E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

LuCiD
The ESRC International Centre for Language and Communicative Development

Lancaster University

Gestalt auditory principles support phrase structure parsing

Tony Trotter, Padraic Monaghan, & Rebecca L.A. Frost

LuCiD Conference 2017, July 6th

Prosody and syntax

- Prosody contains information about syntax; syntactic structure is automatically mapped onto prosodic structure during production (Nespor & Vogel, 1986).

- Variation in duration, intensity and pitch systematically relate to the hierarchical structure of syntax (Nespor & Vogel, 1986; Nespor et al., 2008)

The Prosodic Bootstrapping Hypothesis

- Prosody may assist infants as they first learn to parse continuous speech into words and syntactic constituents:
    - Stress and other prosodic cues facilitate early word segmentation in English (Shukla, White, & Aslin, 2011)
    - Prosody supports syntactic segmentation in infants, and may underlie infants' early conceptualizations of syntactic constituency (e.g. Nazzi et al., 2000; Hawthorne & Gerken, 2014)
    - Prosody continues to influence word segmentation & syntactic processing in older children (Snedeker & Yuan, 2008) and adults (Langus et al., 2012)

Language Specificity?

- Do learners benefit only from the prosodic cues of their native language?
  - Yes:
    - Seidl (2007): English and Dutch 6-month olds can recognise clauses signaled in their native prosody, but not with non-native prosody
  - No:
    - Langus et al. (2012): Adult Italian speakers; both native vs. non-native prosody allowed subjects to learn a hierarchically-organised artificial grammar
    - Hawthorne, Mazuka, and Gerken (2015): English infants trained with AG strings with Japanese-like prosody can recognise grammatical movement of clauses

Universals in auditory perception?

- *If prosody's acoustic manifestations are sufficient to support learning*, experience with the target language should not be required to parse speech into constituent-like chunks:
  - E.g. The Iambic-Trochaic law
    - Cooper & Meyer (1960), Hay & Diehl (2007), Boll-Avertisyan et al. (2017)

- Music Perception literature has highlighted the role of Auditory Perceptual Gestalts for grouping rhythmic and tonal sequences
  - Pitch Similarity (e.g. The scale illusion; Deutsch, 1975a, 1975b)
  - Temporal Proximity (Lehrdahl & Jackendoff, 1983; Deliège, 1987; Deutsch, 1980)

Acoustic cues and the comprehension of speech

- For sentence comprehension, this fits in well with the *sequential processing theory* proposed by Frank et al. (2012):
  - During comprehension, listeners would have to rely on superficial, low-level cues to parse its semantics, then assign syntax accordingly

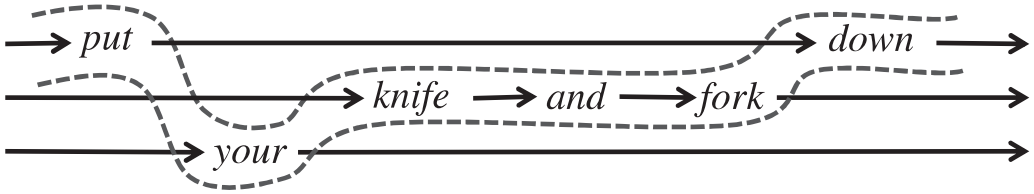


Figure 1. Combining constructions into a sentence by switching between parallel sequential streams. Note that the displayed vertical order of constructions is arbitrary.

