N-gram frequency effects on speech production in German and Mandarin Chinese

Ching Chu Sun[1], Peter Hendrix[2]

[1]) Eberhard Karls University
ching-chu.sun@uni-tuebingen.de
[2]) Eberhard Karls University

Frequency effects on acoustic durations have been reported, both at the word level (Bell, Brenier, Gregory, Girand, & Jurafsky, 2009; Gahl, 2008; Pluymaekers, Ernestus, & Baayen, 2005) and at the phrase level (Tremblay & Tucker, 2011; Arnon & Cohen Priva, 2013, 2014), with shorter acoustic durations for more frequent words and phrases. Whereas previous work focused on frequency effects for carefully selected phrases that form well-defined lexical units, the word n-grams in the current study are random word trigrams without a predefined linguistic structure or a well-established lexical status. We investigate the pronunciation duration of word n-grams in both elicited speech and spontaneous speech. For elicited speech, we report the results of two reading aloud experiments in German and in Mandarin Chinese. For spontaneous speech, we look into the acoustic durations of multi-word sequences in datasets extracted from a corpus of spontaneous speech in both languages.

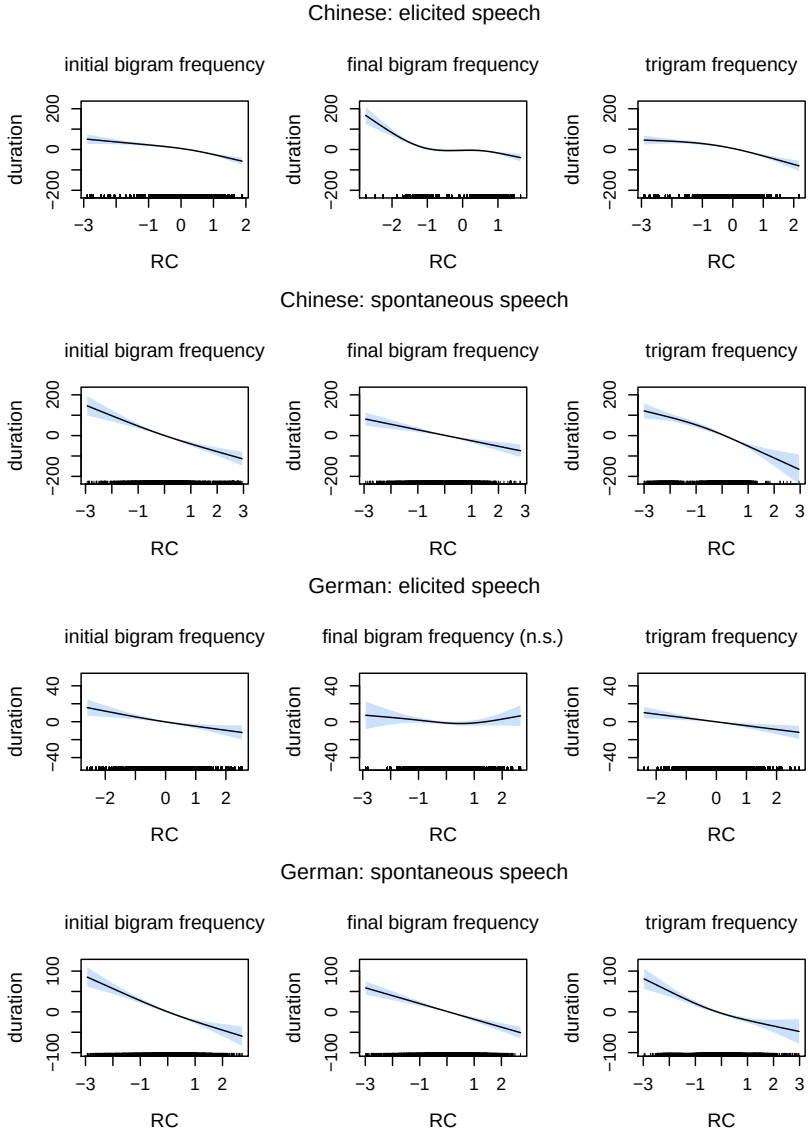
The stimuli for the experiments in German and in Mandarin Chinese were sequences of three words. The response variable is the acoustic duration of the pronunciations. The predictors of primary interest are the frequencies of the word unigrams, bigrams, and trigrams. We obtained these frequencies through Google searches that were restricted to documents in simplified Chinese and German, respectively. The frequency counts for word unigrams, word bigrams and word trigrams are highly correlated. To obtain independent frequency measures, we entered the (log-transformed) Google frequency counts into a principal components analysis with varimax rotation. For both German and Mandarin Chinese, this procedure yielded a set of orthogonal rotated components that each represented one of the six frequency measures of interest.

The residuals of a regular multiple regression model were not normally distributed. We therefore fit quantile regression models of the median with a quantile generalized-additive mixed-effect model (qgamm) (Fasiolo, Goude, Nedellec, & Wood, 2018). The reported results, however, were highly similar in a standard regression of the mean. We entered the rotated components for the unigram, bigram, and trigram frequency counts into the qgamms. Furthermore, we added a random effect of speaker and a number of control variables to each model.

The qgamms fit to the experimental data and the spontaneous speech data in Mandarin Chinese and in German revealed significant effects of (the rotated components that correspond to) the word unigrams on word trigram durations. With the exception of the trigram-final bigram in the experimental data in German, we furthermore observed significant effects of the frequency for all word bigrams in all data sets. Finally, we found significant effects of the frequency of the trigram as a whole over and above the effects of the frequency of the component unigrams and bigrams in Mandarin Chinese as well as in German, both in elicited speech and in spontaneous speech. The effects of bigram frequency and trigram frequency are presented in Figure 1. As can be seen in Figure 1 the bigram and trigram frequency effects tend to be linear or near-linear in nature.

Whereas previous studies focused on acoustic durations of fixed expressions or well-defined lexical bundles, we imposed very few restrictions on the word trigrams selected in the studies reported here. Nonetheless, we observed robust trigram frequency effects in both elicited speech and spontaneous speech. The effects were present in both Mandarin Chinese and German, and were highly significant across the predictor range. The current results suggest that the combinatorial properties of words play a pivotal role in the production of multi-word sequences, even when these sequences do not form fixed expressions or lexical bundles and are therefore unlikely to be stored holistically in the mental lexicon.

Figure 1: Results of the QGAMM fitted to the acoustic durations in the experimental data and in the spontaneous speech data in Mandarin Chinese and German. Plotted are partial effects of the frequency of the trigram-initial bigram (left panels), the trigram-final bigram (middle panels) and the word trigram (right panels).



References

Arnon, I., & Cohen Priva, U. (2013). More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech*, *56*, 349–371.

Arnon, I., & Cohen Priva, U. (2014). Time and again: The changing effect of word and multi-word frequency on phonetic duration for highly frequent sequences. *The Mental Lexicon*, *9*, 377–400.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language*, *60*(1), 92–111.

Fasiolo, M., Goude, Y., Nedellec, R., & Wood, S. N. (2018). Fast Calibrated Additive Quantile Regression [Computer software manual]. (https://arxiv.org/abs/1707.03307v2)

Gahl, S. (2008). Time and thyme are not homophones: the effect of lemma frequency on word durations in spontaneous speech. *Language*, *84*(3), 474–496.

Pluymaekers, M., Ernestus, M., & Baayen, R. H. (2005). Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, *118*, 2561-2569.

Tremblay, A., & Tucker, B. (2011). The effects of n-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon*, *6*, 302–324.