

Weighing auditory, visual and semantic cues in lexical stress perception in Spanish

Floris Cos^{1,2}, Lieke van Maastricht^{1,2}, Hans Rutger Bosker², Matteo Maran² & Esther Janse^{1,2}

¹Centre for Language Studies, Radboud University, Nijmegen, ²Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen

Simple up-and-down hand movements known as beat gestures influence lexical stress perception: with the same ambiguous acoustic item, listeners may perceive either *CONtent* or *conTENT*, depending on which syllable has a beat gesture aligned to it. This “Manual McGurk effect” has been demonstrated using isolated words in Dutch (Bujok et al., 2025) and notably in Spanish (Rohrer et al., 2025), where lexical stress distinguishes between 1st person present (*repreSENto*, “I represent”), and 3rd person past tense verb forms (*representÓ*, “(s)he represented”). The current study investigates how listeners weigh auditory and visual cues to lexical stress in a full, tense-biasing sentence context:

- (1) *Anteriormente, *repreSENto / representÓ a la ciudad de Málaga.*
“Before, *I represent / (s)he represented the city of Málaga.”

In our stimuli, we combined 3 biasing contexts (past tense vs. present tense vs. no bias), 5 auditory versions of the verb form, following an acoustic lexical stress continuum, and 2 visual beat conditions, timed to one of the last two syllables of the verb form. Ninety Spanish natives (42F, 46M, 2 Other; $M_{\text{Age}} = 28.69$, range = 21–40) indicated for each stimulus which pronoun (*yo* “I” or *ella* “she”) they thought fit best with the sentence.

The results indicated that the proportion of *yo*-responses decreased as the audio became more *representÓ*-like. Compared to the neutral sentence context, the proportion of *yo*-responses increased in present-bias contexts, and decreased in past-bias contexts (Figure 1). Although beat alignment did not influence categorization responses, reaction-time analyses showed that participants responded faster whenever the beat gesture was congruent with the semantic bias (Figure 2). The results thus suggest that in full sentences, beat gestures do not change the ultimate interpretation of the utterance, but congruent (vs. incongruent) gestural timing does facilitate word recognition.

Figure 1.

Proportions of “yo”-responses as a function of Audio Step, Sentence Context and Beat Alignment.

