The Role of Narrative Coherence on Speech Motor Planning for Breathing

Jeffrey E. Kallay, Ulrich Mayr, & Melissa A. Redford, *University of Oregon*

THE PROBLEM

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; 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.

STUDY HYPOTHESIS

Whereas the automatic coordination of language and breathing during speaking follows from language structure, controlled coordination references the conceptual information that is to be conveyed (i.e., language content); speech motor planning for breathing during a pause is thus more likely to be observed when the task promotes richer conceptual structure than when it promotes weaker conceptual structure.

METHODS

SPEAKERS were 40 healthy college-aged adult (29F;11M).

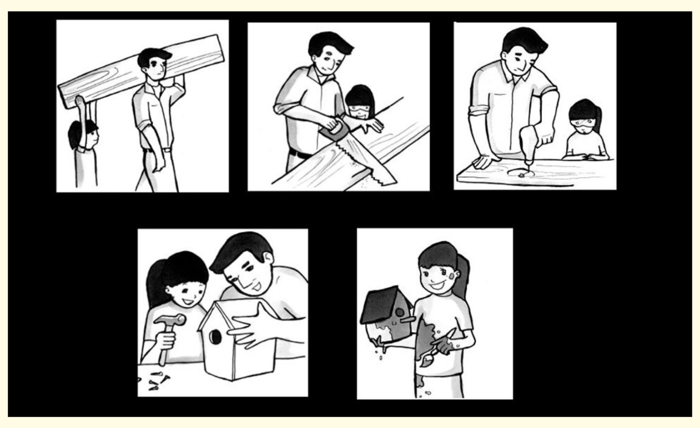
ELICITATION CONDITIONS (from Kallay, 2020):

1.Highly Coherent
- sequence of local actions within single event schema.

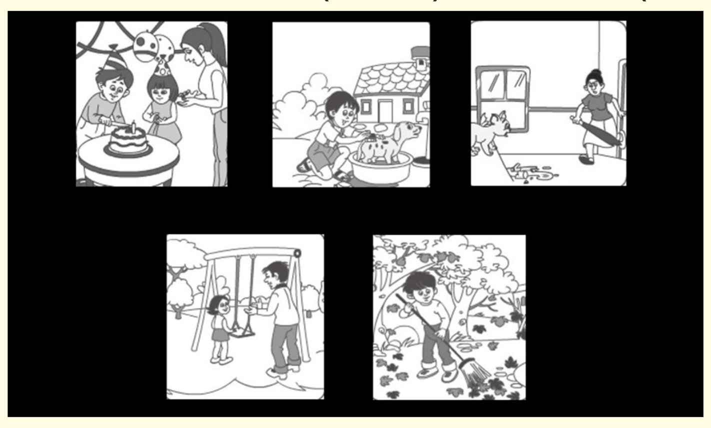
2.Moderately Coherent:
- sequence of event schemas within larger episode.

3.Incoherent:
- unrelated events performed by different actors.

See Schank & Abelson (1977), Mandler & Johnson (1977) Mandler (1985)



Example COHERENT sequence



Example INCOHERENT sequence

CODING:
- Narratives transcribed & segmented into pause-delimited utterances.
- Juncture strength coded on 6-point scale from ungrammatical (=0) to discourse boundary (=5)
- Breaths w/in pauses acoustically identified.

