Language-modulated event segmentation examined through interruption detection and boundary tapping

Xiaotao Liu[1] & Norbert Vanek[1] (norbert.vanek@auckland.ac.nz), [1]University of Auckland

Introduction: Languages differ in how they encode dynamic events. For instance, English and Mandarin diverge lexically, in their use of determiners (e.g., the) [1,2] and verb-satellite constructions (e.g., cut off) [3,4] and also phrasally, in how they represent telicity, that is, whether an event is construed as having an inherent endpoint [5,6]. While prior work has examined these contrasts, only a few studies [7,8,9] have explored how language modulates event segmentation. The present research takes an original approach by examining both internal (detecting interruptions within an event) and external (tapping boundaries of an event) event segmentation across Mandarin and English speakers. We tested two hypotheses using two event types (*cutting* and *blending*). Hypothesis 1 (H1): In internal segmentation, English speakers would attend more to the early stages of *blending* events, using more determiners to describe initial object states, whereas Mandarin speakers would focus on later stages, employing more verb-satellite constructions to specify outcomes. H2: In external segmentation, English speakers' greater use of telic descriptions than Mandarin speakers would boost sensitivity to the endpoints of *cutting* events, which are inherent and clear, while endpoints in *blending* events are not.

Methods: Three experiments were conducted using 22 pairs of events, which involved cutting an object in half or blending an object into pieces (Fig. 1). In Exp. 1, participants described the events in videos as soon as they recognized them, with the aim to investigate encoding differences between Mandarin and English. In Exp. 2, participants watched the events with interruptions inserted at their early (20%), middle (50%), and late (80%) stages (Fig. 2). Their task was to detect interruptions as soon as possible. In Exp. 3, the task was to tap the start and end of object-state changes while watching videos.

Exp. 1 Event description: Forty native speakers (20 English, 20 Mandarin) watched and described 22 videos. Results showed that at the lexical level, English speakers used significantly more determiners than Mandarin speakers (odds ratio (OR) = 5.33, $z = 7.08$, $p < .001$). However, English speakers used fewer verb-satellite constructions than Mandarin speakers only in *blending* events (OR = 0.06, $z = -3.46$, $p = .003$) (Fig. 3a). At the phrasal level, English speakers used more telic descriptions for *cutting* events than Mandarin speakers did (OR = 2.29, $z = 5.73$, $p < .001$) (Fig. 3b).

Exp. 2 Interruption detection: 128 native speakers (64 English, 64 Mandarin) watched 22 videos with the task to detect interruptions. No significant accuracy differences were found between groups in *cutting* events. In *blending* events, English speakers were more accurate than Mandarin speakers at detecting interruptions at early points (OR = 2.96, $z = 3.01$, $p = .009$), but less accurate at late points (OR = 0.42, $z = -2.39$, $p = .048$) (Fig. 4). These results suggest that cross-linguistic encoding differences at the lexical level can modulate internal event segmentation in *blending* events, which aligns with H1.

Exp. 3 Boundary tapping: Forty native speakers (20 English, 20 Mandarin) watched 20 videos and tapped event boundaries. No significant between-group differences were found for RTs in *blending* events. Significant differences were observed when tapping the endpoints in *cutting* events (Fig. 5), with English speakers being faster than Mandarin speakers (OR = 0.81, $z = -2.27$, $p = .047$). These results suggest that cross-linguistic differences in phrasal-level telicity encoding can modulate external event segmentation in *cutting* events, aligning with H2.

Conclusion: Our findings support the idea that cross-linguistic differences at the lexical level and the phrasal level can modulate how speakers segment object-state change events internally and externally. These findings fit well within the theme of this year's HSP conference and can stimulate discussion on the intersections between language and thought.

Fig. 1 Paired stimuli example (Exp. 1): Cut an orange in half vs. blend an orange into substances

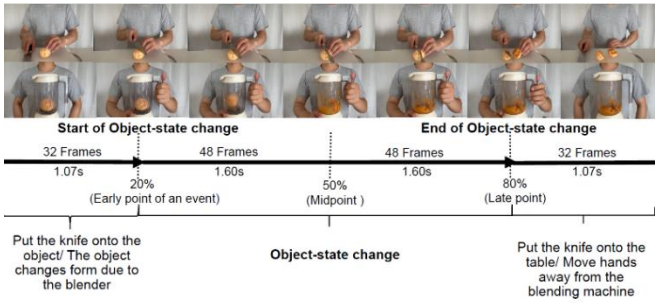


Fig. 2 Example of interrupted events (Exp. 2): Cut an orange in half

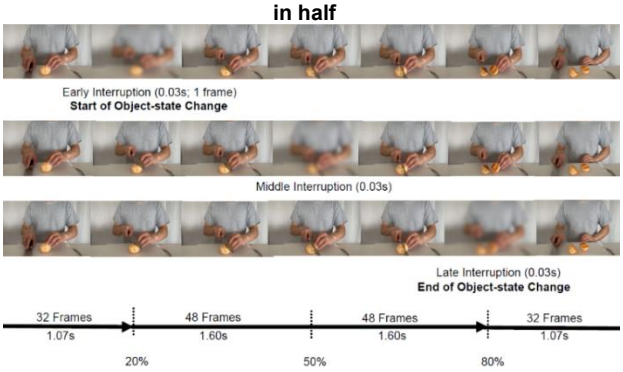


Fig. 3a Proportion of determiners and verb-satellite constructions (Exp. 1) Error bars represent ±1 SE

