THE COMPREHENSION OF ACOUSTICALLY REDUCED MORPHOLOGICALLY COMPLEX WORDS: THE ROLES OF DELETION, DURATION, AND FRQUENCY OF OCCURRENCE

Mirjam Ernestus and R. Harald Baayen

Radboud University Nijmegen & Max Planck Institute for Psycholinguistics
mirjam.ernestus@mpi.nl, harald.baayen@mpi.nl

ABSTRACT

This study addresses the roles of segment deletion, durational reduction, and frequency of use in the comprehension of morphologically complex words. We report two auditory lexical decision experiments with reduced and unreduced prefixed Dutch words. We found that segment deletions as such delayed comprehension. Simultaneously, however, longer durations of the different parts of the words appeared to increase lexical competition, either from the word's stem (Experiment 1) or from the word's morphological continuation forms (Experiment 2). Increased lexical competition slowed down especially the comprehension of low frequency words, which shows that speakers do not try to meet listeners' needs when they reduce especially high frequency words.

Keywords: Speech comprehension, acoustic reduction, frequency of occurrence, lexical competition, speech production

1. INTRODUCTION

In spontaneous conversations, words are often realized much shorter and with fewer segments than in formal speech [5, 9]. Several studies have shown that especially words of a higher frequency of occurrence tend to be acoustically reduced [e.g., 8, 10]. Two different hypotheses have been formulated for explaining this role of frequency. The first, speaker-driven, hypothesis is based on the fact that speakers have had more practice producing words of a higher frequency. More practice typically results in smoother, overlapping, articulatory gestures, which may lead to reduced realizations [4]. The second hypothesis is listener-driven. It argues that listeners have fewer difficulties recognizing reduced realizations of high frequency words than of low frequency words, and that as a consequence speakers can afford to reduce high frequency words to a greater extent without loss of comprehension [1].

Evaluation of the listener-driven hypothesis is difficult since little is known about the comprehension of acoustically reduced words. It has been shown that listeners rely on fine phonetic detail signaling the presence of highly reduced segments [11], and that the recognition of highly reduced words is difficult out of context [6]. Nothing is known about the role of a word's frequency of occurrence in the comprehension of acoustically reduced realizations.

The present study addresses the roles of segment deletion, durational reduction, and a word's frequency of occurrence in comprehension. We report two auditory lexical decision experiments with reduced and unreduced prefixed words in Dutch.

2. EXPERIMENT 1

2.1. Method

2.1.1. Participants

Forty-seven native speakers of Dutch were paid for their participation.

2.1.2. Materials

We constructed 36 pairs of words, each word consisting of a prefix (*be-* /bə/, *ge-* /xə/, *ont-* /ɔnt/, or *ver-* /vər/), a verbal stem, and a suffix ([t] for past-participles and [ə] for infinitives). An example is the pair *bestralen* - *bestraten*, [bəstralə] - [bəstratə], 'to irradiate' - 'to pave'. The members of a pair were phonologically as similar as possible, but differed substantially in their frequency of use, as listed in CELEX [2].

A female speaker produced two versions of each word. The first version was a careful pronunciation with all segments present. In the second version, the prefix was reduced in a prescribed way: *be-* as [b], *ge-* as [x], *ont-* as a nasalized [ɔ], and *ver-* as [f]. The unreduced prefix realizations were on average 42 ms longer (138 ms) than the reduced ones (96 ms). The stems of the words were always unreduced.

The experiment contained as fillers 24 existing words and 96 pseudowords with the same four prefixes. Half of these were reduced. The experiment started with four existing words and four pseudowords to familiarize the participants with the task.

Every participant heard both unreduced and reduced realizations but only one realization of a given word. Target and filler items were randomized.

2.1.3. Procedure

Participants were instructed to decide as quickly as possible whether the form they heard was a word or a pseudoword. Participants responded by pressing the *yes* button on a button box with their preferred hand or the *no* button with their non-preferred hand. Stimuli were presented through closed headphones. Reaction times were measured from stimulus onset. Each new trial was initiated 2500 ms after offset of the previous stimulus. If a participant did not respond within these 2500 ms, a time-out response was recorded.

