The effect of speech planning and prosodic structure on kinematic properties of gesture

Jungyun Seo[1], Ruaridh Purse[1] & Jelena Krivokapić[1]

[1] *University of Michigan*
sjungyun@umich.edu, rupurse@umich.edu, jelenak@umich.edu

Previous research has established that pauses are related to speech planning, in that they provide time for speakers to plan the upcoming utterance [1-4]. Pauses are also markers of prosodic boundaries [1,4,5]. It is known that these two sources of pauses are difficult to distinguish, but only a few studies have tried to tease them apart [1,4,6]. Ferreira (1991) observes that there is no final lengthening before pauses that are inserted to facilitate planning of upcoming material and points out that this is different from the pattern at typical prosodic boundaries. She suggests [1,4] the possibility that there are two different types of pauses: structural and planning pauses, with planning pauses not being part of prosodic structure. The first question of the present study examines the effect of planning on phrase boundaries, specifically testing whether planning changes the temporal properties of prosodic boundaries. The second question examines the effect of planning on word boundaries, testing whether, when speakers need more planning time, a planning pause is inserted (as suggested in [1,4]) or prosodic boundaries are inserted, as assumed in Levelt's (1989) speech production model [7]. The answers to these questions will inform us about whether and how the effect of speech planning differs at different types of boundaries (word boundaries and prosodic boundaries), about the role of prosodic boundaries in speech planning, and about the nature of prosodic representation.

Articulatory kinematic data of 7 monolingual American English speakers, each producing 270 sentences, were recorded using electromagnetic articulometry (EMA). To test the effect of Boundary (phrase, word boundaries) and Planning (planning, no planning), the target word (Mima [mimə], Biba [bibə]) occurred in IP-final and IP-initial (Table 1-1) or IP-medial positions (Table 1-2), testing the effect of boundary. In the no planning condition, participants simply read orthographically presented sentences. To induce planning at the boundary in the planning condition, sentences included a blank space and were presented alongside one picture. Participants were instructed to start speaking as soon as they saw the sentence, which triggered a second picture to appear. Participants could then fill in the blank by comparing the two pictures, producing, e.g., "Mima" if the pictures were the same, and "Biba" if they differed. Lip closing and opening duration was measured for the three bilabial consonants surrounding the boundary [mimə, mi/bi] (C1, C2, C3), using a semi-automatic labeling procedure (mview; Tiede, Haskins Laboratories). Linear Mixed-effect Models and pairwise comparisons for different conditions were conducted on measured gesture and pause (identified from the kinematic signal) durations.

The results show similar scope of boundary-related lengthening in two boundaries (prosodic boundaries and prosodic boundaries with planning), shown in Fig. 1a and 1b, indicating that planning does not change the scope of the existing boundaries. However, pause duration is longer in prosodic boundaries with planning compared to prosodic boundaries, suggesting that speakers take additional time in pauses to accommodate an increased planning load. When word boundary conditions are compared, word boundaries *with* planning are longer in C2-opening, pause and C3-closing (Fig. 2a)—showing an identical scope to the boundary-driven lengthening in prosodic boundaries (Fig. 1a). Furthermore, there is no difference, in lengthening or pause duration, between word boundaries with planning and prosodic boundaries with planning. These results indicate that speakers use prosodic boundaries for planning, either by extending the pause duration of existing boundaries or inserting new

boundaries to accommodate planning time. Implications for models of speech production and for the larger question of what constitutes a prosodic boundary are discussed.

Table 1. Experiment stimuli, a subset. The full experiment included one more sentence type for the prosodic boundary condition. Target words are presented in bold font (they were not bolded in the experiment). The boundary under consideration always occurs after the (first) "Mima". Prompt sentences are given in italics above the sentence containing the target word. The number in square brackets indicates the number of repetitions.

