Semantic context effects in the recognition of acoustically unreduced and reduced words

Marco van de Ven[1], Benjamin V. Tucker[2], Mirjam Ernestus[3,1]

[1]Max Planck Institute for Psycholinguistics, The Netherlands
[2] University of Alberta, Canada
[3] Radboud University Nijmegen, The Netherlands

Marco.vandeVen@mpi.nl, bvtucker@ualberta.ca, Mirjam.Ernestus@mpi.nl

Abstract

Listeners require context to understand the casual pronunciation variants of words that are typical of spontaneous speech [1]. The present study reports two auditory lexical decision experiments, investigating listeners' use of semantic contextual information in the comprehension of unreduced and reduced words. We found a strong semantic priming effect for low frequency unreduced words, whereas there was no such effect for reduced words. Word frequency was facilitatory for all words. These results show that semantic context is relevant especially for the comprehension of unreduced words, which is unexpected given the listener driven explanation of reduction in spontaneous speech.

Index Terms: acoustic reduction, word recognition, speech perception, semantics, latent semantic analysis

1. Introduction

In spontaneous speech, words are often pronounced much shorter than in careful speech [2, 3]. For example, the English words *yesterday* and *ordinary* can be pronounced like *yeshyay* and *onry*. The deletion of single or multiple segments or syllables is highly common in spontaneous speech. For instance, in the Buckeye Corpus of American English conversational speech (306,652 word tokens) [4], complete syllables were deleted in 6% of the words [3]. Previous research has shown that listeners have difficulty understanding highly reduced words in isolation [1] and that listeners simultaneously use acoustic and semantic/syntactic cues in the context to predict reduced words [5]. The present study compares the contribution of semantic context in the recognition of reduced words and unreduced words.

We predicted that listeners would rely more on any cue, including semantic information, in the comprehension of reduced words compared to unreduced words, since reduced words are more difficult to recognise. However, recent studies on spoken word recognition have shown that this is not necessarily the case.

For example, [6] conducted two auditory lexical decision tasks with reduced and unreduced prefixed words, in order to test whether frequency plays a role in the recognition of these words. As expected, the results for the unreduced words showed that participants performed better for more frequent words. Surprisingly, however, no frequency effect was found for the reduced words. This finding suggests that the role of frequency cues in the recognition of reduced words is (at best) marginal. Thus, cues that facilitate the recognition of unreduced speech do not necessarily facilitate the recognition of reduced speech.

Similar results were obtained by [5]. These authors conducted two visual and two auditory cloze tasks, in order to investigate which aspects of the context contribute to the understanding of reduced words. In these experiments, participants were asked to predict the (reduced) words on the basis of just the context (the reduced words were masked), choosing from four semantically and syntactically plausible options. Their results showed that the effects of trigram frequency and word frequency were restricted to the visual presentation mode; in the auditory presentation mode these frequency effects appeared to be overruled by acoustic cues. Thus, in line with [6], this study suggests that the role of frequency cues becomes marginalised in the comprehension of reduced speech.

No research has investigated the role of semantics in the comprehension of acoustically reduced words. The present study investigates how semantic (separate from syntactic) context affects the comprehension of reduced and unreduced words. Further, we tried to replicate the absence of frequency effects for reduced speech (e.g., [5, 6]. We describe two auditory lexical decision experiments, in which participants were presented with either unreduced (Experiment 1) or reduced (Experiment 2) target words.

We used Latent Semantic Analysis (LSA) to estimate to what extent words are semantically related [7]. This computational technique rests on the assumption that semantically related words tend to occur in similar contexts. LSA scores are based on the frequencies with which words co-occur in similar paragraphs/texts, using large written corpora to estimate these frequencies. LSA places words in a multi-dimensional vector space on the basis of these co-occurrence frequencies; the cosine distance between the vectors for these words is then taken as a measure of the words' semantic similarity. LSA scores range from -1 to 1, where values close to 1 indicate a high semantic relatedness, and values close to -1 indicate a low semantic relatedness. Previous research has shown that LSA can simulate human behaviour in various types of experiments. For example, LSA can simulate human similarity rating [8, 9]. More importantly, LSA scores can simulate semantic priming effects in visual lexical decision experiments [8].

Furthermore, we estimate word frequency using frequency counts for the spoken portion of the Corpus of Contemporary American English (385 million word tokens) [10]. This frequency measure can be taken as a rough estimate of how frequently a given word occurs in American English.

6 – 10 September, Brighton UK

2. Experiment 1

2.1. Introduction

We used an auditory lexical decision task to investigate the effects of the semantic relation of a word with its preceding word and of word frequency on speech comprehension. We manipulated the semantic relation between the words in our experiment by constructing word pairs with LSA scores ranging from 0.35 (mildly related) to 0.93 (highly related). The members of a word pair were presented directly after each other in consecutive trials.

