

The Role of Affective Prosody in the Perception of Emotional States

Presenter: **Madeline MacLean** madeline.maclean1@ucalgary.ca

PIs: Angeliki Athanasopoulou & Dimitrios Skordos

School of Languages, Linguistics, Literatures, and Cultures



1. Emotion Perception in Language

- Many linguistic and non-linguistic cues from which we infer the emotional states of others (Ben-David et al., 2016; Kikutani & Ikemoto, 2021):
 - facial expressions
 - voice quality
 - prosodic cues**
 - context / words used in conversation**
 - knowledge about our interlocutor
 - etc.

Kaufeld et al. (2020)

- Our processing system integrates and weights these cues based on their accuracy in a given environment
- The weight of a cue changes based on the environment
 - e.g., we cannot see the interlocutor's face → no facial cues
 - e.g., we are in a noisy environment → limited prosodic cues

Ben-David et al. (2016)

- Proposed two alternative theories:
 - Prosodic Dominance:** listeners are more attuned to prosodic cues in language comprehension
 - Semantic Dominance:** listeners are more attuned to pragmatic inference (or contextual cues)
- Participants were exposed to audio stimuli of a sentence
- Sentence had either congruent or incongruent emotions in the semantic/pragmatic info/prosodic info of the sentence
- asked to rate sentences when presented with:
 - general rating task** → required to rely on one cue over the other
 - semantic rating task** → required to rely on context cue
 - prosody rating task** → required to rely on prosodic cue
- On a 6-point Likert scale: How much do you agree that the speaker is (emotion)?
- Results** = Prosodic Dominance → prosodic cues played larger role than semantic cues overall

Kikutani and Ikemoto (2022)

- Replicated & extended Ben-David et al. (2016) in other languages
- Results** → Prosodic / Semantic Dominance depends on the specific emotion expressed
- E.g., in their study they found:
 - Sadness → **semantic dominance** < **prosodic dominance**
 - due to low intensity and low volume → ambiguous pronunciations = harder to comprehend words
 - Anger & Happiness → **semantic dominance** > **prosodic dominance**
 - due to high intensity, volume and faster speech rate → pronunciations of words are likely to be clearer and easier to comprehend

Summary:

- Mixed consensus: prosodic AND semantic dominance in the literature (whether other cues guide cue hierarchy or not)
- Scarce number of studies on affective prosody and emotionally relevant context and how they are processed
- Use of rating tasks instead of replicating natural speech scenarios

2. Research Question

- Is emotional prosody or situational context the primary cue for inferring other's emotions?

3. Methods

- Participants:** 9 native English speakers from linguistics courses
- Stimuli:**
 - Simple visual scenes paired with audio
 - Non-synthetic speech produced by a male and a female speaker
 - Recorded short stories of either a happy/sad context by female narrator and happy/sad utterances by the two (male and female) characters in the stories
- Software:** Praat and PsychoPy
- Task:**
 - Hear a story (context) that with cues for the emotional state of one of the two characters (happy or sad)
 - Then, they hear an utterance from the same character with prosodic cues for their emotional state (happy or sad)
 - They select whether the character is happy or sad
- Congruent Condition:** context and prosody express the same emotion
- Incongruent Condition:** context and prosody express opposite emotion

Congruent Condition (happy)

Story: Max brought Grace to an apple orchard.

Apples are Grace's favourite fruit.

Grace said: We went to an orchard today. 😊

How does Grace feel?
Happy 😊 ---- Sad 😞

Congruent Condition (sad)

Story: Max brought Grace to an apple orchard.

Apples are Grace's least favourite fruit.

Grace said: We went to an orchard today. 😞

How does Grace feel?
Happy 😊 ---- Sad 😞

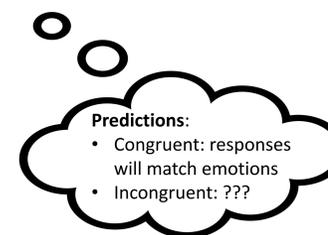
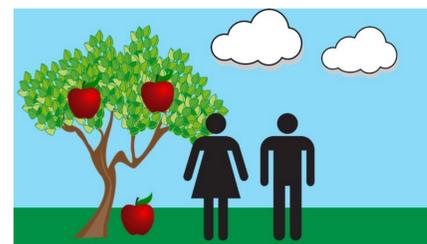
Incongruent Condition

Story: Max brought Grace to an apple orchard.

