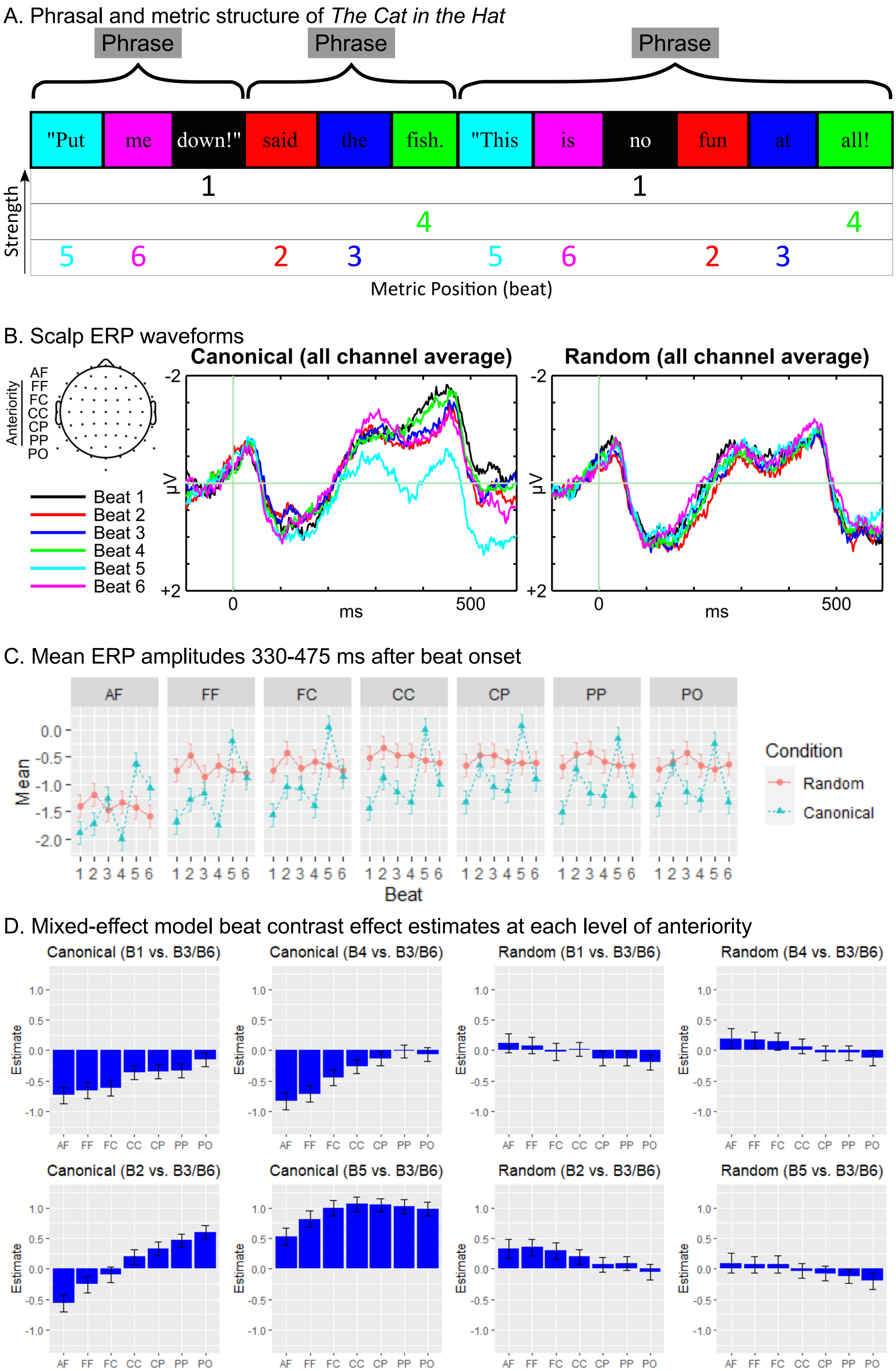
Untangling neural responses to implicit phrasing and meter in children's poetry

Ahren Fitzroy (Mount Holyoke College, UMass Amherst), Mara Breen (Mount Holyoke College)
afitzroy@mtholyoke.edu