FULL LOGISTIC MIXED EFFECTS MODEL

Dependent Variable
- 0/1 = presence/absence of breath pause

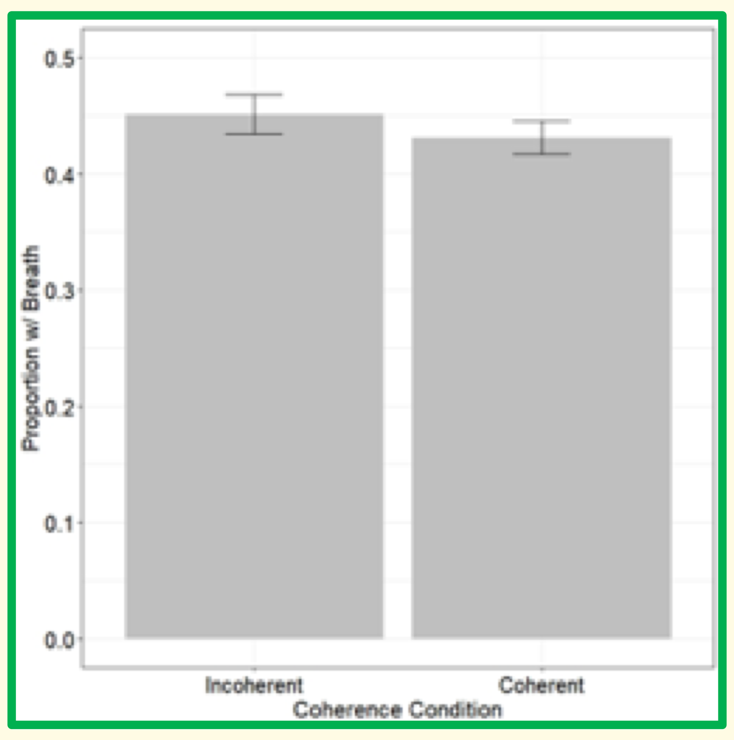
Predictor Variables:
1. Coherence condition (conceptual structure)
2. Juncture strength (linguistic structure)
3. Preceding utterance length (buffering)
4. Following utterance length (planning)
5. Time since breath intake (recovery)

SPEECH SAMPLE CHARACTERISTICS

	Coherent	Incoherent
Mean narrative duration (SD) in seconds	61.94 (39.09)	85.34 (49.16)
Mean *N* of words per narrative (SD)	45.79 (29.41)	62.20 (35.39)
Mean utterance duration (SD) in seconds	1.81 (1.23)	1.73 (1.16)
Total number of pauses	1211	882
Total number of breaths	522	398

