Multimodal prominence in head gestures (or how accent and phrasing influence the gesture timing)

Núria Esteve-Gibert[1], Joan Borràs-Comes[2], Eli Asor[2], Marc Swerts[3], Pilar Prieto[4,2]

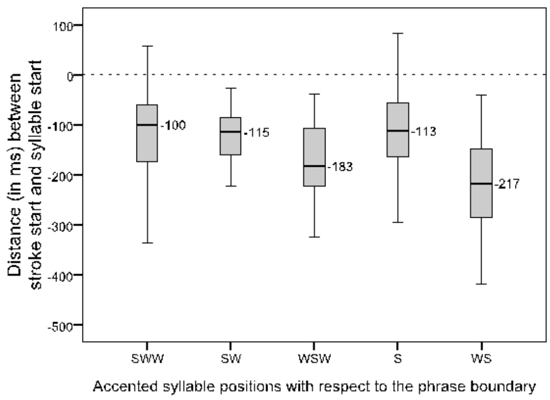
[1]Aix Marseille Université, CNRS, LPL UMR 7309, Aix-en-Provence, France, [2]Universitat Pompeu Fabra, Barcelona, Spain, [3]Tilburg University, Tilburg, The Netherlands, [4]ICREA

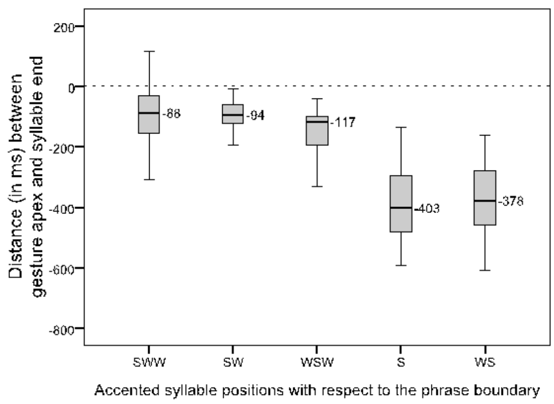
In the last decades, researchers have shown that gesture and speech are tightly temporally coordinated. Most gestures contain a phase of prominence (the interval of gesture stroke or the peak of gesture apex) that has been found to correlate with accented syllables (Kendon, 1980; McNeill, 1992). These correlations have been reported for manual gestures (Esteve-Gibert & Prieto, 2013; Guellaï et al., 2014; Leonard & Cummins, 2011) and articulatory movements (Rochet-Capellan et al., 2008; Turk & Shattuck-Hufnagel, 2014), but not for head gestures. Head gestures have been reported to correlate with information structure, since speakers tend to produce them associated with focal elements (e.g. Kim et al., 2014) and listeners perceive an element as focalized if it is accompanied by a head movement (Krahmer & Swerts, 2007).

This study investigates the role of prosodic structure on the timing of head gestures, focusing on the potential effects of metrical patterns (S, SW, WS, WSW, and SWW) and phrase boundaries. In Experiment 1 we analyzed the timing spontaneous head gestures co-occurring with words produced in a focus position (N=114), and in Experiment 2 we obtained head gestures from a semi-controlled setting to control for the speakers' communicative intent and to be able to analyze the same amount of cases per metrical pattern (N=155). We propose that the timing of gesture movements parallels that of rising-falling pitch movements in phrase-final positions, where the peak of prominence in both modalities is retracted if there is an upcoming phrase boundary (e.g., Prieto & Ortega-Llebaria, 2009).

Results from Exp1 showed that head movements started some milliseconds after the target word and they finished some milliseconds before the target word ended, independently of the metrical pattern of the target word (respectively, $F(3, 113)=.821$, $p=.485$; $F(3, 113)=.703$, $p=.552$). Metrical patterns did influence, though, the position of the gesture apex ($F(2, 111) = 5.72$, $p = .004$). Post-hoc tests revealed that gesture apexes were retracted in words with final stress. Results from Exp2 confirmed that the scope of the head movement was the focalized word, because gestures started before the target word and ended before the target word ended, independently of the metrical pattern (respectively, $F(4, 150) = 1.644$, $p = .166$ and $F(4, 150) = 1.995$, $p = .098$). Also, that metrical patterns influenced the position of the gesture apex ($F(4, 150) = 72.186$, $p < .001$). Finally, we found that metrical patterns effected the beginning of the gesture prominence ($F(4, 150) = 11.577$, $p < .001$). Pair-wise comparisons showed that gesture prominence began earlier when pre-accented material was available, and that gesture apexes occurred before if there was no post-accented material (Figures 1 and 2).

In conclusion, this study shows that: (1) speakers produce head gestures to accompanying focalized words; (2) the timing of gesture prominence depends on the timing of the prosodic prominence; and (3) the left-hand phrase boundary decides the anchoring point for the start of the gesture prominence while the right-hand phrase boundary decides the anchoring point for the gesture apex.





Figures 1 and 2. In experiment 2, the effect of metrical patterns and phrase boundaries on the distinct measures analyzed.

References

Esteve-Gibert, N., & Prieto, P. (2013). Prosodic structure shapes the temporal realization of intonation and manual Gesture Movements. *Journal of Speech, Language, and Hearing Research 56*(3), 850–865.

Guellaï, B., Langus, A., & Nespor, M. (2014). Prosody in the hands of the speaker. *Frontiers in Psychology*, *5*, 1–8.

Kendon, A. (1980). Gesticulation and speech: two aspects of the process of utterance. In M. R. Key (Ed.), *The Relationship of Verbal and Nonverbal Communication* (pp. 207–227). The Hague: Mouton.

Kim, J., Cvejic, E., & Davis, C. (2014). Tracking eyebrows and head gestures associated with spoken prosody. *Speech Communication*, *57*, 317–330.

Krahmer, E., & Swerts, M. (2007). The effects of visual beats on prosodic prominence: Acoustic analyses, auditory perception and visual perception. *Journal of Memory and Language*, *57*(3), 396–414.

Leonard, T., & Cummins, F. (2011). The temporal relation between beat gestures and speech. *Language and Cognitive Processes*, *26*(10), 1457-1471.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.

Prieto, P. & Ortega-Llebaria, M. (2009). Do contour tones induce syllable lengthening in Catalan and Spanish.” In M. Vigário, S. Frota and M.J. Freitas. (Eds.). Interactions in Phonetics and Phonology, John Benjamins: Amsterdam/Philadelphia.

Rochet-Capellan, A., Laboissière, R., Galván, A., & Schwartz, J. (2008). The Speech Focus Position Effect on Jaw-Finger Coordination in a Pointing Task. *Journal of Speech, Language, and Hearing Research, 51*(6), 1507–1521.

Turk, A., & Shattuck-Hufnagel, S. (2014). Timing in talking: what is it used for, and how is it controlled? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *369*, 20130395.