The Role of Narrative Coherence on Speech Motor Planning

Jeffrey E. Kallay[1], Ulrich Mayr[2], and Melissa A. Redford[1]
Departments of Linguistics[1] & Psychology[2], University of Oregon, Eugene OR 97403

Speech production studies have demonstrated a robust, positive correlation between inhalation depth and the length of a subsequent utterance (e.g., Winkworth et al., 1994; Mitchell et al., 1996; Whalen et al., 1997; Huber, 2008; Fuchs et al., 2013), which has been interpreted as evidence for extended lookahead in speech motor planning. This interpretation is problematic because the finding is based on read speech, which allows for the possibility that external cues to length drive breath intakes. This possibility is strengthened by the finding that, when we asked participants to memorize sentences in order to control for visual cues to utterance length, only the preceding utterance length predicts breath intake patterns (a recovery effect); the length of a subsequent utterance does not (Kallay et al., 2019).

Of course, the memorization paradigm introduces its own problems of interpretation. In particular, rote speech is atypical in that it is fairly monotonous (at least in our experiments). The monotonous prosody gives the impression that speakers are more focused on the form of what they are saying than on the content. This focus should not matter under the standard psycholinguistic assumption that speech motor planning references a phonetically-specified speech plan; but what if extended lookahead actually relies on access to conceptual information? In this case, a planning effect on breath intakes may only emerge when the to-be-conveyed information can be assessed as more or less dense in the context of an on-line production task. The results reported here are consistent with this hypothesis.

The current results are based on speech data collected from 40 undergraduate students who completed several elicitation tasks as part of an extended study on pausing (Kallay, in prep). Here, we focus specifically on spontaneous speech that was elicited to investigate effects of narrative coherence on pause frequency and duration. The materials used were sets of 5 wordless cartoon pictures, which were arranged to create individual storyboards with different levels of coherence. The high-coherence storyboards depicted a sequence of well-defined, causally-related sequential actions with a limited temporal scope (i.e., an event schema; see Schank & Abelson, 1977; Mandler, 1984). The moderate-coherence storyboards depicted several such schemas, strung together to create an "episode" (see Mandler & Johnson, 1977). The no-coherence storyboard depicted unrelated individual scenes performed by completely separate actors. Each participant produced a narrative in response to each storyboard given. In this way, a total of 240 very short narratives were elicited (3 conditions x 2 stories x 40 participants). Each was transcribed and hand-segmented into pause-delimited utterances. Pauses were then coded for the presence (1) or absence (0) of breath intakes based on perceptual and visible acoustic cues (see Kallay et al., 2019). These binary data were modeled using logistic regression. The predictor variables were narrative coherence (high, moderate, none), preceding utterance length (in duration and syllables), and following utterance length (in duration and syllables). A significant effect of preceding utterance length on the presence of breath intakes suggests that physiological recovery drives breath intakes; a significant effect of following utterance length suggests that breath intakes are planned to optimize speech production. The results indicated independent effects of both recovery [$z = 5.501$, $p < .001$] and planning [$z = 3.671$, $p < .001$], but only the planning effect varied with coherence (see Figure 1): narratives produced in the high coherence condition showed a significantly stronger effect of subsequent utterance length on breath intakes than those produced in the no coherence condition [$z = -2.721$, $p = .007$]. The difference between narratives in the moderate and no coherence condition was not significant.

Our interpretation of these results, given our previous results and skepticism about the read-speech results, is that speech motor planning at the level of the utterance references the density of to-be-conveyed conceptual information, where information density is a relative notion that can only be defined with reference to a discourse-level plan.