	(3) No planning condition	**(4) Planning condition**
Target word position	**(a) Prosodic boundary** (structural boundary)	**(b) Prosodic boundary with planning** (planning boundary)
(1) Prosodic boundary condition	*What would you like?* I want a **Mima**, **Biba**'s mom, and a cat. [15] I want a **Mima**, **Mima**'s mom, and a cat. [15]	*What would you like?* I want a **Mima**, ____'s mom, and a cat. [60] (**Mima** or **Biba**)
	(c) Word boundary	**(d) Word boundary with planning**
(2) Word boundary condition	I want a **Mima meeting** a banana. [15] I want a **Mima beating** a banana. [15]	I want a **Mima** _____ a banana. [60] (**meeting** or **beating**)

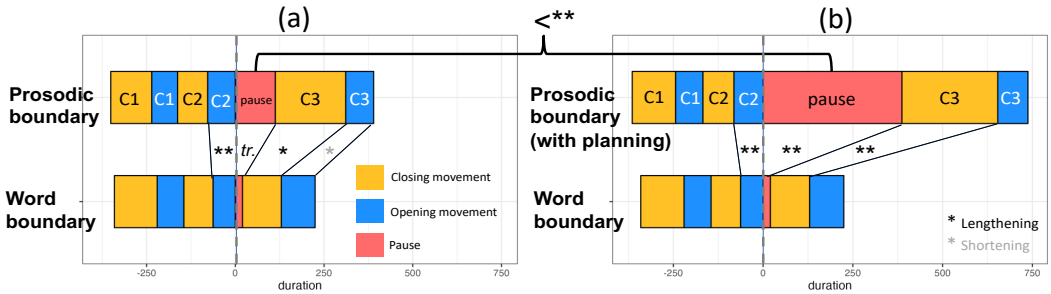

Fig. 1. Results of pairwise comparison between Boundary and Planning conditions. (a) Comparison between prosodic boundary and word boundary conditions, (b) Comparison between prosodic boundaries with planning and word boundaries. The line crossing (a) and (b) indicates the comparison between prosodic boundaries and prosodic boundaries with planning. Results for the second sentence type (not shown in the stimuli) are almost identical.

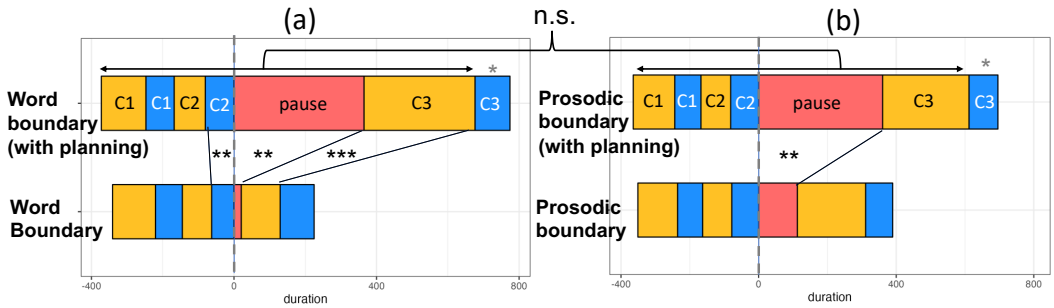

Fig. 2. Results of pairwise comparison between Boundary and Planning conditions. (a) Comparison between word boundaries with planning and word boundaries, (b) Comparison between prosodic boundaries with planning and prosodic boundaries. The line crossing (a) and (b) indicates the comparison between word boundaries with planning and prosodic boundaries with planning. Results for the second sentence type are almost identical.

References.

[1] Ferreira, F. (2007). Prosody and performance in language production. *Language and Cognitive Processes*, 22(8), 1151-1177. [2] Fuchs et al. (2013). Acoustic and respiratory evidence for utterance planning in German. *JPhon*, *41*(1), 29-47. [3] Kentner et al. (2023). The final lengthening of pre-boundary syllables turns into final shortening as boundary strength levels increase. *JPhon*, *97*, 101225. [4] Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *J. of Memory and Language*, 30(2), 210-233. [5] Krivokapić et al. (2020). Pause postures: The relationship between articulation and cognitive processes during pauses. *JPhon*, *79*, 100953. [6] Ramanarayanan et al. (2009). Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. *JASA*, *126*(5), EL160-EL165. [7] Levelt, W. J. (1989). *Speaking: From intention to articulation*. MIT press.