How are words reduced in spontaneous speech?

Holger Mitterer

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Abstract

Words are reduced in spontaneous speech. If reductions are constrained by functional (i.e., perception and production) constraints, they should not be arbitrary. This hypothesis was tested by examing the pronunciations of high- to mid-frequency words in a Dutch and a German spontaneous speech corpus. In logistic-regression models the "reduction likelihood" of a phoneme was predicted by fixed-effect predictors such as position within the word, word length, word frequency, and stress, as well as random effects such as phoneme identity and word. The models for Dutch and German show many communalities. This is in line with the assumption that similar functional constraints influence reductions in both languages.

Key words: reduction, spontaneous speech, Dutch, German, functional constraints

Introduction

With the availability of large corpora, it has become clear that spontaneous speech contains many reduced word forms (e.g., yesterday   "yesay"). This gives rise to the questions how words are actually reduced (i.e., why is it "yesay" and not "seday"). Are the reduced form arbitrary "contracts" between speaker and listener, just as the contract that the sound chain "dog" refers to a canine, or are reductions constrained by functional factors?

It has been argued that functional factors influence phonological assimilations (Kohler, 1990; Mitterer, Csépe and Blomert, 2006). Steriade (2001), for instance, argued that directional asymmetries in assimilation are related to the perceptual salience of the assimilated segment. Retroflex stops assimilate mostly in pre-vocalic position, while apical stops assimilate mostly in post-vocalic position. This pattern correlates (negatively) with the salience of place cues for apical and retroflex stops. The formant transitions for retroflex stops are more salient in postvocalic position, while the formant transitions for apical stops are more salient in prevocalic position. Hence, both types of stops assimilate in the position in which their place cues are less salient. This led Steriade to propose the following hypothesis (p. 222): "The likelihood that a lexical representation R will be realized as R' is a function of the perceived similarity between R and R'."

If this constraint also functions for more severe reductions than assimilations, we should find commonalities in reductions over different words and even different languages. To investigated this hypothesis, pronunciations of mid-to-high frequency words with two or three syllables

and six to nine phonemes were investigated in two corpora, a Dutch (CGN, Corpus of Spoken Dutch) and a German (the Kiel Corpus) corpus.

Method

From the Dutch corpus, 15 words each with 2 or 3 syllables and 6 to 9 phonemes were selected that had a lexical frequency of more than 1 per million and less than 100 per million. For each word, we randomly selected 15 tokens from the corpus which were then phonetically transcribed.

The Kiel Corpus has a more limited vocabulary, so that it was impossible to have a similarly rigorous data gathering approach as for Dutch. Therefore, we selected all words with 2 or 3 syllables and 6 to 9 phonemes from the same frequency range as in the Dutch corpus ($1/10^6 <$ frequency $< 100/10^6$).

For each phoneme, it was coded whether it was present, weakened, or absent. A linear mixed-effect model with a logistic linking function (Jaeger, in press) was then used to test how well deletion can be predicted by structural (fixed-effect) predictors such as: number of syllables and phonemes, relative position in the word (3d order polynomial), stress, Syllable Part (Onset - Nucleus - Coda), Complexity (is the phoneme part of a consonant cluster) and Word frequency. Additionally phoneme and word were added as random effects. Insignificant factors were deleted by backward elimination.

Results

Table 1 shows the deletion likelihood in the various major "cells" of the design. Phonemes are more likely to be deleted in words with many phonemes, especially if the word has only two syllables.

Table 1. Deletion likelihood for phonemes in words with different number of phonemes and syllables in English and Dutch.

Language	Number of Syllables	Number of Phonemes			
		6	7	8	9
Dutch	2	0.07	0.14	0.14	0.22
	3	0.09	0.09	0.12	0.14
German	2	0.04	0.07	0.13	--
	3	0.08	0.05	0.08	0.08

The regression analysis (see Table 2) shows that the differences in Table 1 are significant. In both data sets, reduction of a phoneme in a 2-syllable word gets more likely as the word contains more phonemes. There are many other communalities: In both data sets, phonemes are especially likely to be deleted in the middle of words, in the syllable coda, and in unstressed

syllables. The only divergence is that phonemes in consonant clusters (=Predictor Complexity) are only more likely to be reduced in Dutch than phonemes in simple syllable onsets and codas.

Table 2. Partial effects in the regression analysis for both the Dutch and the German samples.

Predictor	Effect in	
	Dutch	German
N(Phonemes) x N(Syllable)	More reduction in long 2-syllable words	More reduction in long 2-syllable words
Position	inverted U-shape	inverted U-shape
Syllable Part	Coda>Nucleus>Onset	Coda>Nucleus>Onset
Stress	Stress inhibits reduction	Stress inhibits reduction
Complexity	More reduction in complex onsets and codas	n.s.
Complexity x Stress	Stronger effect of Complexity with stress	n.s.
Position x Syllable Part	Less reduction for onsets in initial position	Less reduction for onsets in initial position
Lexical Frequency	More reduction in frequent words	More reduction in frequent words
R^2	0.20	0.27

The analysis of the random effects also shows an interesting pattern. Phonemes that are likely to reduce in one language are also likely to reduce in the other language ($r = 0.69$, $p < 0.001$). In both languages, "s" is very unlikely to be reduced and the alveolar stops /t/ and /d/ as well as the low-amplitude /h/ are likely to be reduced.

Discussion

The results reveal some expected effects and some unexpected effects. The effect of stress - phonemes in stressed syllables are unlikely to be reduced - is very much expected (Shockey, 2003). However, the fact that phonemes are especially likely to be reduced in word-medial positions is surprising. Traditionally, it is assumed that especially word endings are likely to be reduced. Descriptively, this is true for the current data set as well. Reduction often occur often at the end of words. This effect is however sufficiently explained by the high deletion likelihood for syllable codas and for unstressed syllables. The final segment of a word is often a syllable coda and

(in Dutch and German) often unstressed. This appear to be the causal factors that lead to the deletion of word-final consonants.

Importantly, the strong commonalities in the two data sets are in line with the assumption that reductions are functionally constrained. Although the models only explain a moderate amount of the variance (20%), it is clear that reduced forms do not arise arbitrarily. Some of the effects clearly seem to have a perceptual basis, such as the relatively high likelihood of /h/-deletion. Mielke (2003) already argued that the low amplitude of /h/ - causing it to be perceptually not salient - is the basis for /h/-deletion. Also the (strong) effect of stress can be linked to amplitude: Stressed syllables tend to be longer and louder than unstressed syllables. Reducing segments in stressed syllables would therefore be perceptually salient. Strong conclusions can, however, only be drawn if these results are complemented with independent perceptual evidence, that may or may not converge with the current results.

Acknowledgements

References

Jaeger, T. F. (in press). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language.

Kohler, K.J. 1990.   Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. J. H. A. Marchal (Ed.), Speech Production and Speech Modelling (pp. 69-92). Dordrecht: Kluwer.

Mielke, J. 2003. The interplay of speech perception and phonology: Experimental evidence from Turkish. Phonetica, 60, 208-223.

Mitterer, H., Csépe, V. and Blomert, L. 2006. The role of perceptual integration in the perception of assimilated word forms. Quarterly Journal of Experimental Psychology, 59, 1305-1334.

Shockey, L. 2003. Sound patterns of spoken English. Cambridge, MA: Blackwell.

Steriade, D. 2001. Directional asymmetries in place assimilation: a perceptual account. In E. Hume and K. Johnson (Eds.), The role of speech perception in phonology (pp. 219-250). New York, NJ: Academic Press.