The present study

- Aim: To assess the degree to which prosody may assist the processing of long-distance dependencies in complex syntactic structures
- Data taken from Montag & MacDonald (2014):
  - Spontaneous relative clause productions (n = 20) from American English speakers (n = 64):
    - E.g. Active/HCE, "$[\text{The bear}]_1$ $[\text{the girl}]_2$ $[\text{is hugging}]_3$ $[\text{is white}]_4$"
    - E.g. Passive, "$[\text{The bear}]_1$ $[\text{being hugged}]_2$ $[\text{by the girl}]_3$ $[\text{is white}]_4$"

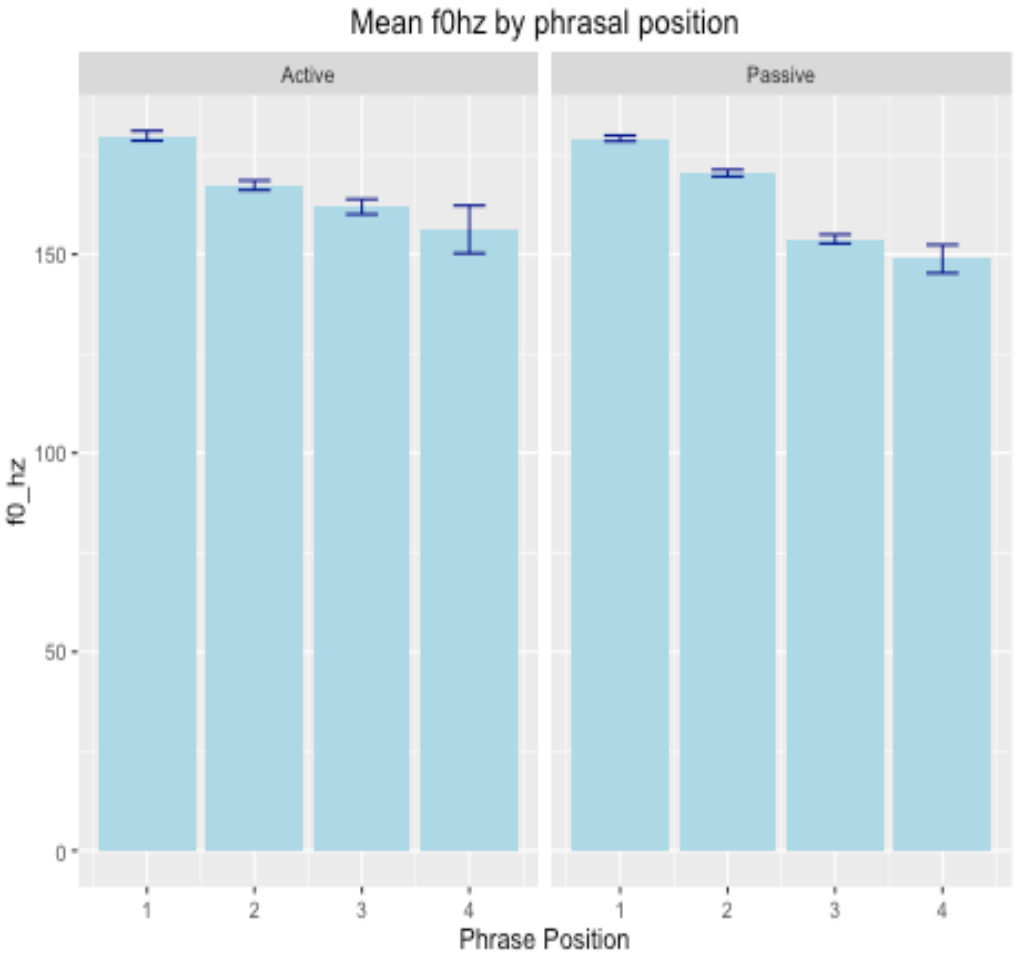
Hypotheses

(1) Phrasal units containing syntactic dependencies will be more similar in pitch, enabling grouping according to the Gestalt similarity principle

(2) Pause duration should reflect the Gestalt principle of proximity; pauses occurring between clauses will render those clauses distinct if they are longer in duration than elsewhere in the speech

