning is continuous and sensitive to informational redundancy

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

[a]Radboud University Nijmegen, The Netherlands

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

Running title: Informational redundancy and reduction

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. For the seven most frequent words ending in the adjectival suffix **-lijk**, 40 occurrences were randomly selected from a large database of face-to-face conversations. Analysis of the selected tokens showed that the degree of articulatory reduction (as measured by duration and number of realized segments) was 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. Repetition effects were limited to suffixes, while effects of predictability from the previous word were restricted to the stems of two of the seven target words. Predictability from the following word affected the stems of all target words equally, but not all suffixes. The implications of these findings for models of speech production are discussed.

Introduction

We speak in order to be understood. Nevertheless, the dynamics 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; Kohler, 2000; Johnson, 2004], there is little evidence that their presence actually hinders communication. This has been explained by the hypothesis 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]. This suggests that it is not so much repetition that matters, but rather whether a word refers to 'given' or 'new' information.

Hawkins and Warren [1994] argued, however, that first and second occurrences of words differ not only in whether they present 'given' or 'new' information, but also in their likelihood of carrying sentence accent. First occurrences of content words are more likely to be accented than second occurrences, which could also explain the observed differences in duration and intelligibility. In their study, Hawkins and Warren tried to disentangle repetition effects from effects due to sentence accent and segmental identity. They found no differences in intelligibility between first and second tokens that could not be accounted for by the presence or absence of accent. This led them to conclude that “local phonetic variables, notably sentence accent and the phonetic and phonological properties of individual segments, exert a greater influence on intelligibility than whether or not a word has been used before in the conversation” [Hawkins and Warren 1994, p. 493].

Does this mean that repetition by itself should no longer be considered a possible predictor of reduction? Recent findings by Gregory et al. [1999] and Aylett and Turk [2004] suggest otherwise. 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 neither the fifth nor the tenth token of a word in a conversation are likely to be accented, these reductions are probably not due to deaccentuation alone. In other words, there seems to be more to repetition effects than just the presence or absence of sentence accent.

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. We will discuss these measures in more detail below.

Previous studies on the effects of contextual predictability on 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, is there an effect of repetition on reduction that is independent of sentence accent 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, 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., in press] and mutual information [Keune et al., in press]. Second, being adverbs and adjectives, **-lijk** words are a priori less likely to carry sentence accent. 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, presenting new information about the discourse topic. Therefore, the possibility that these words are accented cannot be completely excluded.

INSERT TABLE 1 APPROXIMATELY HERE

As is clear from Table 1, the seven target words investigated in this study differ in frequency, phonemic content, meaning, and the number and type of discourse functions they can perform. Since all of these factors can be expected to affect reduction, failure to control for them may limit the possibility of finding effects of repetition or contextual predictability. To over-come this problem, we incorporated the factor 'word' as a fixed effect in our analyses.

Of course, this does not control for the discourse function performed by a particular token of a word. There is no reason to assume, however, that discourse function is systematically correlated with either repetition or contextual predictability. It seems very unlikely that, for instance, the 15th occurrence of a word is always associated with one particular discourse function, while the 20th occurrence always performs another. We assume that for our target words, discourse functions are more or less randomly distributed among different tokens, independent of the number of times the word has been mentioned before or how predictable the token is on the basis of neighbouring words. Furthermore, classifying words according to their discourse function is a notoriously difficult activity, which can be regarded as a research topic in itself. Therefore, we considered this beyond the scope of the current paper.

Still, there were a lot of other variables that had to be controlled for. To this end, we used multiple regression analysis. In such an analysis, it is also easy to check whether the effects of two or more variables are additive or interactive. If, for example, repetition effects were to be limited to non-accented words, this would surface in our analyses as a significant interaction between repetition and accent. If, on the other hand, both repetition and accent show significant main effects, this implies that their effects are additive and not confounded. This kind of information is necessary for answering the two research questions formulated above.

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 often 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. All acoustic measurements were made by a trained phonetician with the help of 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 [namək] for ‘namelijk’ /namələk/), it was considered part of the suffix. In addition, the phonetician determined for each token which segments were realized in the speech signal. This transcription did not start from the citation forms of the target words, but was purely based on the auditory evidence in the signal and the visual information in the waveform. A particular segment was only included in the transcription if there was both visible and auditory evidence for its presence. Finally, the labeller coded for each token whether the stem carried pitch accent or not.

