The effect of speech planning and prosodic structure in speech production: How do planning and prosodic structure interact?

Jungyun Seo[1], Ruaridh Purse[1] & Jelena Krivokapić[1] ([1]*University of Michigan)*
sjungyun@umich.edu, rupurse@umich.edu, jelenak@umich.edu

The present study examines the interaction between speech planning and prosodic structure in order to understand the role of prosodic structure in speech planning. Pause durations are known to be affected by the amount of planning required for upcoming material. For example, pause duration increases for long or syntactically complex material [1-2, 4-7]. On the other hand, pauses are also salient temporal phonetic markers of prosodic boundaries, along with final and initial lengthening [1,2,8]. The relationship between these two different sources of pauses (structural, planning) has not been systematically examined [1,2,3], but there are two hypotheses: Ferreira [1,2] proposes that there are two different types of pauses. Structural pauses are part of prosodic boundaries, and as such are accompanied by final lengthening and occur based on the prosodic structure of the utterance. Planning pauses are independent of prosodic structure, occur to provide planning time for an upcoming unit, and are not accompanied by final lengthening. A second hypothesis, from Levelt [10], suggests that speakers insert prosodic boundaries when they need more time to plan, in which case pauses from different sources (structural, planning) cannot be distinguished. The two hypotheses will be tested by examining the effect of planning on different boundaries: (1) prosodic boundaries, (2) word boundaries. The results will inform us whether and how the effect of speech planning differs at different types of boundaries and whether and how speakers modify prosodic structure in planning (i.e., whether speakers insert a prosodic boundary when they need more planning time).

Articulatory kinematic data from seven monolingual American English speakers were collected using Electromagnetic Articulometry (EMA), with sensors for movement tracking attached to the main articulators. To test the effect of the Boundary (word, prosodic IP boundary), the target words “Mima/Biba” [mimə, bibə], “meeting/beating” were embedded in carrier sentences, forming a sequence of three bilabial consonants, with an IP boundary (Table 1a) or word boundary (Table 1b) after the first “Mima”. To test the effect of Planning, conditions varied in the amount of planning load (easy, difficult). In the easy condition (Table 1c), speakers read orthographically presented sentences. In the difficult planning condition (Table 1d), sentences were presented with a blank space alongside one picture. Participants were instructed to start speaking as soon as they saw the sentence, and a second picture automatically appeared in reaction to their voice. Participants inserted “Mima/meeting” if the pictures were the same, and “Biba/beating” if the pictures differed, in this way inducing planning at the targeted boundary (after the first “Mima”). Lip closing and opening durations (defined as the Euclidean distance between two sensors attached on the upper and lower lip) were measured on the three bilabial gestures [mimə, mi/bi] (Fig. 1). Linear Mixed Effect Models and corrected pairwise comparisons were conducted on the measured durations.

The results show a similar scope of boundary-related lengthening in the prosodic boundary conditions (easy and difficult planning), spanning the region of C2 opening to C3 closing (Fig. 2a and 2b). This indicates that planning does not extend the scope of the existing boundaries further into the phrase. However, pause duration is longer in prosodic boundaries with difficult planning compared to prosodic boundaries with easy planning, suggesting that speakers use additional time in pauses to accommodate an increased planning load. Word boundaries with difficult planning (compared to word boundaries with easy planning) are longer in C2-opening, pause and C3-closing (Fig. 3a)—showing an identical scope to the boundary-driven lengthening in prosodic boundaries (Fig. 2a). Furthermore, there is no difference between word boundaries (difficult planning) and prosodic boundaries (difficult planning). These results indicate that planning pauses are prosodic boundaries, and that speakers use prosodic boundaries for planning, either by extending pause duration of existing boundaries or by inserting new boundaries to accommodate planning time. Implications for speech production models are discussed.

Table 1. A subset of experiment stimuli. The full experiment included one more sentence type for the prosodic boundary condition and is not presented due to space limitations (but the results are almost identical to the ones presented). Target words are bolded only for the presentation purpose, but they were not bolded in the experiment. The boundary of interest always occurs after the (first) "Mima". To elicit natural production of the stimuli, prompt questions were given (shown in italics). The number of repetitions for each sentence are presented with the number in square brackets. 270 sentences were recorded.