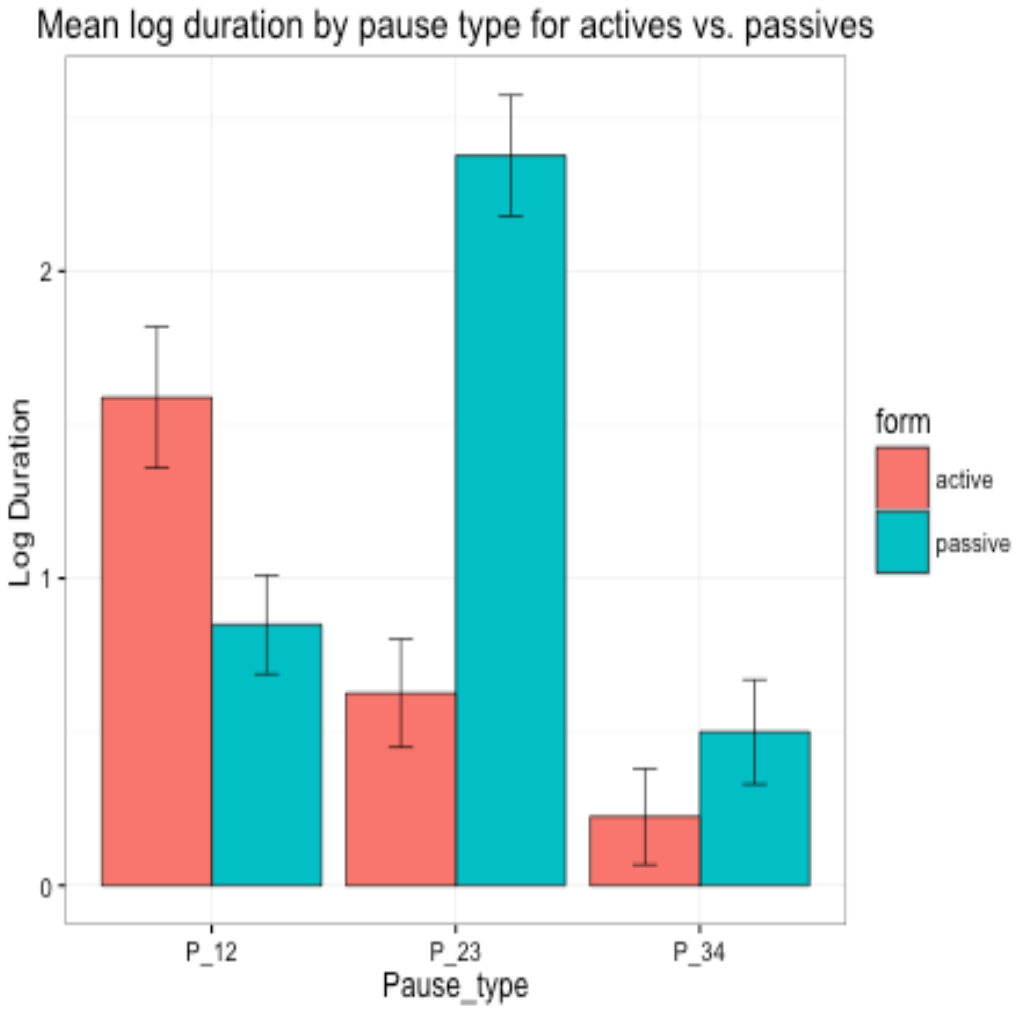
(3) Pauses should be more likely to occur between clauses than elsewhere in the speech