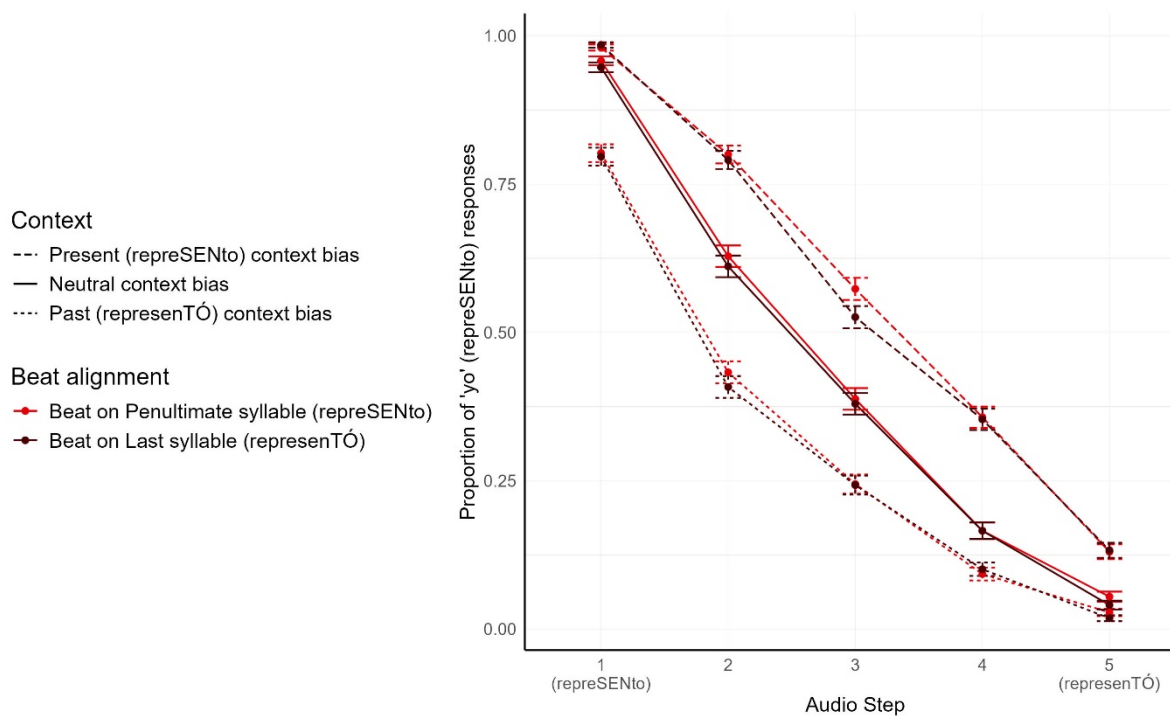
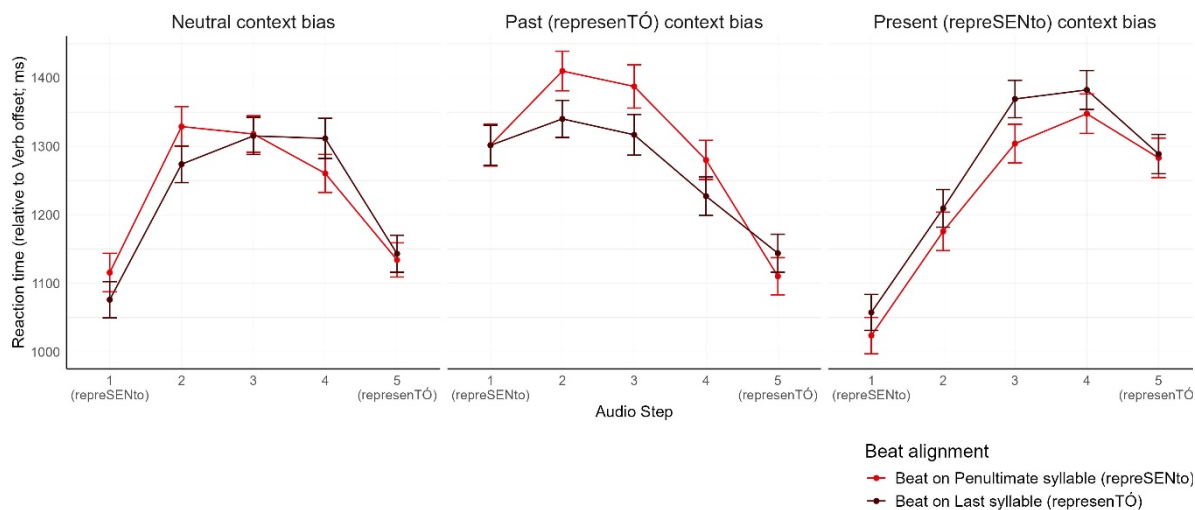


Figure 2.

Reaction times (RTs) for each step of the acoustic continuum and Beat Alignment, separated for each Sentence Context. RTs were measured from the offset of the verb form.



References

- Bujok, R., Meyer, A. S., & Bosker, H. R. (2025). Audiovisual perception of lexical stress: Beat gestures and articulatory cues. *Language and Speech*, 68(1), 181-203. <https://doi.org/10.1177/00238309241258162>
- Rohrer, P. L., Bujok, R., Van Maastricht, L., & Bosker, H. R. (2025). From “I dance” to “she danced” with a flick of the hands: Audiovisual stress perception in Spanish. *Psychonomic Bulletin & Review*, 1-10. <https://doi.org/10.3758/s13423-025-02683-9>