Since most recordings contained at least some background noise, it was hard to establish clear-cut segmentation criteria [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, the phonetician placed boundaries where she could see visible changes in the waveform pattern supported by abrupt formant transitions in the spectrogram.

INSERT FIGURE 1 APPROXIMATELY HERE

The fact that all tokens were measured only once may have implications for the generalizability of our results. After all, there is no guarantee that a second measurement of the same tokens, even if performed by the same labeller, would yield exactly the same results. On the other hand, it is unlikely that our results are completely due to the labeller's idiosyncracies, as she was naive with respect to the goals of the study and used similar criteria for all tokens.

To assess the effects of repetition, we determined for each randomly selected token how often the target word had been uttered during the conversation before the selected token occurred. We coded the selected item for the time point during the conversation at which it occurred (for example, after 54 seconds), and counted how often the same word had been uttered before that time point. 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. The original counts varied between 0 and 20, while the transformed values ranged between 0 and 3.

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 between the target word and the previous word and mutual information between the target word and the following word. The mutual information between two words is a measure of the reduction in uncertainty about one word due to knowing about the other [e.g. Manning and Schütze, 1999]. Therefore, the higher the value for mutual information, the easier one word can be predicted on the basis of the other. To compute mutual information, we used the following equation (X and Y denote either the previous word and the target word, or they denote the target word and the following word; XY denotes the combination of the two words):

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

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. Furthermore, language users are known to be sensitive to logarithmic values rather than raw frequencies [Rubenstein and Pollack, 1963]. All frequency estimates were obtained from the Corpus of Spoken Dutch.

The reliability and stability of the mutual information measure depends crucially on corpus size. If the corpus is too small, the frequency counts for many two-word combinations (the numerator in the equation above) will approach zero, leading to unstable estimates. This was not the case in our sample, as the frequencies of the sampled word combinations ranged from 1 to 1520. As a consequence, the distribution of the log transformed values entering the equation was reasonably symmetric.

Eight other variables known (or expected) to affect reduction were taken into account and designated control variables. First, there were the speaker characteristics Sex, Year of Birth, Education Level, and Region of Secondary Education. Second, Speech Rate (in syllables per second) was computed over the largest chunk of speech containing the target word that did not contain an audible pause. The number of syllables in the chunk was determined on the basis of two sources of information. For all words in the chunk except the target word, we counted the number of vowels in the orthographic transcription. For the target word itself, we counted the number of vowels in the manual segmentation. The total number of vowels was then divided by the overall duration of the chunk. The remaining three control variables were the presence of Pitch Accent on the stem, whether the segment following the target word was a consonant or a vowel (henceforth, Following Segment), and, for reasons explained above, Word. Table 2 gives an overview of the most important sample characteristics for each of the seven target words separately.

INSERT TABLE 2 APPROXIMATELY HERE

Results

Analysis

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 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. During bootstrapping, the proposed model was fitted 200 times to different random selections of our data points. If a particular variable in the model failed to reach significance in too many of these fitting cycles, it was removed from the model. Only those predictor variables that remained significant throughout this whole procedure are reported below.

Regression results

INSERT TABLE 3 APPROXIMATELY HERE

The results of the six regression models are summarized in Table 3. 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 meaning, 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 illustrated 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 or audible trace of the suffix **-lijk** and were therefore excluded from the analyses for the suffix. The duration of the suffix was predicted by Sex ($\hat{\beta} = 11.8, t(248) = 2.30, p < 0.05$), Speech Rate ($\hat{\beta} = -16.4, t(248) = -7.79, p < 0.0001$) and Following Segment ($\hat{\beta} = 14.1, t(248) = 2.72, p < 0.01$). Suffixes were longer if they were produced by women, longer if they were followed by a vowel, and shorter at higher Speech Rates. There was also a significant interaction between MI Following and Word ($F(7, 248) = 2.30, p < 0.05$). A higher value for MI Following led to shorter realizations of the suffix, but only for the target words 'eigenlijk' ($\hat{\beta} = -41.0, t(248) = -2.30, p < 0.05$) and 'namelijk' ($\hat{\beta} = -37.5, t(248) = -3.12, p < 0.005$). Finally, we found an effect of Mentions ($\hat{\beta} = -9.4, t(248) = -2.16, p < 0.05$): The more often 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