2.2. Results

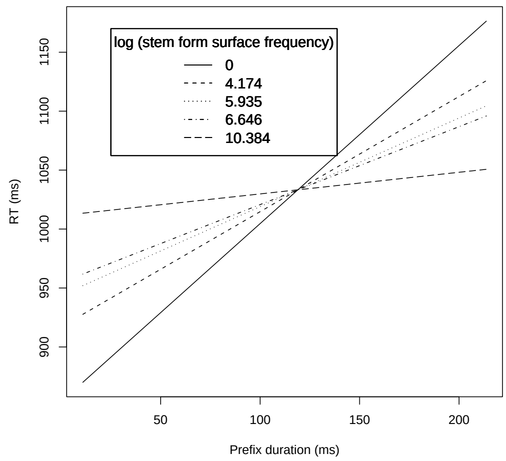
Participants produced 2836 correct responses, 532 incorrect responses, and 15 timeouts. We analyzed the correct versus incorrect responses by means of generalized linear mixed-effect models [3] with participant and word as crossed random factors, and with contrast coding for factors. We included as predictors the reduction of the prefix (yes/no), the duration of the prefix (in ms), the distance from the prefix to the uniqueness point (UP) of the word (in ms), and the distance from the UP to the end of the word (in ms). Moreover, we entered four frequency measures: the surface and lemma frequencies of the complete word and of the word minus the prefix (henceforth stem form), all logged and based on CELEX. Prefix was added as control variable.

We observed an interaction of the reduction of the prefix with the surface frequency of the word: For unreduced realizations, correct responses were more likely the higher the frequency of the word (estimated coefficient, henceforth $\beta, = 0.41, F(1, 3364) = 26.62, p < 0.0001$). Frequency did not predict accuracy for reduced realizations.

We analyzed the reaction times (RTs) for the correct responses, using linear mixed effect models, also with participant and word as crossed random factors [3]. We excluded trials following timeouts and RTs longer than 1000 ms post stimulus offset. We entered the predictors mentioned above as well as the RT on the preceding trial. After the initial fit, data points for which the absolute standardized residuals were greater than 2.0 were removed and the model was refitted.

Participants tended to maintain their local speed (RT on preceding trial: $\beta = 0.04, F(1, 2566) = 12.98, p < 0.001$). Participants were delayed by reduced prefixes ($\beta = 114.49, F(1, 2566) = 20.24, p < 0.0001$). Surprisingly, longer prefix durations, reflecting less reduction, also elicited longer RTs ($\beta = 1.51, F(1, 2566) = 8.04, p < 0.01$). In addition, participants responded more slowly to words with a greater distance from the prefix to the UP ($\beta = 0.11, F(1, 2566) = 7.60, p < 0.01$), especially if the prefix was unreduced (interaction $\beta = 0.25, F(1, 2566) = 10.49, p < 0.01$).

Figure 1: The combined effects of prefix duration and the surface frequency of the word's stem form on the RTs in Experiment 1.



Two frequency measures reached significance. First, listeners responded faster to words with higher surface frequencies ($\beta = -19.83, F(1, 2566) = 15.10, p < 0.0001$). Second, the surface frequency of the stem form showed a main effect ($\beta = 15.21, F(1, 2566) = 6.33, p < 0.05$), in interaction with the duration of the prefix ($\beta = -0.13, F(1, 2566) = 4.86, p < 0.05$). A lower frequency of the stem form facilitated comprehension at shorter prefix durations, as illustrated in Figure 1. The lines in this figure represent the minimum, the three quartiles, and the maximum values of stem form frequency.

2.3. Discussion

A higher surface frequency of the complete word improved participants' accuracy but only if the prefix was unreduced. In the RTs, the benefit from higher frequencies emerged both for the reduced and the unreduced realizations. Hence, there is no evidence that surface frequency would play a greater role in the comprehension of reduced than in the comprehension of unreduced realizations.

Longer durations of the prefix and of the part between the prefix and the UP (Prefix2UP) delayed participants' responses. That is, listeners responded more slowly when the information disambiguating the word from its lexical competitors came in over a longer stretch of time, increasing lexical competition. Importantly, the response delay resulting from a longer Prefix2UP was greater for unreduced than for reduced words. This suggests that unreduced prefixes facilitate lexical access to the words consistent with the prefix and the following segments, which also results in increased lexical competition.