	(c) Easy planning condition	**(d) Difficult planning condition**
Target word position	**(1) Prosodic boundary with easy planning** (structural boundary)	**(2) Prosodic boundary with difficult planning** (planning boundary)
(a) Prosodic boundary condition	*What would you like?* I want a **Mima**, **Mima**'s mom, and a cat. [15] I want a **Mima**, **Biba**'s mom, and a cat. [15]	*What would you like?* I want a **Mima**, ____'s mom, and a cat. [60] (**Mima** or **Biba**)
	(3) Word boundary with easy planning	**(4) Word boundary with difficult planning**
(b) 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**)

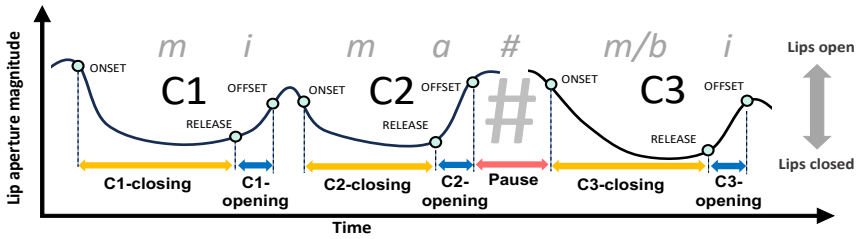

Fig 1. Schematic representation of the lip movement and measurement. Data were labeled by using a semi-automatic labeling procedure (mview: Tiede, Haskins Laboratories). C1, C2, C3 refers to the labeled consonants.

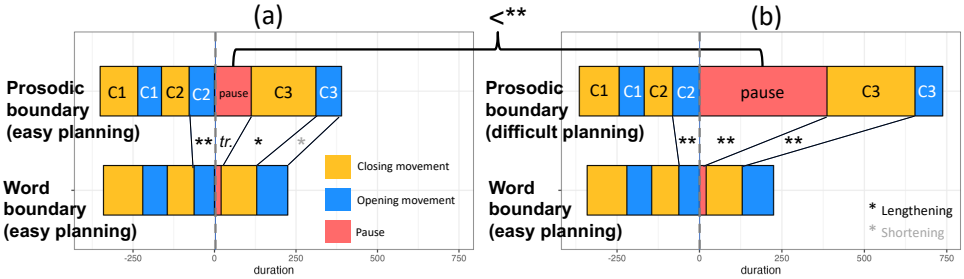

Fig 2. (a) Comparison between the prosodic boundary with easy planning and word boundary with easy planning, (b) Comparison between prosodic boundaries with difficult planning and word boundaries with easy planning. The line crossing (a) and (b) indicates the comparison between prosodic boundaries with easy/difficult planning.

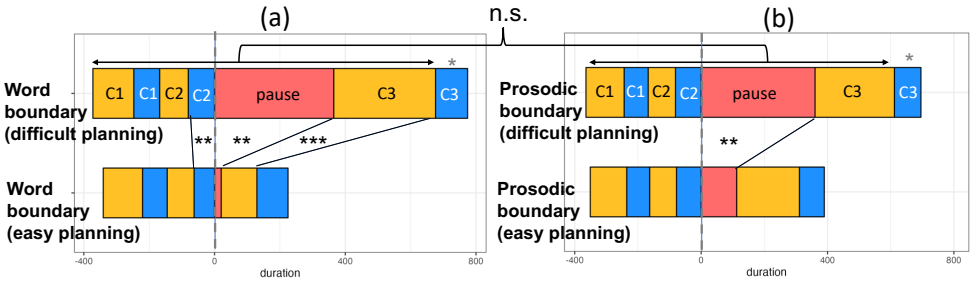

Fig 3. Results of pairwise comparison between Boundary and Planning conditions.

[1] Ferreira (2007). Prosody and performance in language production. *Lang Cognitive Proc*, 22(8). [2] Ferreira (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *JML*, 30(2). [3] Ramanarayanan et al. (2009). Analysis of pausing behavior in spontaneous speech using real-time magnetic resonance imaging of articulation. *JASA,* 126(5). [4] Fuchs et al. (2013). Acoustic and respiratory evidence for utterance planning in German. *JPhon*, 41(1). [5] Krivokapić (2007). Prosodic planning: Effects of phrasal length and complexity on pause duration. *JPhon*, *35*(2). [6] Watson & Gibson (2004). The relationship between intonational phrasing and syntactic structure in language production. *Lang Cognitive Proc. 19*(6). [7] Kentner et al. (2023). The final lengthening of pre-boundary syllables turns into final shortening as boundary strength levels increase. *JPhon*, 97. [8] Krivokapić et al. (2020). Pause postures: The relationship between articulation and cognitive processes during pauses. *JPhon*, 79. [9] Shattuck-Hufnagel & Turk (1996). A prosody tutorial for investigators of auditory sentence processing. J. *of Psycholinguist. Res.*, 25(2). [10] Levelt (1989). *Speaking: From intention to articulation*. MIT press.