As expected, 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 durations and number of realized segments of the seven most 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 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). Apparently, 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 there is an effect of repetition on reduction that is independent of sentence accent. This is true for several reasons. First of all, we focused on words that are unlikely to be accented, either because they are discourse markers ('eigenlijk', 'natuurlijk', and 'namelijk'), or because they seldom introduce new entities to a discourse ('duidelijk', 'waarschijnlijk', 'moeilijk', and 'makkelijk'). More importantly, we found that even if these words are accented, they still show an effect of the number of previous mentions on the duration of the suffix.

At first glance, these results may appear contrary to the conclusions of Hawkins and Warren [1994] mentioned earlier in this paper. However, some reservations are in place here. First of all, Hawkins and Warren measured intelligibility, while we were concerned with durations and the number of realized segments. Moreover, the word type used in the current study also differed from the one used by Hawkins and Warren. It may be possible that pure repetition effects are restricted to adverbs and adjectives, while nouns or verbs, which were the focus of Hawkins and Warren's attention, show no such effects. Further research is needed here.

How can our findings 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 morphological 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 straightforward.

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.

For mutual information with the **following** word, the picture was somewhat more complicated. We again found a significant interaction with word: Mutual information with the following word only affected the duration of the suffix in the discourse markers ‘eigenlijk’ and ‘namelijk’. However, the effects observed for this variable on the number of segments in the stem and the word 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 always limited to high-frequency words, while effects of following context affected the stems of all words. Additional evidence against the chunking hypothesis was provided 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, words can be excluded as possible units, since stems and suffixes differed in their sensitivity to different measures. Morphemes, however, do not appear too suitable either, 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. [in press] 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 that 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, and so on. 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 this 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 possible that some reductions are mainly 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.

Acknowledgments

This research was supported by Netherlands Organization for Scientific Research (NWO) grant number 360-70-130 to the third author. We thank Randy Diehl, Sarah Hawkins, and Klaus Kohler for their useful comments on an earlier version of this paper.

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.

Kohler, K.: Investigating unscripted speech: implications for phonetics and phonology. Phonetica *57:* 85–94 (2000).

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).

Manning, C. and Schutze, H.: Foundations of Statistical Natural Language Processing (MIT Press, Cambridge 1999).

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. To appear in J. acoust. Soc. Am.

Rubenstein, H. and Pollack, I.: Word predictability and intelligibility. J. Verbal Learn. Verbal Behav. *2:* 147–158 (1963).

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 and in total) in the Corpus of Spoken Dutch , and citation forms of the seven words investigated in this study.]

[Table 2: Information about the sampled tokens for each of the target words separately (N = 40 for each target word). This information includes the number of tokens carrying pitch accent, the number of different previous and following words observed, the lower and upper values for both Mutual Information measures, and the maximum number of previous mentions observed.]

[Table 3: 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 (per million)	Frequency (in total)	Citation form
eigenlijk	actual/actually	1922	18320	/ɛixələk/
natuurlijk	natural/naturally	1440	13602	/natyrlək/
waarschijnlijk	probable/probably	335	3098	/ʋarsxɛinlək/
moeilijk	difficult	320	2736	/mujlək/
duidelijk	clear/clearly	272	2056	/dœydələk/
namelijk	namely	135	1154	/namələk/
makkelijk	easy/easily	96	872	/mɑkələk/