The interaction illustrated in Figure 1 shows that lower stem form frequencies were beneficial for the comprehension of realizations with shorter prefixes. Since shorter prefixes are more difficult to identify, the following stems may have become more activated than the full forms. These stem forms had to be suppressed to make comprehension of the full forms possible. Suppression was easier when the stem forms were less activated, that is, when the stem forms were of a lower frequency of occurrence.

The question arises whether these results generalize from words with highly reduced prefixes and unreduced stems to more natural, less reduced realizations with segment deletions in both the prefix and the stem. We investigated this in Experiment 2.

3. EXPERIMENT 2

Experiment 2 investigated the comprehension of prefixed words that were produced at a low or at a high speed rate. The high speed rate led above all to shorter segment durations.

3.1. Method

3.1.1. Participants

Twenty-four native speakers of Dutch, who had not participated in Experiment 1, were paid for their participation.

3.1.2. Materials

We selected 127 prefixed words starting with one of the 11 prefixes (or particles) *aan-*, *be-*, *bij-*, *ge-*, *in-*, *om-*, *onder-*, *ont-*, *op-*, *over-*, or *ver-*, with 16 words for the prefixes that were also tested in Experiment 1, and 9 words for each new prefix. The words covered a broad range of frequencies.

We also selected 89 morphologically simple existing words as fillers and created 218 pseudowords with the same morphological structure as the existing words in the experiment. Finally, we selected 7 existing words and 7 pseudowords, with varying morphological structure, to familiarize the participants with their task. All words, except those starting with *ge-*, were infinitives.

The same female speaker as in Experiment 1 recorded the words, but this time in a naming experiment, in which she read aloud words appearing on a computer screen. The experimental words were first presented at a low presentation rate (interstimulus interval of 15000 ms) and then at a high rate (700 ms). Our speaker realized the existing prefixed words on average 163 ms longer at the low (average duration: 609 ms) than at the high rate (446 ms).

Participants heard words produced at both presentation rates, but only one realization of every word. Target and filler items were randomized.

3.1.3. Procedure

The procedure was the same as in Experiment 1.

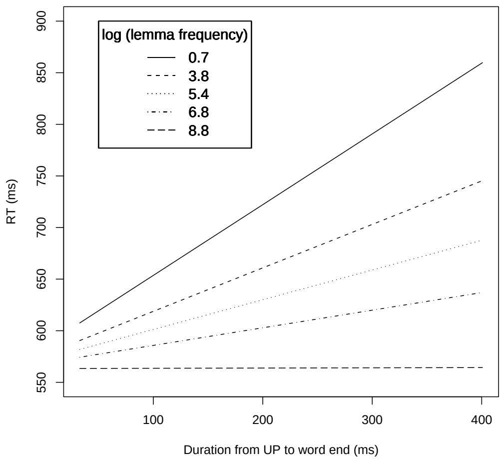
3.2. Results

Participants produced 2577 correct and 386 incorrect responses, and 50 timeouts. We analyzed the correct versus incorrect responses, using generalized mixed effect models with the same variables as in Experiment 1. Correct responses were more likely for words realized at the slow presentation rate ($\beta = 0.09, F(1, 2593) = 56.30, p < 0.0001$) and for words of a higher lemma frequency ($\beta = 0.61, F(1, 2593) = 13.13, p < 0.0001$). Importantly, the effect of lemma frequency was greater for words produced at the low than at the high presentation rate ($\beta = 0.24, F(1, 2593) = 7.96, p < 0.01$).

We analyzed the RTs for the correct responses, excluding trials following timeouts and RTs longer than 1000 ms post stimulus offset. In addition, we removed data points for which the standardized residuals of the initial fit were smaller than -2.0 or greater than 2.0. We then refitted the model.