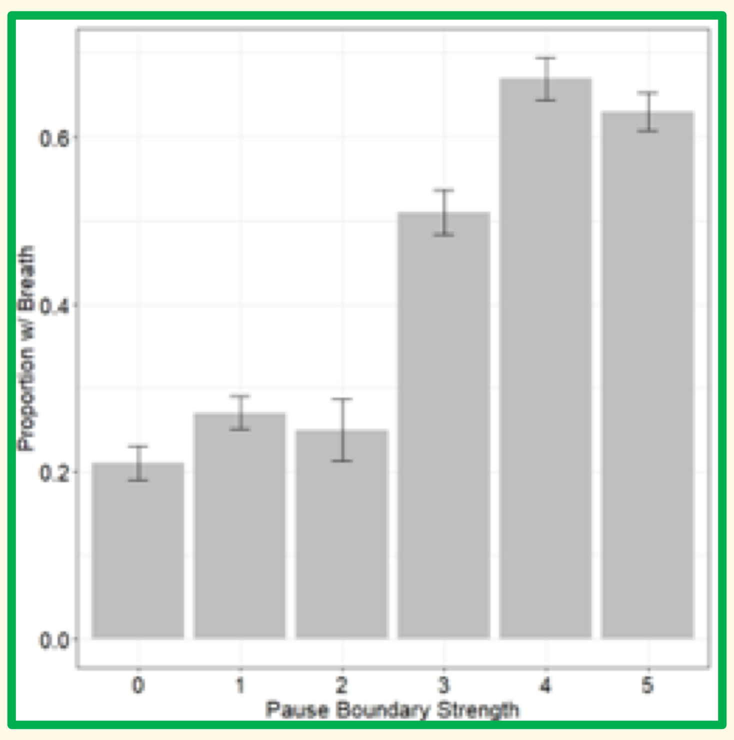
RESULTS

CONCEPTUAL STRUCTURE



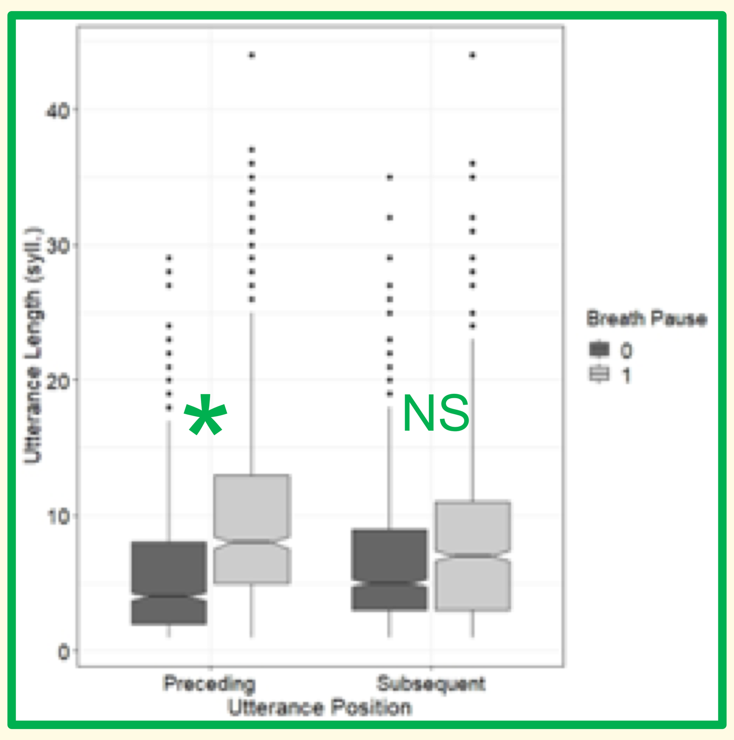
Probability of breath intake higher at pauses when narrative is unstructured than when structured.

LINGUISTIC STRUCTURE



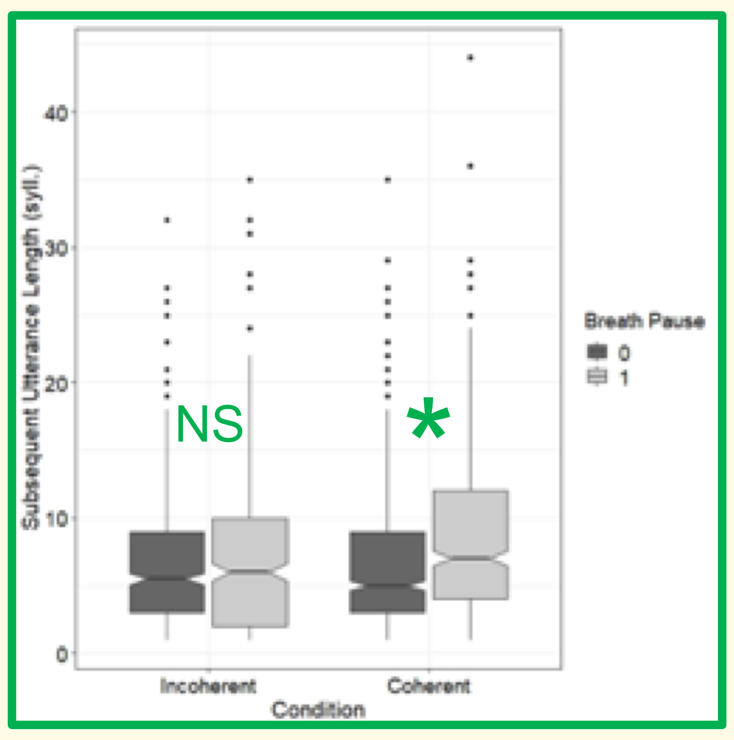
Probability of breath intake increases at pauses with grammatical juncture strength.

BUFFERING



Probability of breath intake increases at pauses with preceding utterance length.

PLANNING



Interaction between coherence cnd and utterance length: **CONCEPTUAL STRUCTURE affects PLANNING.**

INTERPRETATION

Given our previous results (Kallay et al., 2019) and skepticism about the read-speech results, our interpretation of the present results is that speech motor planning for breathing during a grammatical pause references the density of to-be-conveyed conceptual information, where information density is defined with reference to the conceptual structure of a discourse-level plan.

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. (2020). *Towards 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.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Institutes of Health (NIH), award #R01HD087452.

Contact: redford@uoregon.edu