2.2. Participants

Twenty native speakers of English from the University of Alberta, Department of Linguistics participants pool took part in the experiment, and received course credit for their participation.

2.3. Materials

We extracted 154 nouns, with varying frequencies (range: 40-58322), from the Corpus of Contemporary American English [10]. These nouns were used to construct 77 word pairs with varying LSA scores (range: 0.35 to 0.93). This range includes words that are highly related (e.g., *saddle - horse*), moderately related (e.g., *table - bowl*), and mildly related (e.g., *tower - statue*). Further, the experiment contained 87 semantically unrelated filler pairs, and 128 pseudowords. We used a limited number of nonword fillers as to induce a "YES"-response bias, which will make it particularly difficult to find any priming effects in our data. As a consequence, any priming effects that show up are robust effects.

Existing words and nonwords were pseudorandomised, such that there were no more than six existing words and no more than three nonwords in succession. We avoided rhyme and/or alliteration between all immediately adjacent words in the experiment.

The materials were produced by a male speaker of Canadian English. The speaker was asked to pronounce the words in a clear citation style. The words were presented in a fully randomised order, in order to prevent that the speaker produced faster realisations for words that were highly related to their preceding word (e.g. a shorter pronunciation for "cat" when preceded by "dog"). A different native speaker of Canadian English verified that all the existing words were pronounced naturally and clearly.

2.4. Procedure

Participants listened to all stimuli and made lexical decisions by means of a button press. The experiment was self-paced. The experiment took place in a soundproof booth, and the materials were presented over closed headphones at a comfortable listening level. The experiment lasted approximately 15 minutes.

2.5. Results

We analysed the response accuracy by means of generalised linear mixed-effects models with the logit link function ([11]). As we are primarily interested in the effects of gradual differences in semantic relatedness on word comprehension (as opposed to semantic match/mismatch priming), we decided to analyse only the target words which have LSA scores (with their preceding word) between 0.34 and 1. This range of semantic similarity roughly represents words that are highly, moderately, and mildly related. Thus, if any effects of semantic similarity show up in our analyses, these effects are caused by subtle (instead of semantic match/mismatch) distinctions in semantic relatedness.

Participants produced 1503 correct responses, 15 incorrect responses, and 42 time outs. We included the fixed effect factors *word frequency* (log word frequency), *lsa* (LSA score of the word with the preceding word), and *previous RT* (log of the response time (RT) from the preceding trial). We included the RT on the preceding trial as an indication of the participants' local response speed, which may reflect at which point in the processing of the word participants made their lexical decision. If participants responded slowly their response may have occurred long after they recognised the word, and consequently may show smaller effects of semantic relatedness and lexical frequency.

In addition, we included several control variables, namely the fixed effect *trial number*, and random effects for *Participant* and *Word*. These variables were included mainly to reduce variance, as to increase the likelihood of finding effects for the main predictors. None of the fixed effects proved significant.

Further, we analysed the log RTs (from stimulus offset) for the correct responses. Log RTs were used in order to obtain a normal distribution. We excluded data points for which the standardised residuals were smaller than -2.5 or larger than 2.5. We included the same predictors as above (i.e. the main predictors *word frequency*, *lsa*, *previous RT*, and the control variables *trial number*, *Participant*, and *Word*). In addition, we included the predictor *word duration* (log of the stimulus duration); we took the log of the duration, such that the RTs and the durations were on the same scale.

A summary of the results is provided in Table 1. The control variables showed the expected effects: participants responded faster towards the end of the experiment, to longer words, and if their preceding RT was also short. More importantly for our research question, we found an effect of *word frequency*: participants responded more rapidly to words of a higher word frequency. In addition, we found an interaction between *word frequency* and *lsa*. This interaction is visualised in Figure 1.