Results 1: Pitch Dynamics



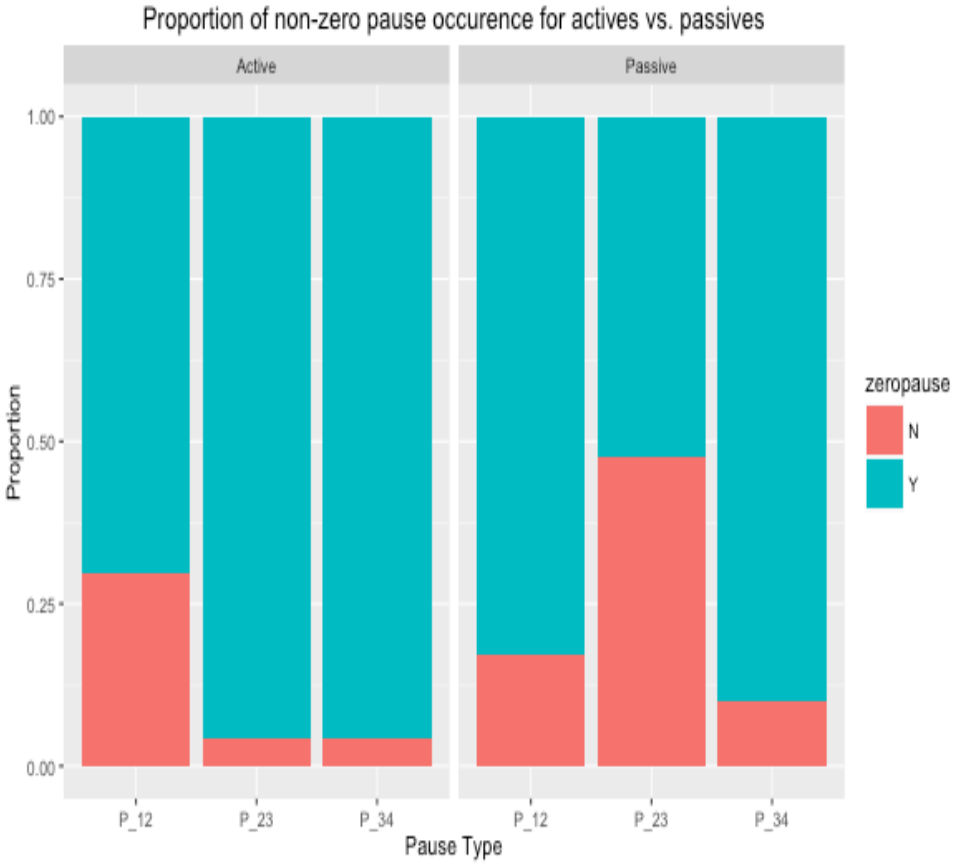
- Main effect of **position**; pos. 1 – 2 (β = -7.83, SE = 1.39, *t* = -5.69), 2 – 3 (β = -11.96, SE = 1.38, *t* = -8.69)
- **Form*Position interaction** for pos. 2 - 3 (β = 12.46, SE = 2.76, *t* = 4.52):
  - ***Pitch reduction between positions 2 & 3 greater for passive structures***

Results 2: Pause Duration



- 3 – 4 pauses are significantly shorter than 1 – 2 pauses (β = -0.119, *t* = -2.239)
- Significant Form*pausetype interaction for 1 – 2 pauses (β = 0.284, SE = 0.11, *t* = 2.518), and 2 – 3 pauses (β = -0.284, SE = 0.11, *t* = -2.518):
  - ***Longer pauses for actives between 1 & 2, and for passives between 2 & 3***

Results 3: Probability of Pause Occurrence



- No significant effect of pausetype or form
- Form*pausetype interaction approached significance for 1 – 2 (p = 0.073), and 2 – 3 (p = 0.073)

Summary

- Active Structures
  - Phrases in positions 2 & 3 are more temporally proximate, and share higher pitch similarity, making them distinct from the initial phrase of the main clause and more likely to be grouped together

- Passive Structures
  - Phrases *1 and 2*, and, *3 and 4* are thus *more temporally proximate*, and have *higher pitch similarity*, suggesting a *two-chunk structure*

Conclusions

- ***Whilst prosodic cues may result from production constraints, they may nevertheless be useful during comprehension by providing reliable, perceptual grouping cues***

- ***Prosodic cues allow auditory perceptual Gestalts to support the processing of active-object relatives, perceptually grouping the dependencies of the embedded clause, distinguishing them from the main clause***

Acknowledgements

- A big thank you for the funding provided by:
  - ESRC NW DTC (ES/L008955/1)
  - LuCiD

- And of course, a huge thanks to:
  - Jessica Montag and Maryellen MacDonald
  - Padraic Monaghan
  - Rebecca Frost
- And lastly, thanks to all of you for listening!