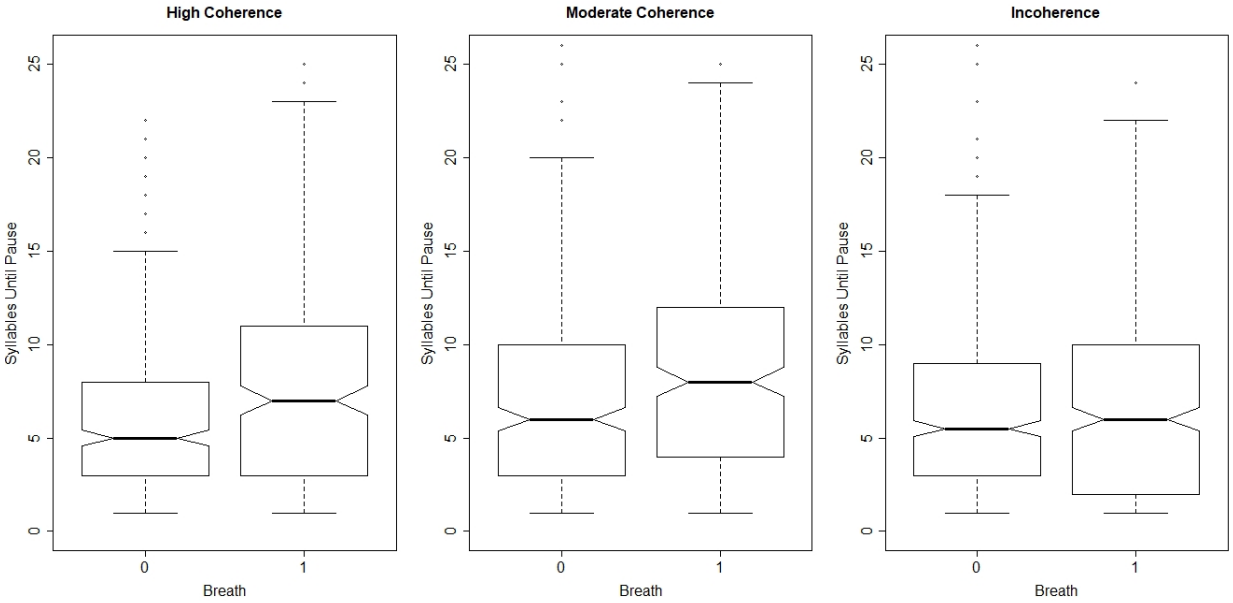

Figure 1. The presence (1) versus absence (0) of a breath intake during pause (x-axis) is shown as a function of the subsequent pause-delimited utterance length (y-axis) and narrative coherence condition (panels).

REFERENCES

Fuchs, S., Petrone, C., Krivokapić, J., & Hoole, P. (2013). Acoustic and respiratory evidence for utterance planning in German. *Journal of Phonetics*, 41(1), 29-47.

Huber, J. E. (2008). Effects of utterance length and vocal loudness on speech breathing in older adults. *Respiratory Physiology & Neurobiology*, 164(3), 323-330.

Kallay, J. (in prep). Modeling pausing patterns in adult narrative speech. PhD thesis, Univ. of Oregon.

Kallay, J., Mayr, U. & Redford, M.A. (2019). Characterizing the coordination of speech production and breathing. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (eds.) *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 1412-16). Canberra, Australia: Australasian Speech Science and Technology Association Inc.

Mandler, J. (1984). *Stories, scripts, and scenes: Aspects of a schema theory*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, 9(1), 111-151.

McFarland, D. H., & Smith, A. (1992). Effects of vocal task and respiratory phase on pre-phonatory chest wall movements. *Journal of Speech, Language, and Hearing Research*, 35(5), 971-982.

Schank, R. C., & Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding*. Hillsdale, NJ: Erlbaum.

Winkworth, A. L., Davis, P. J., Ellis, E., & Adams, R. D. (1994). Variability and consistency in speech breathing during reading: Lung volumes, speech intensity, and linguistic factors. *Journal of Speech, Language, and Hearing Research*, 37(3), 535-556.

Whalen, D.H., & Kinsell-Shaw, J.M. (1997). Exploring the relationship of inspiration duration to utterance duration. *Phonetica*, 54, 138-152.