Table 1: *Results for the RTs in Experiment 1.*

Predictor	$\hat{\beta}$	F	p
trial number	-0.0005	15.57	$< .0001$
word duration	-1.048	40.69	$< .0001$
previous RT	0.239	89.27	$< .0001$
word frequency	-0.333	10.79	$< .01$
lsa	-3.191	0.39	n.s.
word frequency:lsa	0.418	7.88	$< .05$

For semantically unrelated words, word frequency has a large influence on participants' RTs. The influence of word frequency decreases for words that are semantically more related to their preceding word. For words in the highest LSA range, there is hardly any effect of frequency, and if there is, higher frequency hinders rather than facilitates comprehension. The interaction of LSA with frequency for words with low or intermediate semantic relatedness with their preceding words is as expected, as it reflects cue trading, which has also been documented by ([12]) for the concreteness of a word's meaning and family size. The unexpected effect for semantically highly related words suggests that a high word frequency inhibits the use of semantic contextual information in word recognition. Possibly, this is due to a stronger suppression of competitors for
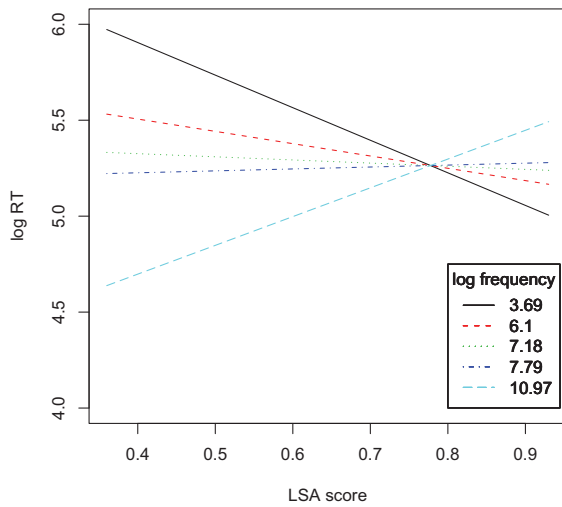
Figure 1: *The combined effects of LSA score and word frequency on RTs in Experiment 1, as estimated by the regression model.*

Figure 2: *Word durations for the reduced and unreduced words.*

higher frequency words, which delays the recognition of semantically related words.

The question arises whether similar word frequency and semantic context effects can be observed in reduced speech, which is characterised by shorter word durations and segment deletions. We address this issue in Experiment 2, which contained more reduced speech materials.

3. Experiment 2

3.1. Participants

Twenty native speakers of English from the University of Alberta, Department of Linguistics participant pool took part in the experiment, each received course credit for participating in the experiment. Participants in Experiment 2 had not participated in Experiment 1.

3.2. Materials

We created another set of recordings for the materials used in Experiment 1. In these recordings, the same speaker of Canadian English was asked to produce the same list of words as quickly as possible, in order to elicit reduced speech. Again, these realisations were verified by a native speaker of Canadian English, especially paying attention to whether the words were reduced in a natural manner. The durations of these reduced words differed significantly from the durations of the unreduced words used in Experiment 1, as illustrated by Figure 2.

3.3. Procedure

The procedure was identical to Experiment 1.

3.4. Results and discussion

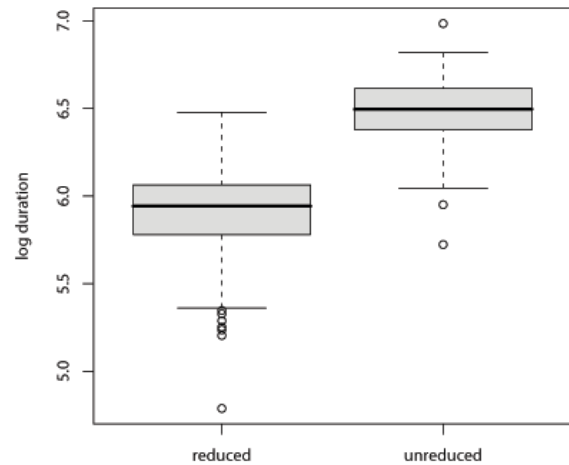We again analysed the response accuracy by means of generalised linear mixed-effects models ([11]), including the same predictors as for Experiment 1, and we only analysed words that have LSA scores between 0.34 and 1. Participants produced 1457 correct responses, 63 incorrect responses, and 40 time outs. None of the fixed effects proved significant.

Further, we analysed the RTs for the correct responses, using the same exclusion criteria and the same predictors as for Experiment 1. A summary of the results is provided in Table 2. The control variables showed exactly the same effects as in Experiment 1. Further, we observed an effect of *word frequency*: participants responded more quickly to higher frequency words. We did not observe an effect of LSA.

Table 2: *Results for the RTs in Experiment 2*

Predictor	$\hat{\beta}$	F	p
trial number	-0.0005	55.02	$< .0001$
previous RT	0.222	97.23	$< .0001$
word duration	-0.803	36.06	$< .0001$
word frequency	-0.042	6.89	$< .05$

Comparing the two experiments, we observed some striking differences in the roles of semantic context and word frequency in auditory word recognition. In Experiment 1, we found that semantic contextual information generally facilitates comprehension, whereas in Experiment 2 there was no effect of semantic context whatsoever. Furthermore, whereas frequency showed a main effect in Experiment 2, we found an interaction with LSA in Experiment 1. The question arises whether these differences attain statistical significance. Therefore, a regression model was fitted for the combined data sets of Experiments 1 and 2.