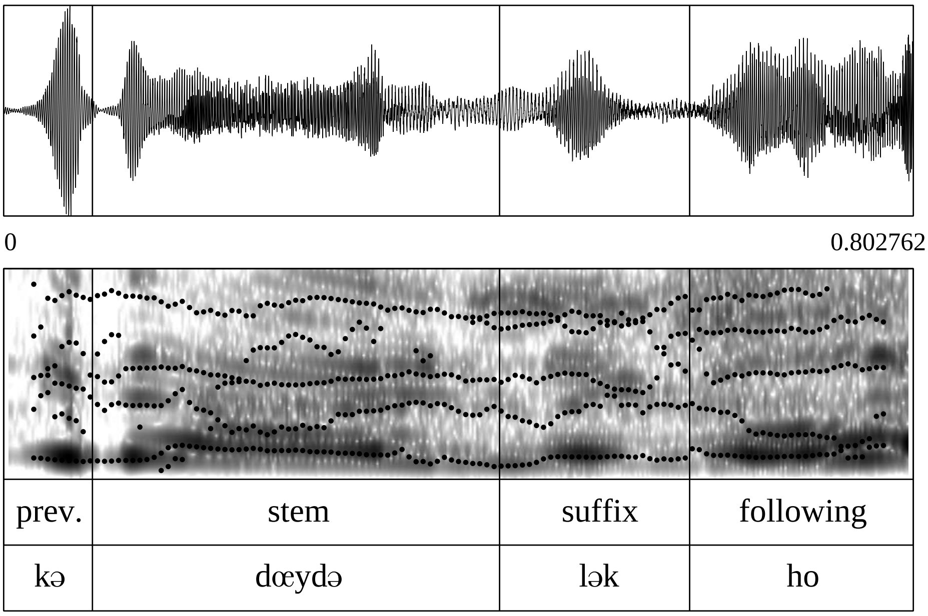
Word	Accented tokens	Previous words	Following words	MI Previous		MI Following		Max. prev. mentions
				lower	upper	lower	upper	
eigenlijk	3	30	25	-4.13	-2.80	-4.53	-2.75	20
natuurlijk	0	27	23	-4.51	-2.73	-4.54	-2.25	15
waarschijnlijk	2	28	26	-4.83	-2.78	-4.48	-2.82	4
moeilijk	4	19	24	-3.81	-2.60	-4.28	-2.68	6
duidelijk	4	25	29	-4.68	-2.03	-4.74	-2.84	2
namelijk	2	27	32	-4.62	-2.88	-4.58	-1.95	1
makkelijk	5	21	31	-4.21	-2.65	-4.54	-2.61	6

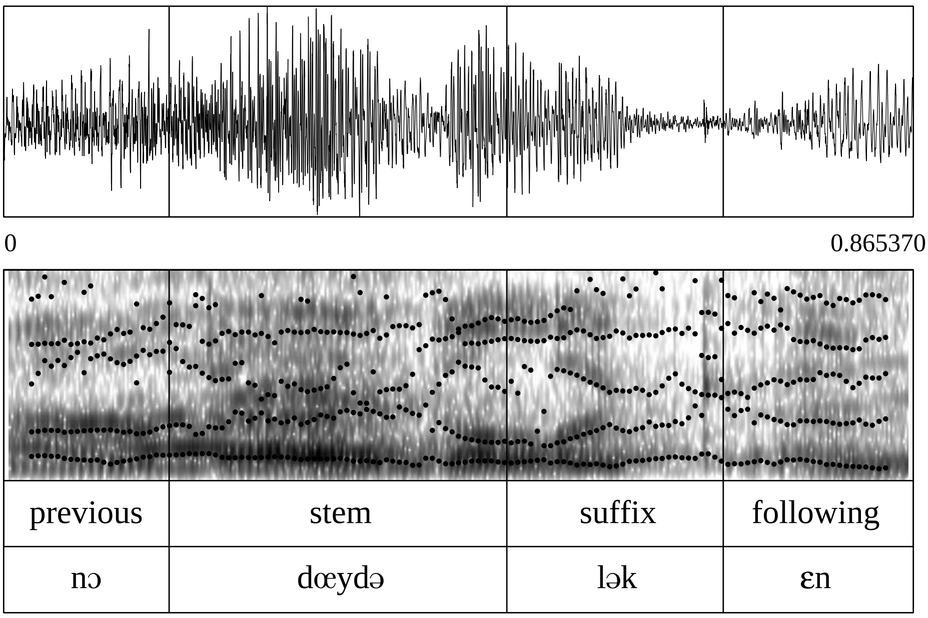
	Stem		Suffix		Entire word	
Predictor variable	Duration	Segments	Duration	Segments	Duration	Segments
Mentions			*			
MI Previous						
MI Previous * Word	*	*			*	*
MI Following		*				*
MI Following * Word			*			
Education Level						
Following Segment			*			
Pitch Accent	*				*	*
Region of Sec. Education						
Sex			*			
Speech Rate	*		*		*	
Word	*	*	*	*	*	*
Year of Birth					*	
Explained variance (R^2)	.58	.70	.41	.39	.51	.56

