Speech rate effects on the vowel length contrast in production and perception: Evidence from Swiss Standard German

Franka Zebe[1], Felicitas Kleber[2]
[1]*Phonetics Laboratory, UZH,* [2]*Institute of Phonetics and Speech Processing, LMU Munich*

It is broad consensus that segmental duration varies as a function of speech rate (e.g. [1, 2]). Perceptual category boundaries along duration continua between long and short segments shift with an alteration of speaking rate in the surrounding segments such that the same stimuli with ambiguous duration are classified as phonemically long when the overall tempo of the carrier phrase is accelerated and as short when it is decelerated [3]. This compensation effect has been found for several typologically diverse languages, including English [4], Italian [1], Maltese [2], and German [5, 6], reflecting a cross-linguistic robustness of perceptual compensation for segmental shortening in fast speech, despite some phoneme-dependent differences in production, with longer segments being more affected by rate than short ones [7, 8]. Similar, though underresearched, fine compensation differences also seem to exist between regional varieties: Swiss listeners of Standard German (SSG) compensated more for rate differences than German listeners [5] which might be linked to a greater role of durational cues to quantity contrasts in SSG compared to German varieties [9]. Additionally, these cues may be differently used in production and perception: Although SSG speakers consistently produced underlying short vowels with a very short duration, the same SSG listeners also accepted longer durations in the classification of short vowels [6]. This study therefore examines the production and perception of the vowel length contrast in a more varied set of SSG speakers and how it is affected by speech rate. Following [6], we predicted the boundary between long and short vowels to occur at a shorter duration in production than in perception (H1); following [5], we expected a significant effect of speaking rate on the perceptual 50% cross-over boundary (COB) (H2). We further predicted the rate effect to be smaller in production than in perception (H3), as shorter segments should be affected less by variation than longer segments in production but not necessarily in perception.

39 diglossic speakers from Zurich and Lucerne participated in both a production and a subsequent perception experiment. Speakers produced in SSG and in random order various minimal pair words, each embedded in a carrier sentence. The set of sentences was repeated ten times, alternating between a normal and a fast speaking rate. The current production analysis focuses on 780 realizations of the minimal pair words *bitter* /ˈbɪ̥tʰɐ/ 'bitter' and *Bieter* /ˈbi̥ːtʰɐ/ 'tenderer' as they most directly match the perceptual judgments of the same participants to 11 stimuli from a *Bieter-bitter* continuum varying only in the duration of the first vowel embedded in a normal and a fast speech carrier phrase. For each participant, we calculated the optimal category boundary (OCB) between [iː] and [i] in production (see [2]) and the 50% COB between [iː] and [i] in perception using the quickpsy package in R [10].

Regarding H1, 23 out of 39 listeners behaved as predicted in normal tempo, but only 17 in fast tempo, suggesting that the use of durational cues in production and perception varies as a function of speaker and tempo. Commensurate with Figure 1 and as predicted by H2, results from a linear mixed-effects model showed a significant main effect of *rate* on both the OCB and the 50 % COB ($F=133.19$, $p<.001$). The model also confirmed H3, revealing a significant interaction between *rate* and *type* (production vs. perception) ($F=5.43$, $p=.022$), with a larger effect on perception than on production (cf. Fig. 1).

Taken together the results suggest an overall greater compensation effect in perception for rate-induced variation in production at the group-level, although the diminished variation in production is predictable by the target sound's underlying length. The findings thus add to the accumulating evidence by which production and perception are not always aligned [11].

