Speech sound errors: Categorical or gradient?

Hugo Quené
Utrecht University

Speech sound errors provide valuable insights into how speakers mentally prepare their speech. Speakers detect and repair some (but not all) errors in their own speech. This detection may happen relatively early, or even immediately (e.g. *v--horizontal*), or it may happen relatively late (*go from left again... eh, from pink again to blue*). The duration of the error-to-interruption interval follows a bimodal distribution, with the peaks separated by about 350 ms (Nooteboom & Quené, 2013). This suggests two different detection processes, and hence two different stages of mental preparation of speech. Presumably, the first (early) detection taps into abstract phonological processing, whereas the second (late) detection taps into articulatory preparation and/or execution. Hence, it is hypothesized that early-detected errors tend to be categorical, that is, without a phonetic trace of the target sound, whereas late-detected and undetected errors tend to be gradient, that is, with articulatory blending resulting in a phonetic trace of the intended target into the produced sound.

This hypothesis was investigated using data from two sources: (1) elicited Dutch sound errors from several experiments involving phonological priming (Quené & Nooteboom, 2024), as well as (2) a sample of spontaneous English sound errors from the Simon Fraser University Speech Error Database (SFUSED, based on English podcasts; Alderete, 2025; Alderete et al., 2021). For both corpora, consonants were taken from correct controls, from early-detected errors, from late-detected errors (Dutch only), and from undetected errors.

The Dutch errors were studied perceptually *via* a consonant identification task using human listeners; the results of this perceptual study confirm the main hypothesis (Nooteboom & Quené, 2013). The English errors were studied acoustically by means of measurements of the Voice Onset Time (VOT) of critical consonants; the results of this acoustic analysis are inconclusive. Although the VOT data are consistent with the main hypothesis, there were too few early-detected errors (n=13) in the sample to corroborate a null effect, and the trace effect in the undetected errors (n=73) was relatively small. An additional sample of 242 English speech errors from the SFUSED also failed to corroborate the predicted pattern in VOT.

In conclusion, the combined results suggest that speech sound errors that are detected early seem to be categorical, without any phonetic trace of the intended speech sound, whereas sound errors detected late seem to be more gradient. The evidence is clear in the Dutch elicited errors, and less convincing for English spontaneous speech errors. These findings suggest that speakers have different mental representations of their speech plan in the earlier vs. later stages of mental preparation of speech.

References

Alderete, J. (2025). Simon Fraser University Speech Error Database – English (SFUSED English): Methods and Design. *Journal of Cognition*, *8*(1), 28. https://doi.org/10.5334/joc.440

Alderete, J., Baese-Berk, M., Leung, K., & Goldrick, M. (2021). Cascading activation in phonological planning and articulation: Evidence from spontaneous speech errors. *Cognition*, *210*, 104577. https://doi.org/10.1016/j.cognition.2020.104577

Nooteboom, S.G. & Quené, H. (2013). Parallels between self-monitoring for speech errors and identification of the misspoken segments. *Journal of Memory and Language*, 69 (3), 417–428. doi:10.1016/j.jml.2013.04.006, Corrigendum (2023): doi:10.1016/j.jml.2023.104448.

Quené, H., & Nooteboom, S.G. (2024). Some properties of mental speech preparation as revealed by self-monitoring. *Speech Communication*, 103043. https://doi.org/10.1016/j.specom.2024.103043

Acknowledgements

My sincere thanks are due to Sieb Nooteboom, for his support and guidance in our past work on speech errors, and for his critical comments on earlier versions of this presentation. In addition, I am very grateful to Matt Goldrick and John Alderete for inspiring discussions about this work, for sharing the full SFUSED corpus of speech errors, and for sharing their VOT data (Alderete et al., 2021).

References

Alderete, J. (2025). Simon Fraser University Speech Error Database – English (SFUSED English): Methods and Design. *Journal of Cognition*, 8(1), 28. https://doi.org/10.5334/joc.440

Alderete, J., Baese-Berk, M., Leung, K., & Goldrick, M. (2021). Cascading activation in phonological planning and articulation: Evidence from spontaneous speech errors. *Cognition*, 210, 104577. https://doi.org/10.1016/j.cognition.2020.104577

Nooteboom, S. G., & Quené, H. (2013). Parallels between self-monitoring for speech errors and identification of the misspoken segments. *Journal of Memory and Language*, 69(3), 417–428. https://doi.org/10.1016/j.jml.2013.04.006

Quené, H., & Nooteboom, S. (2024). Some properties of mental speech preparation as revealed by self-monitoring. *Speech Communication*, 103043. https://doi.org/10.1016/j.specom.2024.103043