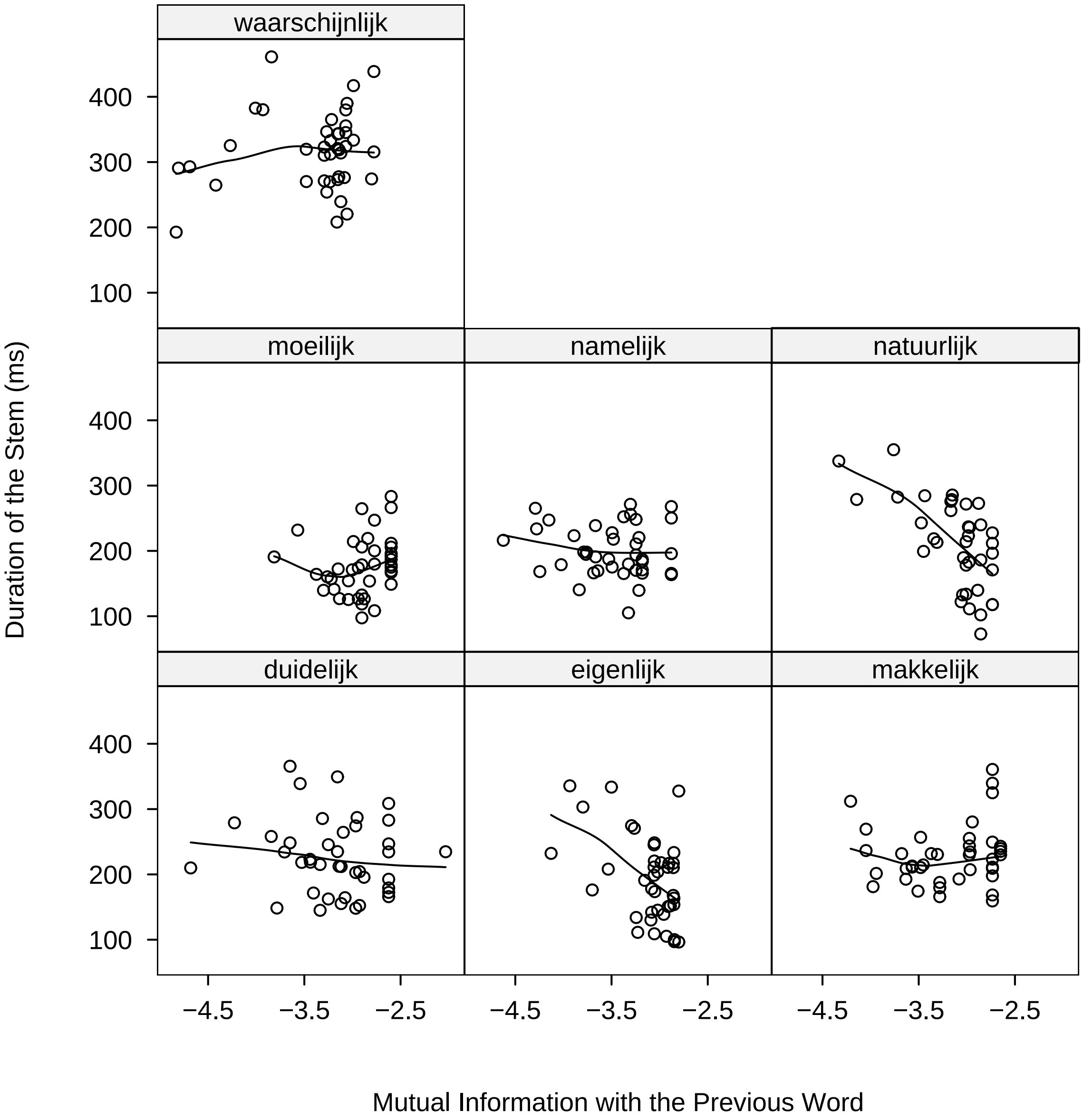
[Figure 1: Two segmentation examples of the word ‘duidelijk’. The top token was produced without background noise or overlapping speech, resulting in a waveform that was relatively easy to segment. In the bottom token, it was much harder to determine segment boundaries. In both cases, we placed boundaries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