Apples are Grace's least favourite fruit.

Grace said: We went to an orchard today. 😊

How does Grace feel?
Happy 😊 ---- Sad 😞



6. Next Steps

- Finish pilot and run statistical analysis
- Currently piloting two other control experiments:
 - Given a context, are participants able to select the utterance (prosody) that matches the emotion of the context?
 - Given an utterance (prosody), are participants able to select the context that matches the emotion of the utterance?

References:

Ben-David, B. M., Multani, N., Shakuf, V., Rudzicz, F., & Van Lieshout, P. H. H. M. (2016). Prosody and Semantics Are Separate but Not Separable Channels in the Perception of Emotional Speech: Test for Rating of Emotions in Speech. *Journal of Speech, Language, and Hearing Research*, 59(1), 72–89.

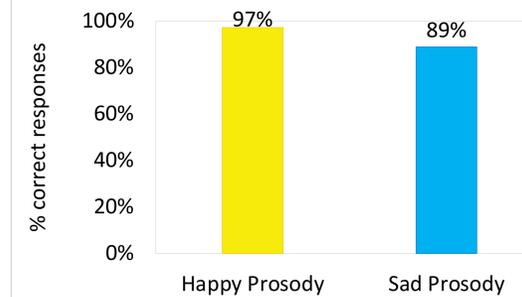
Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *The Journal of the Acoustical Society of America*, 119(5), 3059–3071.

Kikutani, M., & Ikemoto, M. (2022). Detecting emotion in speech expressing incongruent emotional cues through voice and content: investigation on dominant modality and language. *Cognition and Emotion*, 36(3), 492–511.

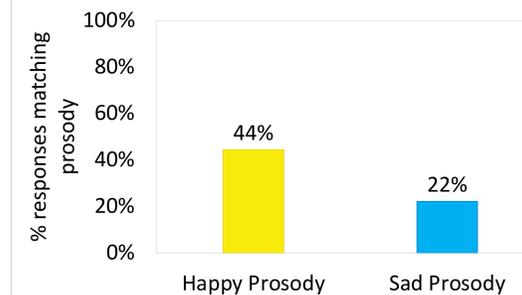
Colston, H. L. (2017). *The Routledge Handbook of Language and Humor*. Routledge.

Kaufeld, G., Ravensschlag, A., Meyer, A. S., Martin, A. E., & Bosker, H. R. (2020). Knowledge-based and signal-based cues are weighted flexibly during spoken language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(3), 549–562.

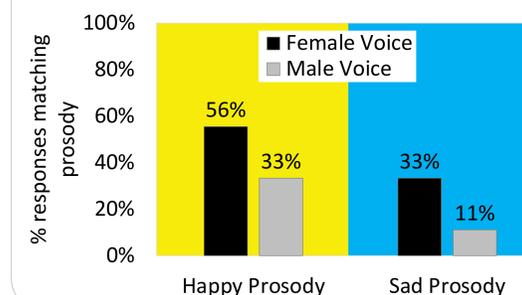
4. Pilot Results (no statistical analysis)



Congruent condition: High accuracy in identifying emotions



Incongruent condition:
• Low reliance on prosody in identifying emotions
• Difference between happy and sad



Incongruent condition:
• Rely on prosody more when happy female voice
• Rely on context more when sad male voice

5. Discussion

- Situational context and emotional prosody are too broad categories
- Perhaps subcomponents of these cues guide how the hierarchy manifests
- As in Kikutani and Ikemoto (2022) account
- How much participants depend on prosody may depend on the voice of the speaker and the specific emotion expressed
 - Male's sad prosodic range close to neutral emotion → male sad may have been perceived as neutral → participants relied on context more heavily
 - It may have been difficult to adjust cue hierarchy since participants had to switch between the male and female voice in the experiment
- Participants could have perceived the incongruent trials as ironic due to conflicting emotions being expressed
 - Verbal Irony: a person saying the opposite of what they truly meant where it's obvious what the true intention is (Colston, 2017)
 - If participants perceive it as irony → should choose context across the board consistently → not what's observed
 - To avoid confound → change instructions: tell participants that the characters are sincere and not sarcastic in nature

Acknowledgements:

