

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Pre-boundary lengthening and pause signal boundaries in action sequences

Permalink

<https://escholarship.org/uc/item/5ks5h1h4>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors

Hilton, Matt
Wartenburger, Isabell
Verrel, Julius
et al.

Publication Date

2022

Peer reviewed

Pre-boundary lengthening and pause signal boundaries in action sequences

Matt Hilton (matt.hilton@uni-potsdam.de)

Department of Psychology, Cognitive Sciences, University of Potsdam, Germany

Isabell Wartenburger (isabell.wartenburger@uni-potsdam.de)

Department of Linguistics, Cognitive Sciences, University of Potsdam, Germany

Julius Verrel (julius.verrel@neuro.uni-luebeck.de)

Max Planck Institute for Human Development, Germany

Institute of Systems Motor Science, University of Lübeck, Germany

Birgit Elsner (birgit.elsner@uni-potsdam.de)

Department of Psychology, Cognitive Sciences, University of Potsdam, Germany

Abstract

In order to probe the production of kinematic cues to signal boundaries in action sequences, adults performed sequences of three actions on an object, with or without an action boundary following the second action. Movement of the hand was recorded via motion tracking, and it was found that the boundary was marked by a lengthening of the pre-boundary action and by a pause. These cues are also found in prosody to signal phrase boundaries in speech, suggesting a close coupling of the mechanisms underlying boundary production in both domains.

Keywords: action segmentation; kinematic boundary cues; motion tracking; domain-generalizability

Introduction

The structure of speech is fundamentally similar to that of action. Both information streams are characterized as sequences of hierarchically organized sub-units. In speech, sentences are formed from sequences of words and these words are in turn formed from sequences of syllables, in the same way that actions are formed from sequences of sub-units at different structural levels. For example, the action of opening a door involves sequences of coarser-grained sub-units (e.g., unlock the door, twist the handle, etc.) which in turn are formed by sequences of finer-grained units (e.g., insert the key, turn clockwise, etc.). In order to make sense of these streams of information, the listener or observer must therefore be able to determine the structure by segmenting the incoming stream into the relevant sub-units.

In the speech domain, it has been argued that the listener segments the speech stream into its constituent words and phrases by integrating knowledge-derived (top-down) and signal-derived (bottom-up) cues (Mattys, White, & Melhorn, 2005). Segmentation via top-down cues involves determining the location of boundaries between parts of speech according to lexical or syntactic knowledge. For example, if a listener has already learned the words *face* and *covering*, then this listener is likely to process these words when heard in sequence as a two-word phrase instead of an unfamiliar single word. Bottom-up cues, on the other hand, are characterized as perceptual cues embedded in the speech stream at boundaries between words and phrases. These bottom-up cues can be found in prosody (Wagner & Watson, 2010), and previous analyses

of intonational phrase boundaries have typically reported on three such prosodic boundary cues: pre-boundary lengthening, pre-boundary pitch rise, and pause (e.g., Peters, Kohler, & Wesener, 2005). Pre-boundary lengthening is characterized as an extension of the final word or syllable before the boundary, pre-boundary pitch rise is characterized by heightened pitch in that segment, and a pause is characterized as a silent interval following the offset of the final word or syllable before the boundary (see for instance, Huttenlauch, de Beer, Hanne, & Wartenburger, 2021). While Huttenlauch et al. found stable prosodic cues, independent of the intended listener, some studies have reported that speakers make particular use of prosodic cues to signal boundary location when they know that any top-down cues are less informative or uninterpretable by the listener, for example when speaking to those with limited experience of the target language, such as infants (e.g., Ludusan, Cristia, Martin, Mazuka, & Dupoux, 2016) or adult non-native language learners (for a review see Piazza, Martin, & Kalashnikova, 2021).

Segmentation in the action domain is also understood in terms of integrating top-down and bottom-up cues (Zacks, Speer, Swallow, Braver, & Reynolds, 2007). Action segmentation via top-down cues involves applying knowledge of intentions and action outcomes (Baldwin & Baird, 1999). For example, the observer knows from experience that the action sequence *tie your shoelaces* is complete when the laces have formed a tight bow, and any subsequent action (e.g., buffing the shoe) will be considered a separate component of the sequence. Alternately, if it is known that the actor intends to clean the window, the end of the first action will be determined as the point at which all dirt has been removed, separate from any subsequent actions (e.g., squeegeeing). There is also some evidence that bottom-up perceptual cues can signal the location of boundaries between actions in an action sequence, and these cues appear to be implemented by modulating properties of the movement that forms the actions. For example, Zacks, Kumar, Abrams, and Mehta (2009) found that when viewing videos of a human performing everyday action sequences (e.g., folding laundry), velocity changes in movement of the hands and head (rapid acceleration or de-

celeration) aligned with boundary-points between sub-units as identified by adult participants. It is thus possible that these movement-based (kinematic) changes signaled the location of boundaries to the participants. However, in contrast to work on prosodic boundary cues, work examining the nature and role of kinematic boundary cues in action segmentation is much less advanced. Given the many parallels between speech and action-processing, the current paper set out to test the prediction that the boundary cues realized in prosody bear fundamental similarities to kinematic boundary cues in action sequences. In principle, pre-boundary lengthening and pause could be realized identically in action as in speech: pre-boundary lengthening as an extension of the action prior to a boundary, and pause as an absence of movement at the point of the boundary. The third prosodic boundary cue, pre-boundary pitch rise, is a purely acoustic property, and it is therefore questionable how this cue could be present in the action domain. The current paper, however, examines whether velocity-related properties of the pre-boundary action offer a potential parallel to pitch-based boundary cues.

In order to probe the presence of kinematic boundary cues in action sequences, this paper adapts a paradigm that was designed to evoke and measure prosodic boundary cues in speech: Petrone et al. (2017) asked adult speakers to read aloud lists of names, and to communicate to an imagined naïve listener a specific grouping structure of the names, as notated by brackets. For example, when reading the list (*Lola or Mona*) and *Lena* aloud, the speakers were required to communicate that the final name should be considered in a separate structural group to the first two names. In other words, participants were required to indicate the presence of a boundary after the second name. This method successfully elicited prosodic boundary cues: when a word directly preceded the boundary, the duration of its final syllable was longer (pre-boundary lengthening) and its pitch rose higher (pre-boundary pitch rise) than when the same word did not precede a boundary. Participants also marked the position of the boundary in the list with a silent pause. This list-with-brackets procedure allowed for a fine-grained examination of the presence and interaction of prosodic boundary cues in naturally-produced speech. Another major strength of this paradigm is that by comparing acoustic properties of identical lists read with or without a boundary, any potential role of top-down cues is controlled for, allowing for a targeted analysis of bottom-up cues.

In order to examine whether actors produce similar bottom-up boundary cues in action sequences, we presented participants with written lists of verbs (e.g., *lift & shake & roll*), and asked them to perform these lists as a continuous action sequence, either with or without boundaries between individual actions as marked by brackets (e.g., [*lift & shake & roll*] vs. [*lift & shake*] [*& roll*]). To allow for precise measurement of the movements, we chose verbs describing actions that could be performed by a single hand on a single object (*lift, roll, shake, slide*). We used 3D motion-tracking technology to

record the movement of the hand as participants performed the action sequence. This procedure allowed us to compare kinematic features of the same action, either when this action preceded a boundary, or when it did not. If pre-boundary lengthening is present as a kinematic boundary cue, we would expect the duration of the action that precedes a boundary to be longer than the same action when it does not precede a boundary. Similarly, we hypothesized that the boundary cue of pause would be demonstrated as an increased time interval between the offset of the pre-boundary action and the onset of the post-boundary action, compared to when there was no boundary between the actions. Finally, we speculated that peak amplitude of the pre-boundary movement may potentially reflect pitch-rise, and so we also examined whether peak amplitude of the movement would be greater in an action prior to a boundary, compared to the same action not preceding a boundary.

Method

Participants

Participants were recruited from the student population at the University of Potsdam. No participant reported history of neurological disorders. Participants gave informed consent and were reimbursed for their participation. Data from 9 participants were included in the final data set (7 female/2 male; age range 19 - 36 years). One additional participant was tested but their data excluded due to technical issues (the markers would not stick firmly to the participant's hand).

Materials and Procedure

The participant was seated at a large table facing a computer screen, on which visual prompts were displayed throughout the experiment. Prior to testing, all participants were trained on the four actions that would constitute the action sequences. All actions involved manoeuvring a weighted oval-shaped plastic ball (approx. size: 9 x 5 x 5 cm). The four actions were LIFT (raise object up and back down), ROLL (with open palm, away from the participant and back), SLIDE (to the right and then back) and SHAKE (object grasped in pincer grip and rattled; Figure 1). Training was complete when the participant could correctly perform each individual action without hesitation when prompted. Test trials then began immediately.

On every trial, participants were required to perform an action sequence formed of three of the trained actions. At the beginning of each trial, the names of the actions forming the sequence were displayed on the computer screen. On no-boundary trials, the three actions were to be performed as one continuous sequence, denoted on the screen by including all words between a single bracket-pair (e.g., [*lift & roll & slide*]). On boundary trials, a boundary was added between the second and third action, denoted by presenting the first two actions grouped in a bracket-pair and the final action in separate bracket-pair (e.g., [*lift & roll*] [*& slide*]). Participants were asked to perform the sequences so as to allow an

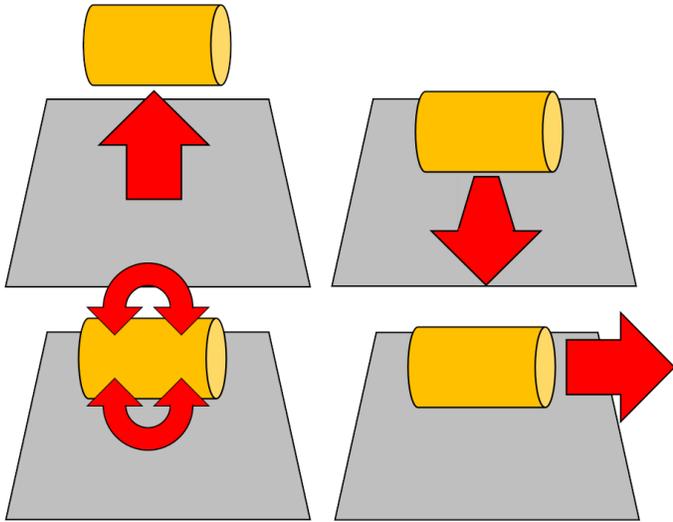


Figure 1: Schematic of individual actions. Clockwise from top-left: LIFT, ROLL, SLIDE, SHAKE. The object used in the experiment was an oval-shaped plastic ball.

imagined observer to determine whether the sequence contained a boundary or not. This procedure was a replication of studies that have used this same notation and instructions to elicit prosodic boundary cues in speech (Petroni et al., 2017; Huttenlauch et al., 2021). Participants were offered no instructions as to how they should mark the boundary between the second and third action, although they were asked not to remove their hand from the object for the duration of the sequence, and to perform each action consecutively, rather than combining actions in one single movement (i.e., shaking and lifting the object together). In an effort to avoid any priming with prosodic boundary cues, the sequences were not read aloud at any point, either by the participant or the experimenter.

The actions formed 24 possible sequences of three actions. Each sequence was performed 4 times, twice as a no-boundary trial, and twice as a boundary trial. The order of sequences was pseudo-randomized: Trial types were presented in blocks, meaning that participants performed 5 boundary trials, followed by 5 no-boundary trials, and so on. In order to present each participant with 20 blocks of 5 sequences each (i.e., 100 trials), 4 additional randomly determined sequences were presented, 2 as a no-boundary sequence, and 2 as a boundary sequence.

The participant's hand was fitted with three white spherical markers: one just below the knuckle of the index finger, one below the knuckle of the little finger, and one just above the wrist (see Figure 2). The 3D-position of the markers was continuously recorded by five Vicon motion tracking cameras mounted in the ceiling. On each trial, the participant saw the notation of the sequence on the screen (e.g., [lift & shake & roll]) and performed the respective action sequence. To avoid distraction, the screen was cleared as soon as the performance

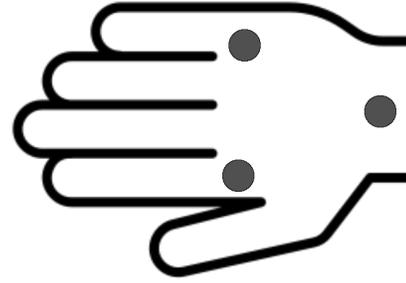


Figure 2: Approximate location of the three markers (depicted as gray circles) on the back of the hand. The position of the markers was recorded by the motion tracking cameras in 3D space.

of the action sequence began.

Data Processing

The Vicon motion tracking system recorded the 3D position of each marker in all axes of movement (transverse, frontal, vertical) at a rate of every 10 ms (100 Hz). These positions were recorded relative to the approximate center of the table-top. The oval-shaped ball was returned to the center of the table at the end of each trial.

The raw data were pre-processed in Matlab (version 2020a). The data were 10 Hz low-pass filtered (e.g., McHugh, Morton, Akhbari, Molino, & Crisco, 2020), and three-point differentiation was applied to compute instantaneous movement speed (i.e., norm of the velocity vector). These data were then exported for further analysis in R (version 3.6.2; R Core Team, 2021). Trials containing missing data or during which the wrong sequence was performed were excluded from analyses. Three of the actions moved the object through a unique axis of movement: LIFT moved through the vertical axis, ROLL moved through the frontal axis, and SLIDE moved through the transverse axis. In order to automatically extract the movement related to each individual action from the data set, we identified the point at which the index-finger marker was furthest from the starting point. For example, the action LIFT was signified by the time of the maximum extension in the vertical axis (Figure 3A). The action SHAKE did not occur in a unique axis of movement, so this action could not be automatically identified in the data. Data related to the action SHAKE were not analyzed¹. All trials were manually checked to ensure that the mid-point identification of the individual actions was accurate.

The onsets and offsets of individual actions were determined from the movement speed data, which demonstrated a

¹ Analyses of pre-boundary lengthening and pre-boundary peak amplitude were restricted to trials in which SHAKE was not performed as the second action (587 usable trials). Analysis of the pause following the second action was restricted to trials in which SHAKE was not performed as the second or final action (375 usable trials). Analysis of the duration of the final action was restricted to the trials in which SHAKE was not performed as the final action (584 usable trials).

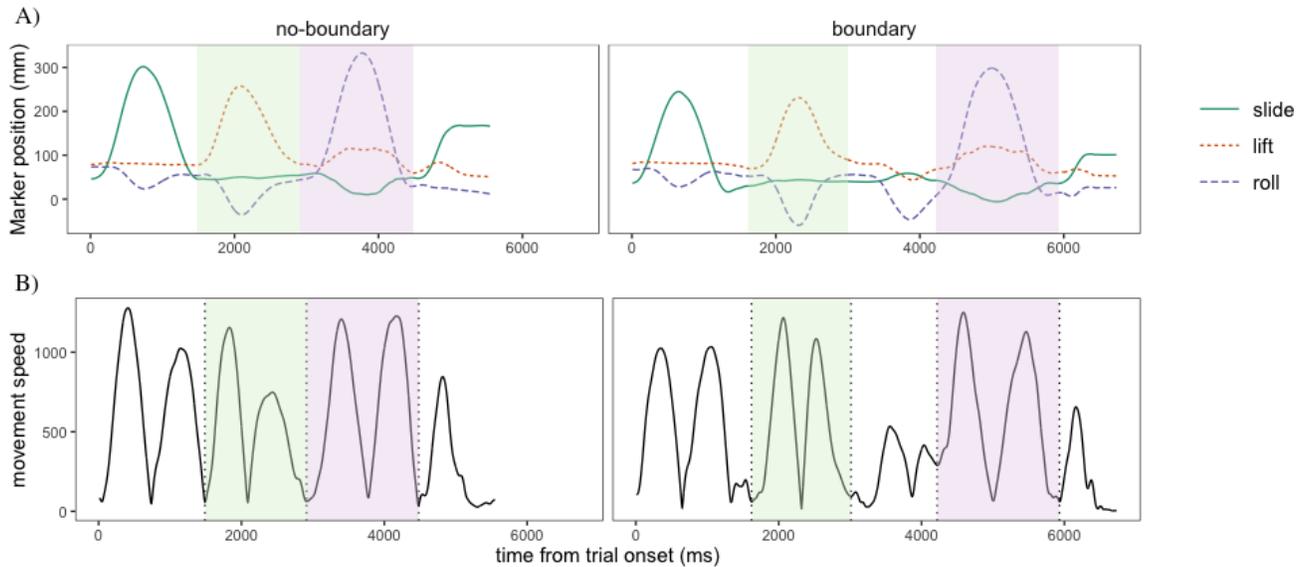


Figure 3: 3D motion tracking data from two example trials of the sequence “*slide, lift and roll*”, performed both as a no-boundary and boundary trial. A) Raw marker-position data (in mm, relative to center of table) from the index-finger. Lines represent the movement through one axis of movement. Each axis was associated with a distinct movement (solid green line = transverse axis/SLIDE, dotted orange line = vertical axis/LIFT, dashed purple line = frontal axis/ROLL). Shaded areas indicate the second and third action as defined by the movement speed data. B) Movement speed of the markers shown in A. Vertical dotted lines represent the onset and the offset of the second and third action, as defined by local speed minima.

series of distinct maxima and minima (Figure 3B). The speed minima corresponded to moments of change of position or direction of the hand, and the time-points of the minima were used as indicators of onset and offset of the individual actions in the sequence. Minima were defined as local minima in a 100 ms rolling window. Given the change in direction, the mid-point of an action also corresponded to a minimum in the movement speed data. The action onset and the action offset were defined as the time of minimum movement speed prior to and following the mid-point of the corresponding action. The following variables were calculated for each trial:

Duration of second action: Time in ms between onset and offset of second action.

Peak amplitude of second action: Maximum value of the movement speed in the interval between the mid-point and the end of the second action.

Duration of interval between second and final action: Time in ms between offset of second action and onset of third action.

Duration of final action: Time in ms between onset and offset of final action.

Statistical analyses were carried out for each variable via Linear Mixed-Effects Models (LMEMs) using the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015), with trial type as fixed effect (dummy coded: no-boundary = 0,

boundary = 1), and maximal random effects structures simplified until convergence for each model separately (Barr, Levy, Scheepers, & Tily, 2013). Inferential statistics were obtained using sequential likelihood ratio tests.