INTRODUCTION: Perception of phrase boundaries in language and music reliably elicits a late positive deflection in event-related potential (ERP) waveforms 400-800 ms after critical word onset known as the closure positive shift (CPS; e.g., Steinhauer et al., 1999), both when phrase boundaries are acoustically realized and when they are implicit (i.e., during silent reading). Conversely, perception of metrically strong beats has been associated with a late negative frontocentral deflection in ERP waveforms from 300-500 ms (LMN; e.g., Potter et al., 2009; Fitzroy et al., 2015), both when metric strength is acoustically realized and when it is implicit (i.e., in a subjective metric hierarchy). The observation of the CPS and LMN to phrase boundaries and metric strength when they are not explicitly present in the evoking stimulus suggests these ERP components are sensitive to implicit prosody (Fodor, 2002), i.e. the top-down imposition of prosodic structure on stimuli without overt prosodic structure. In the current study, we investigated neural correlates of implicit phrase boundary and metric strength processing during the perception of aprosodic synthesizations of *The Cat in the Hat* (CITH; Seuss, 1959), a popular poetic children's book containing strong phrasal and metric structure. Comparing simultaneous implicit phrasing and metric processing indicates whether these processes are separable, and how similar they are to explicit prosodic processing.

METHOD: Twenty-seven young adults listened to two MBROLA synthesizations of CITH while 64-channel electroencephalogram (EEG) data were recorded: one with syllables in canonical order, one with syllables pseudorandomly shuffled. Presentation order was counterbalanced across participants. Intensity and pitch contours were flat, and syllables were presented isochronously with 425 ms inter-onset intervals (300 ms syllable, 125 ms silence). The metric structure of CITH groups syllables into six-beat groups (Fig. A): beat 1 has high strength; beat 4 has intermediate strength; beats 2, 3, 5, and 6 have low strength. Phrasal structure in CITH interacts with metric structure, as phrase boundaries are most likely to occur on beat 4 (69%), second most likely to occur on beat 1 (17%), and uniformly unlikely to occur on beats 2, 3, 5, or 6 (~3.5%). ERPs were averaged by synthesization (canonical, random) and beat (1, 2, 3, 4, 5, 6) (Fig. B). Mean ERP amplitudes were measured between 330-475 ms of syllable n for LMN effects and syllable $n+1$ for CPS effects (Fig. C). Mean amplitudes were analyzed using mixed-effects models with fixed effects of synthesization, beat, electrode anteriority, and electrode laterality, and participant as a random effect. Beats 1, 2, 4, and 5 were each contrasted against the mean of beats 3 and 6, and electrode anteriority and laterality were both coded with centered linear and quadratic contrasts. Interactions of beat contrasts with other fixed effect contrasts were followed up with additional models testing beat at each level of the interacting factor.

RESULTS: An LMN was evident 330-475 ms after beat 1 and 4 onsets in the canonical synthesization (Fig. B, C, D). The LMN was largest over anterior regions for beats 1 and 4, and larger over left regions for beat 1. Additionally, a positivity was evident 330-475 ms after beat 2 and 5 onsets (i.e., 455-600 ms after beat 1 and 4 offsets) in the canonical synthesization (Fig. B, C, D). Because the CPS is typically observed 400-800 ms after a phrase boundary, we interpret this positivity as a CPS to beats 1 and 4. The CPS to beat 1 was topographically constrained, present only over central and parietal regions. The CPS to beat 4 was present over the entire scalp, and largest over right central and parietal regions. The distinct timing and topography of the LMN and CPS responses supports the idea that these components are not unitary. Further, the CPS is larger in amplitude for beat 4 than beat 1, whereas the LMN estimate sizes do not differ between beats (Fig. D). Together, these results indicate that implicit phrasal and metric structure processing during artificial speech perception are separable processes, and support previous findings that the underlying neural substrates of phrasal and metric processing are similar across explicit and implicit contexts.



A. Phrasal and metric structure of The Cat in the Hat
Phrase
Phrase
Phrase
"Put me down!" said the fish. "This is no fun at all!
Strength
1
1
4
4
5 6 2 3 5 6 2 3
Metric Position (beat)
B. Scalp ERP waveforms
Anteriority
AF
FF
FC
CC
CP
PP
PO
Beat 1
Beat 2
Beat 3
Beat 4
Beat 5
Beat 6
Canonical (all channel average)
Random (all channel average)
µV
ms
C. Mean ERP amplitudes 330-475 ms after beat onset
AF FF FC CC CP PP PO
Mean
Beat
Condition
Random
Canonical
D. Mixed-effect model beat contrast effect estimates at each level of anteriority
Canonical (B1 vs. B3/B6)
Canonical (B4 vs. B3/B6)
Random (B1 vs. B3/B6)
Random (B4 vs. B3/B6)
Canonical (B2 vs. B3/B6)
Canonical (B5 vs. B3/B6)
Random (B2 vs. B3/B6)
Random (B5 vs. B3/B6)
Estimate