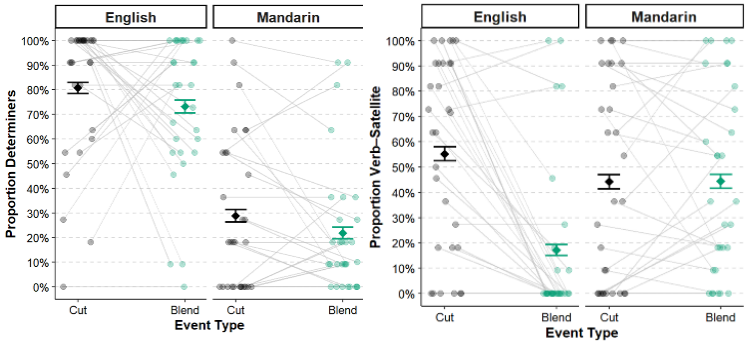


Fig. 3b Proportion of congruent descriptions[1] (Exp.1) Error bars represent ±1 SE

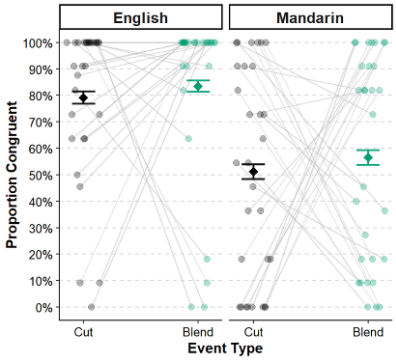


Fig. 4 Accuracy distribution of the interruption detection (Exp. 2)

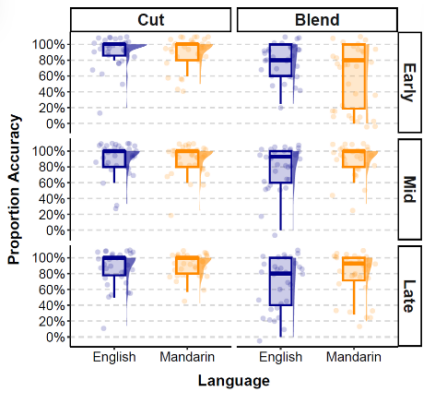
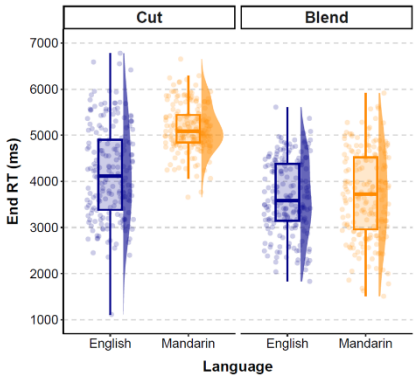


Fig. 5 RT distribution of endpoint taps (Exp. 3)



[1] Bloom, P. (1999). The role of semantics in solving the bootstrapping problem. In R. Jackendoff, P. Bloom, & K. Wynn (Eds.), *Language, logic, and concepts: Essays in memory of John Macnamara* (pp. 285–310). Cambridge, MA: MIT Press.

[2] Chierchia, G. (1998). Reference to kinds across Language. *Natural Language Semantics*, 6, 339–405.

[3] Bohnemeyer, J. (2007). Morpholexical Transparency and the argument structure of verbs of cutting and breaking. *Cognitive Linguistics*, 18(2), 153-177.

[4] Chen, J. (2007). 'He cut-break the rope': Encoding and categorizing cutting and breaking events in Mandarin. *Cognitive Linguistics*, 18(2), 273-285.

[5] Chen, P. (1988). Lun Xiandai Hanyu shijian xitong de sanyuan jiegou [On tripartite organization of the temporal system in Modern Chinese]. *Zhongguo Yuwen [Studies of the Chinese Language]*, (6), 401–422.

[6] Filip, H. (1999). *Aspect, eventuality types and nominal reference*. New York: Routledge.

[7] Gerwien, J., & Stutterheim, C. (2018). Event segmentation: Cross-linguistic differences in verbal and non-verbal tasks. *Cognition*, 180, 225–237.

[8] Ji, Y., & Papafragou, A. (2025). Representation of event boundedness in English and Mandarin speakers. *SSRN*.

[9] Vanek, N. & Zhang, H. (2024), Event boundaries stretched and compressed by aspect: Temporal segmentation in a first and a second language. *Language Learning*, 74: 104-135.

[1] Here, *congruent* means describing cutting events with telic sentences and describing blending events with atelic sentences.