Results

Pre-boundary Lengthening

Duration of the second action was submitted to an LMEM with by-participant and by-action (LIFT, ROLL or SLIDE) random intercepts. This model revealed a significant effect of trial type, $\beta = 88.70$, $SE = 17.52$, $t = 5.06$, $\chi^2(1) = 25.09$, $p < .001$, meaning that the second action was on average 89 ms longer on boundary trials than on no-boundary trials, an indication that participants signalled the presence of the boundary by lengthening the pre-boundary action. To examine whether this extended duration affected the speed with which the second action was performed, the motion-path distance of the index-finger marker (in mm) was divided by action duration, and was submitted to an LMEM with by-participant random intercepts, which revealed a significant effect of trial type, $\beta = 0.019$, $SE = 0.0060$, $t = -3.17$, $\chi^2(1) = 9.99$, $p < .002$, indicating that, as a result of pre-boundary lengthening, hand movement during the second action was slower on boundary trials than no-boundary trials.

Pre-boundary Peak Amplitude

Peak amplitude of the second action was submitted to an LMEM with by-participant random intercepts, which re-

vealed no significant main effect of trial type, $\beta = -4.43$, $SE = 18.55$, $t = -0.24$, $\chi^2(1) = 0.057$, $p = .81$. This finding suggests either that there was no analogous kinematic cue to the prosodic boundary cue pre-boundary pitch rise in our motion data, or that pre-boundary peak amplitude was not a suitable analogous measure.

Pause Following the Second Action

Duration of the final action was submitted to an LMEM with by-participant and by-action (defined by the second action) random intercepts, which revealed a significant effect of trial type, $\beta = 786.92$, $SE = 33.88$, $t = 23.23$, $\chi^2(1) = 331.23$, $p < .001$, meaning that the onset of the final action was delayed by an average of 787 ms on boundary trials relative to no-boundary trials. This finding indicates that a pause in performance of the action sequence was present between the second and third action on boundary trials.

Duration of the Final Action

We calculated the duration of the final action and submitted this to an LMEM with by-participant random slopes and intercepts, and by-action random intercepts. Results revealed no significant main effect of trial type, $\beta = 44.8$, $SE = 22.96$, $t = 1.95$, $\chi^2(1) = 3.21$, $p = .073$. Thus, the durational cues to the presence of the boundary appear to be specific to the second action and to the interval between the second and third action.

Discussion

The current study adapted a method used to examine production of prosodic boundary cues in speech, to examine whether adults produce analogous kinematic boundary cues to signal boundaries in action sequences. Participants performed sequences of three actions on an object while the movement of their hand was recorded via 3D motion-tracking. On half of the trials, participants were asked to communicate the presence of a boundary between the second and final action, and on the other half of trials, participants were asked to perform the three actions without any boundaries. Data revealed that participants communicated the presence of a boundary between two actions by extending the duration of the pre-boundary action, and by extending the duration of the interval between the two actions. As expected, these cues are highly similar to the prosodic boundary cues pre-boundary lengthening and pause. Previous work has found evidence that boundaries in naturally-produced action sequences are marked by changes in motion velocity (Zacks et al., 2009) and we argue that these changes could be correlates of pre-boundary lengthening and pause. It is therefore possible that observers capitalize on these kinematic boundary cues to determine the structure of action sequences.

Previous research has already shown that pre-boundary lengthening and pause as boundary markers are not specific to speech: Music (i.e., a non-speech auditory stimulus) can contain pre-boundary lengthening and pause to signal the location of boundaries between musical phrases, and adult

participants show a similar brain response as measured by EEG (known as the Closure Positive Shift; Steinhauer, Alter, & Friederici, 1999) to these boundary cues in music as when they appear in speech (Glushko, Steinhauer, DePriest, & Koelsch, 2016). Critically, recent work has found a Closure Positive Shift-like response to boundaries in action sequences in both adults and 12-month-old infants (Hilton, Råling, Wartenburger, & Elsner, 2019; Hilton, Wartenburger, & Elsner, 2021). Taken together, this work strongly suggests that the cognitive processes underlying boundary cue detection are domain-general. The current finding, confirming that pre-boundary lengthening and pause are produced to mark boundaries in action sequences, indicates that the domain-general cognitive processes supporting segmentation across domains may in part reflect processing of these bottom-up boundary cues.

In the speech domain, prosodic boundary cues have been found to be sufficient to signal boundary position, for example when top-down knowledge-based cues are removed. Hence, participants are still able to detect boundary location according to prosodic boundary cues when the speech is nonsensical (e.g., Jabberwocky) or lacks syntactic or lexical information (e.g., hummed speech; Pannekamp, Toepel, Alter, Hahne, & Friederici, 2005). Importantly, kinematic boundary cues might similarly be sufficient to signal boundary position in the absence of top-down cues: Hemeren and Thill (2011) presented participants with videos of everyday action sequences (e.g., opening a bottle) that had been converted to constellations of moving-light points. These videos were produced via motion-tracking recording similar to that of the current study. One group were shown these point-light movements alongside the object that was relevant in each action (e.g., the bottle) and asked to indicate in the video points that correspond to boundaries between actions of the action sequence. This group were reliably able to describe the actions that corresponded to the light-point movements. A second group saw the same videos, but inverted, and without any knowledge of the object being used. This manipulation served to disrupt top-down conceptual processing of the movement, as evidenced by the finding that participants in this group were unable to determine the nature of the actions being shown. Critically, however, boundaries were identified in similar positions in both groups. Even if inverted, the movement itself appeared to provide kinematic cues to action boundary location independent of any top-down knowledge-based information available. Based on the current findings, we contend that pre-boundary lengthening and pause are two such kinematic bottom-up cues.

We speculated that the pre-boundary peak velocity amplitude may act as a kinematic boundary cue parallel to the pre-boundary pitch rise found in speech. However, we found no evidence that peak velocity amplitude of an action differed when the action was prior to a boundary or not. It is therefore unclear whether no parallel exists for pre-boundary pitch rise in the action domain, or whether pre-boundary peak velocity

is an inappropriate candidate for this comparison. The current work also does not allow us to pinpoint how kinematic boundary cues interact. Future work should consider whether pre-boundary lengthening and pause operate independently, as has already been studied with regards to these cues in prosody during infancy (Wellmann, Holzgrefe, Truckenbrodt, Wartenburger, & Höhle, 2012; Holzgrefe-Lang, Wellmann, Höhle, & Wartenburger, 2018). If pre-boundary lengthening and pause operate only together as kinematic boundary cues, it would be possible that one cue is a necessary product of the other. For example, pre-boundary lengthening may only be present as an artefact of preparation for the upcoming pause. Clarification of these questions would bolster our understanding of the complex interplay of action segmentation cues.

Overall, the finding that pre-boundary lengthening and pause are present as boundary cues in speech and action raises the prospect that the processes operating on these cues are similar or shared across domains. Work can now begin to examine how these kinematic boundary cues interact in shaping the observer's action segmentation, and to pinpoint the role of these cues in action processing during infancy and early childhood.

Acknowledgments

This work was supported by a grant from the German Research Foundation (DFG, project number 258522519) within the Research Unit Crossing the Borders FOR 2253 to IW (WA 2969/7-1) and BE (EL 253/6-1).

References

- Baldwin, D. A., & Baird, J. A. (1999). Action analysis: A gateway to intentional inference. In P. Rochat (Ed.), *Early social cognition: understanding others in the first months of life*. (pp. 215–240). Mahwah, NJ: Lawrence Erlbaum Associates.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278. doi: 10.1016/j.jml.2012.11.001
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi: 10.18637/jss.v067.i01
- Glushko, A., Steinhauer, K., DePriest, J., & Koelsch, S. (2016, 05). Neurophysiological correlates of musical and prosodic phrasing: Shared processing mechanisms and effects of musical expertise. *PLOS ONE*, *11*(5), 1–27. doi: 10.1371/journal.pone.0155300
- Hemeren, P. E., & Thill, S. (2011). Deriving motor primitives through action segmentation. *Frontiers in Psychology*, *1*, Article 243. doi: 10.3389/fpsyg.2010.00243
- Hilton, M., Råling, R., Wartenburger, I., & Elsner, B. (2019). Parallels in processing boundary cues in speech and action. *Frontiers in Psychology*, *10*, Article 1566. doi: 10.3389/fpsyg.2019.01566
- Hilton, M., Wartenburger, I., & Elsner, B. (2021). Kinematic boundary cues modulate 12-month-old infants' segmentation of action sequences: An ERP study. *Neuropsychologia*, Article 107916. doi: 10.1016/j.neuropsychologia.2021.107916
- Holzgrefe-Lang, J., Wellmann, C., Höhle, B., & Wartenburger, I. (2018). Infants' processing of prosodic cues: Electrophysiological evidence for boundary perception beyond pause detection. *Language and Speech*, *61*(1), 153–169. doi: 10.1177/0023830917730590
- Huttenlauch, C., de Beer, C., Hanne, S., & Wartenburger, I. (2021). Production of prosodic cues in coordinate name sequences addressing varying interlocutors. *Laboratory Phonology*, *12*(1), 1. doi: 10.5334/labphon.221
- Ludusan, B., Cristia, A., Martin, A., Mazuka, R., & Dupoux, E. (2016). Learnability of prosodic boundaries: Is infant-directed speech easier? *The Journal of the Acoustical Society of America*, *140*(2), 1239–1250. doi: 10.1121/1.4960576
- Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, *134*, 477–500. doi: 10.1037/0096-3445.134.4.477
- McHugh, B. P., Morton, A. M., Akhbari, B., Molino, J., & Crisco, J. J. (2020). Accuracy of an electrogoniometer relative to optical motion tracking for quantifying wrist range of motion. *Journal of Medical Engineering & Technology*, *44*, 49–54. doi: 10.1080/03091902.2020.1713240
- Pannekamp, A., Toepel, U., Alter, K., Hahne, A., & Friederici, A. D. (2005). Prosody-driven sentence processing: an event-related brain potential study. *Journal of Cognitive Neuroscience*, *17*(3), 407–421. doi: 10.1162/0898929053279450
- Peters, B., Kohler, K. J., & Wesener, T. (2005). Phonetische Merkmale prosodischer Phrasierung in deutscher Spontansprache. In K. J. Kohler, F. Kleber, & B. Peters (Eds.), *Prosodic structures in German spontaneous speech* (pp. 203–345). Kiel: IPDS.
- Petrone, C., Truckenbrodt, H., Wellmann, C., Holzgrefe-Lang, J., Wartenburger, I., & Höhle, B. (2017). Prosodic boundary cues in German: Evidence from the production and perception of bracketed lists. *Journal of Phonetics*, *61*, 71–92. doi: 10.1016/j.jwocn.2017.01.002
- Piazza, G., Martin, C., & Kalashnikova, M. (2021, May). *The acoustic features and didactic function of foreigner directed speech: A literature review*. PsyArXiv. Retrieved from psyarxiv.com/scnke doi: 10.31234/osf.io/scnke
- R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*(2), 191–196. doi: 10.1038/5757
- Wagner, M., & Watson, D. G. (2010). Experimental and theoretical advances in prosody: A review. *Language and Cognitive Processes*, *25*(7–9), 905–945. doi: 10.1080/03091902.2020.1713240

10.1080/01690961003589492

Wellmann, C., Holzgrefe, J., Truckenbrodt, H., Wartenburger, I., & Höhle, B. (2012). How each prosodic boundary cue matters: Evidence from German infants. *Frontiers in Psychology*, 3. doi: 10.3389/fpsyg.2012.00580

Zacks, J. M., Kumar, S., Abrams, R. A., & Mehta, R. (2009). Using movement and intentions to understand human activity. *Cognition*, 112, 201–216. doi: 10.1016/j.cognition.2009.03.007

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin*, 133, 273–293. doi: 10.1037/0033-2909.133.2.273