We first analysed the response accuracy for the combined results. We included the same predictors as in the previous analyses, in addition to the predictor *Experiment* (reduced/unreduced). We will only report relevant interactions with *Experiment*. Participants produced more correct responses in the unreduced than in the reduced experiment ($\beta =$

-1.73, $F(1, 3036) = 30.08$, $p < .0001$). This result confirms that the materials in Experiment 2 were significantly reduced, and generally more difficult to understand.

We then analysed the RTs for the correct responses, again for the combined results of Experiments 1 and 2. We excluded data points for which the standardised residuals were smaller than -2.5 or larger than 2.5. We included the same random and fixed effects as above, except for word duration, as it was highly correlated with *Experiment*, see Figure 1. We found a three-way interaction between *word frequency*, *lsa*, and *Experiment* ($F(1, 2630) = 8.04$, $p < .005$). This interaction confirms the different roles of semantic context in the comprehension of unreduced and reduced speech: whereas in Experiment 1 it shows an effect in interaction with *word frequency*, it shows no effect at all in Experiment 2. There was no two-way interaction between *Experiment* and *word frequency*, which suggests that the effect of frequency is similar in the two experiments.

4. General discussion

This study investigated the roles of semantic context and lexical frequency in the comprehension of unreduced and reduced words, by means of two auditory lexical decision experiments. Semantic context generally facilitated the comprehension of unreduced words, but, surprisingly, did not play a role in the comprehension of reduced words. Apparently, in adverse listening conditions, listeners do not rely more heavily on the available cues. Moreover, this finding is unexpected given a listener driven account of reduction, which states that speakers reduce especially those words that can be easily understood by the listener because they are predictable. Such an account would assume that semantic context plays a significant role in the comprehension of reduced speech as it codetermines the likelihood of words.

A higher lexical frequency facilitated the comprehension of both reduced and unreduced words. Also this finding is unexpected, since previous studies suggested that frequency effects become marginal for reduced speech. Interestingly, for unreduced words the effect of frequency is modulated by the word's semantic relation with the preceding word: Frequency facilitated comprehension especially for words presented in unrelated semantic contexts. Apparently, listeners do not use all available cues, but rely on cues especially when other cues are less informative.

In conclusion, our results indicate that lexical frequency also plays a role in the comprehension of reduced speech, at least under the conditions of our experiments. In contrast, semantic context appeared to play a smaller role for reduced compared to unreduced words. Further research is required to test whether this is also the case if the semantically related words occur in the same sentence, instead of in independent trials of a lexical decision experiment. Furthermore, whereas small differences in semantic relatedness appear not to play a role in reduced speech, bigger differences may still show an effect.

5. Acknowledgements

6. References

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

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

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

[4] Pitt, M.A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., Fosler- Lussier, E. (2007). Buckeye Corpus of Conversational Speech (2nd release) [www.buckeyecorpus.osu.edu] Columbus, OH: Department of Psychology, Ohio State University (Distributor).

[5] Van de Ven, M.A.M., Ernestus, M., Schreuder, R., 2009. Contextual influences in the understanding of spontaneous speech. Manuscript submitted for publication.

[6] Ernestus, M., Baayen, R.H., 2007. The comprehension of acoustically reduced morphologically complex words: The roles of deletion, duration and frequency of occurrence. In: Proceedings of the 16th International Congress of Phonetic Sciences, Saarbruecken, Germany, pp. 773-776.

[7] Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., Harshman, R. (1990). Indexing By Latent Semantic Analysis. Journal of the American Society For Information Science, 41, 391-407.

[8] Landauer, T. K., Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, 104(2), 211-240.

[9] Landauer, T. K., Foltz, P., Laham, D. (1998). An introduction to Latent Semantic Analysis. Discourse Processes, 25, 259-284.

[10] Davies, Mark. (2008-) The Corpus of Contemporary American English (COCA): 385 million words, 1990-present. Available online at http://www.americancorpus.org.

[11] Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language, 59 (4): 434-446.

[12] Feldman, L. B., Brown, D. B and Pastizzo, M. J. (2006). Semantic Influences on Morphological Facilitation: Concreteness and Family Size. The Mental Lexicon Inaugural Issue, 1, 59 - 84.

[13] Tucker, Benjamin V., Warner, Natasha. (2007). Inhibition of processing due to reduction of the American English flap. Proceedings of the 16th International Congress of Phonetic Sciences, Saarbrücken, August 2007.