**Supplementary Materials 3: Scoring details**

The following exceptions were made in the literal transcription of participants’ word productions:

1. Because of their omnipresence and productivity in Dutch (Shetter, 1959), adding a diminutive suffix to a noun was not regarded as insertion (e.g. *slabje* for *slab*, bib).
2. If a participant modified their production in such a way that it was clear that they were only expressing insecurity with regard to their utterance (example: *samba-iets*, “maraca-something”) and not an actual memory representation, this modification (*iets*, “something”) was not regarded as insertion.
3. Sometimes participants gave multiple productions for one word (e.g. “Is it a *stomp*? Or a *stolp*?” for *stolp*, “bell jar”). In these cases, the last production was transcribed.
4. If a participant mispronounced a phoneme just because of their German-accented Dutch, it was not marked incorrect (an accent does not reflect a false memory representation).
5. In Dutch, syllable-final obstruents get devoiced. Therefore, upon hearing the word /slɑp/ (*slab*, bib), one cannot know whether the true underlying form of the final consonant is /p/ or /b/. Therefore, productions such as /slɑpər/ or /slɑbər/ would receive the same score.

Now, we will further discuss the word length issue mentioned under *Scoring*. In the *ramlert* example, a consequence of using the long alignment is that the sum of correct and incorrect phonemes amounts to 8. However, for participants who did not produce any insertions, the total number of phonemes would be 7 (i.e., equal to the word length). In 3.6% of data points, the sum of the number of scored phonemes was larger than the total word length. This is problematic, as the binomial probability distribution for a particular word is characterised by a fixed parameter *N* for the number of trials (i.e., the number of phonemes), which should not vary over participants. We resolved this issue by rescaling the number of correct and incorrect phonemes, so that they would always add up to the (fixed) word length of the target word, in this case 7. Rescaling was done by multiplying the word length of the target word by the percentage correct (e.g. 7 \* 0.63 = 4.38, rounded off as 4), and subtracting this number from the total number of phonemes to arrive at the rescaled number of incorrect phonemes (7 – 4 = 3). Thus, the final vector for *ramlert* would be (4,3).

**Reference**

Shetter, W. Z. (1959). The Dutch diminutive. *The Journal of English and Germanic Philology, 58*, pp. 75–90.