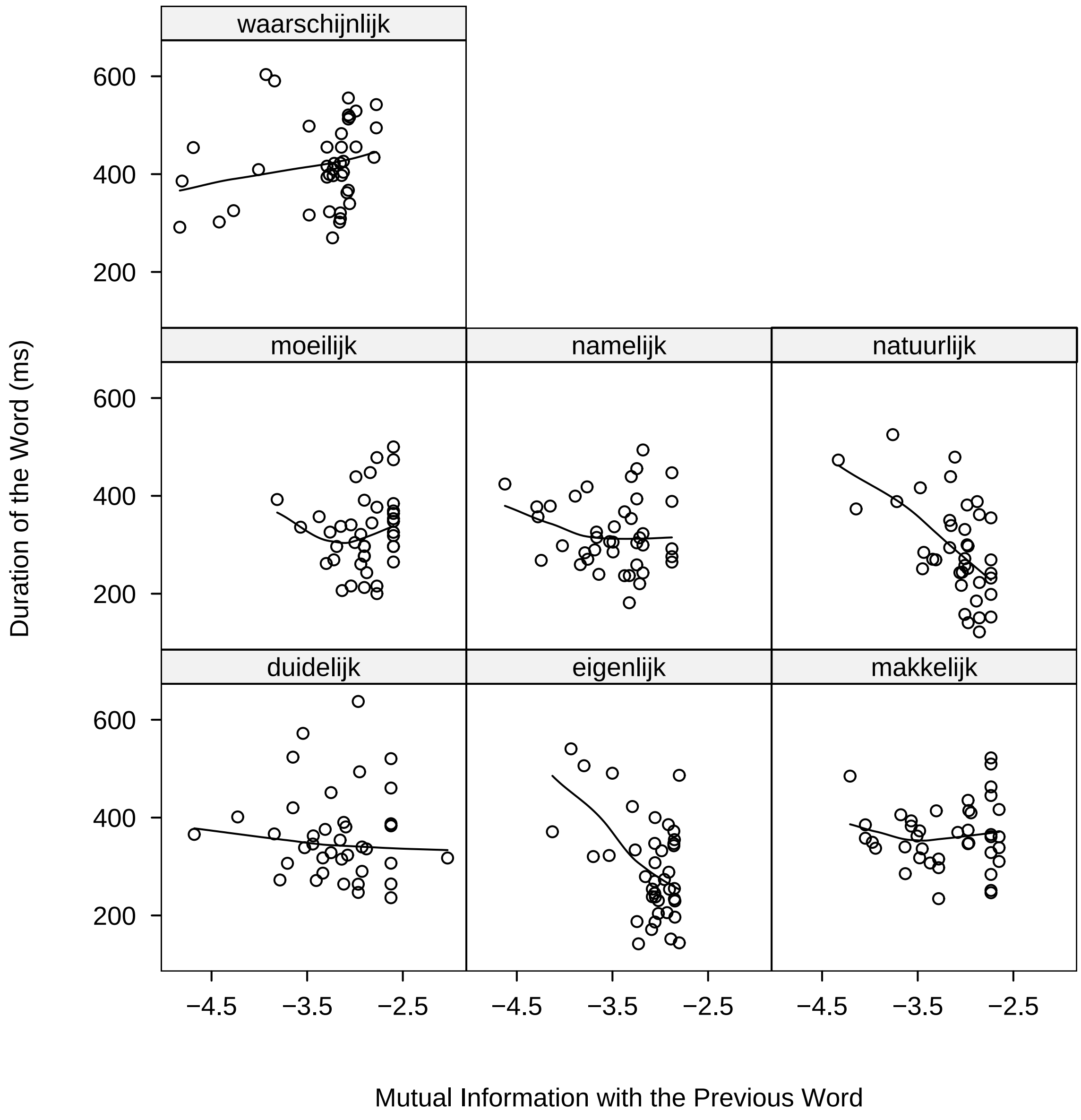
[Figure 2: Duration of the Stem plotted against Mutual Information with the Previous Word for each of the seven target words separately. As can be seen from the descending lines, a higher Mutual Information led to shorter realizations for ‘natuurlijk’ and ‘eigenlijk’. The lines for ‘namelijk’ and ‘duidelijk’ 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
Duration of the Stem (ms)
100
200
300
400
−4.5
−3.5
−2.5
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