doi:10.11576/pundp2022-1058

Figures

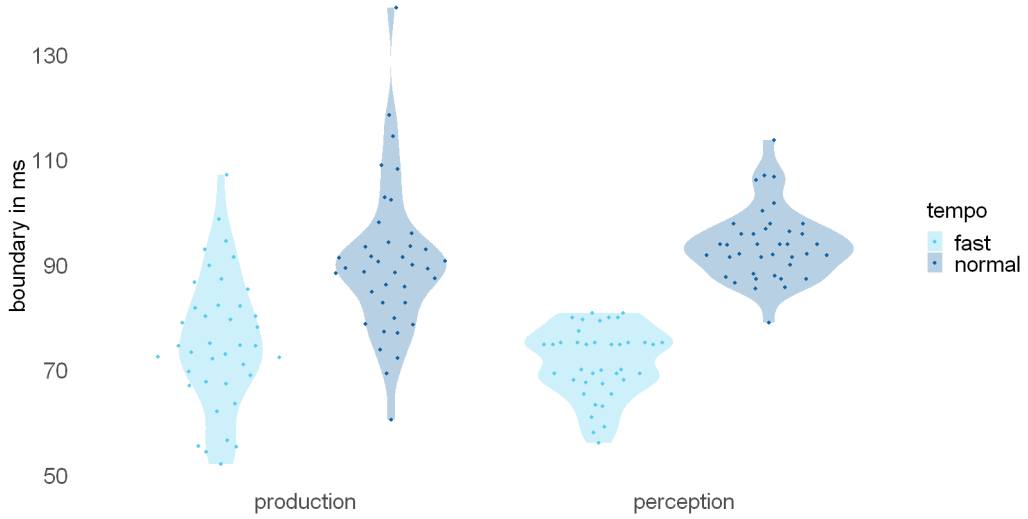

Figure 1. *Boundaries in ms between bitter and Bieter in fast (light blue) and normal (dark blue) speech tempo in production (left) and perception (right).*

[1] Pickett, E. R., Blumstein, S. E. & Burton, M. W. Effects of Speaking Rate on the Singleton/Geminate Consonant Contrast in Italian. *Phonetica*, 56(3–4), 135–157, 1999. doi: 10.1159/000028448.

[2] Mitterer, H. The singleton-geminate distinction can be rate dependent: Evidence from Maltese. *Journal of the Association for Laboratory PhonologyLaboratory Phonology*, 9(6), 2018. doi: https://doi.org/10.5334/labphon.66.

[3] Reinisch, E., Jesse, A., & McQueen, J. M. Speaking rate from proximal and distal contexts is used during word segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, 37(3), 978–996, 2011.

[4] Port, R. F. The influence of tempo on stop closure duration as a cue for voicing and place. *Journal of Phonetics*, 7(1), 45–56, doi: 10.1016/S0095-4470(19)31032-0, 1979.

[5] Jochim, M., Kleber, F., Klingler, N., Pucher, M., Schmid, S., & Zihlmann, U. Measuring the Role of Hypoarticulation in a Sound Change in Progress in Southern German. *14. Tagung Phonetik und Phonologie im deutschsprachigen Raum, Wien*, 2018.

[6] Klingler N., Kleber, F. Jochim, M., Pucher, M., Schmid, S. & Zihlmann, U. Temporal organization of vowel plus stop sequences in production and perception: Evidence from the three major varieties of German. *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia*, 825–829, 2019.

[7] Arvaniti, A. Effects of speaking rate on the timing of single and geminate sonorants. *Proceedings of the 14th International Congress of Phonetics Sciences, Berkeley, University of California 1*, 595–598, 1999.

[8] Klatt, D. H. Interaction between two factors that influence vowel duration. *The Journal of the Acoustical Society of America*, 54(4), 1102–1104, 1973. doi: 10.1121/1.1914322.

[9] Kraehenmann, A. Quantity and prosodic asymmetries in Alemannic. Berlin, New York: Mouton de Gruyter, 2003.

[10] Linares D & López-Moliner J. quickpsy: An R Package to Fit Psychometric Functions for Multiple Groups. The R Journal, 8(1), 122–131, 2016.

[11] Kleber, F., Harrington, J. & Reubold, U. The relation between perception and production of coarticulation during a sound change in progress. *Language and Speech*, 55, 383–405, 2012.

doi:10.11576/pundp2022-1058