As in Experiment 1, participants tended to maintain their local response speed (RT on preceding trial: $\beta = 0.07, F(1, 2177) = 38.92, p < 0.0001$). In addition, longer prefixes ($\beta = 0.45, F(1, 2177) = 90.68, p < 0.0001$), greater distances between the prefix and the UP ($\beta = 0.51, F(1, 2177) = 97.43, p < 0.0001$), and greater distances from the UP to the end of the word ($\beta = 0.74, F(1, 2177) = 132.53, p < 0.0001$) slowed listeners. Participants were speeded up by a higher lemma frequency of the word ($\beta = -2.68, F(1, 2177) = 47.49, p < 0.0001$), especially if the distance from the UP to the end of the word was greater ($\beta = -0.08, F(1, 2177) = 10.78, p < 0.01$). This interaction is illustrated in Figure 2, with the lines representing the minimum, the three quartiles, and the maximum values of the lemma frequency.

Figure 2: The combined effects of the distance from the UP to the end of the word and the word's lemma frequency on the RTs in Experiment 2.



3.3. Discussion

Experiment 2 replicated the finding in Experiment 1 that longer durations of (the parts of) the word lead to slower responses. Listeners responded faster when the information distinguishing the word from its lexical competitors came in earlier.

We also observed a role for the lemma frequency of the word. Listeners were delayed if the distance from the UP to the end of the word was longer. At the UP, listeners posses all information necessary to identify the current morpheme, but they are still in uncertainty about exactly which lemma is presented. For instance, once they have heard [ɔntplɔ], they know that the word contains the morphemes [ɔnt] and [plɔf], but they cannot yet choose between [ɔntplɔfə] 'to explode', [ɔntplɔfɪŋ] 'explosion', and [ɔntplɔfbar] 'explosive'. The competition between such lemmas grows stronger, the longer it takes before the disambiguating information comes in, that is, the greater the distance from the UP to the end of the word. The competition is especially harmful for lemmas of a low frequency of occurrence, since most of them have competitors of a higher frequency, which slows down recognition [7]. This explains the interaction illustrated in Figure 2.

4. GENERAL DISCUSSION

This study demonstrates that the deletion of segments delays word recognition, independently of the frequency of the word or of its stem form. Simultaneously, however, shorter durations may speed up word comprehension, since disambiguating information becomes available faster, quickly terminating lexical competition either from the word's stem or from the word's morphological continuation forms. Fast incoming disambiguating information is especially important for the comprehension of low frequency words, since it is these words that typically suffer most from competition with words of a higher frequency of occurrence [cf. 7].

Speakers typically reduce words to a greater extent, the higher their frequency of occurrence [10]. Since we have shown that a word's frequency of occurrence is less important for the recognition of its reduced than of its unreduced realizations, this frequency effect in production is unlikely to be listener driven. It suggests an important role for speaker-specific processes such as the selection and articulation of the word.

5. REFERENCES

[1] Aylett, M., 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.

[2] Baayen, R. H., Piepenbrock, R., Gulikers, L. 1995. *The CELEX lexical database (CD-ROM)*. University of Pennsylvania, Philadelphia: Linguistic Data Consortium.

[3] Bates, D.M. 2005. Fitting linear mixed models in R. *R News* 5, 27–30.

[4] Bybee, J. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

[5] Ernestus, M. 2000. *Voice assimilation and segment reduction in casual Dutch: A Corpus-Based Study of the Phonology-Phonetics Interface*. Utrecht: LOT.

[6] Ernestus, M., Baayen, R.H., Schreuder, R. 2002. The recognition of reduced word forms. *Brain and Language* 81, 162-173.

[7] Goldinger, S.D., Luce, P.A., Pisoni, D.B. 1989. Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language* 28, 501-518.

[8] Jespersen, O. 1922. *Language: its nature, development and origin*. London: George Allen & Unwin.

[9] Johnson, K. 2004. Massive reduction in conversational American English. *Spontaneous speech: data and analysis. Proceedings of the 1st session of the 10th international symposium*. Tokyo: The National International Institute for Japanese Language, 29–54.

[10] Jurafsky, D., Bell, A., Gregory, M., Raymond, W. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In: Bybee, J., Hopper, P. (eds.), *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 229–254.

[11] Mitterer, H., Ernestus, M. 2006. Listeners recover /t/s that speakers reduce: Evidence from /t/-lenition in Dutch. *Journal of Phonetics* 34, 73-103.