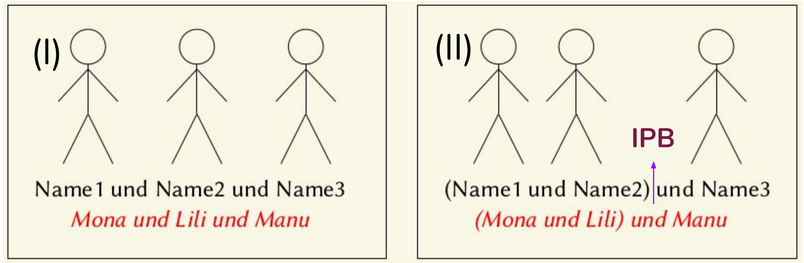
How Much Prosody is Needed? - Perception of Prosodic Grouping in Gated Ambiguous Name Sequences

Marie Hansen*, Valentina Aivazian*, Clara Huttenlauch, Carola de Beer, Sandra Hanne & Isabell Wartenburger
(* equal contribution) *SFB 1287, Cognitive Sciences, Department of Linguistics, University of Potsdam*

BACKGROUND

- in structurally ambiguous **coordinate structures**, such as (I) and (II), prosodic cues aid ambiguity resolution [1, 3]
- **intonation phrase boundaries** (IPB, see (II)) can indicate intended grouping [2, 4]
- speakers also modulate prosodic cues located before the IPB (i.e., at Name1) [2, 4]

leng = syllable duration
pau = pause
f0 = f0 rise



RESEARCH GOALS

Is a reliable detection of intended grouping possible already before the IPB?

➤ Test successive exploitation of prosodic cues in human population experimentally and through diffusion modeling

➤ Compare results to performance of machine learning (ML) models

METHODS

Human listeners (n=43):

- Gating Paradigm: 192 stimuli split into seven parts ("**gates**") each
- Gated stimuli presented successively with increasing length & amount of prosodic information
- Two alternative forced choice decision task: **grouping or no grouping**?

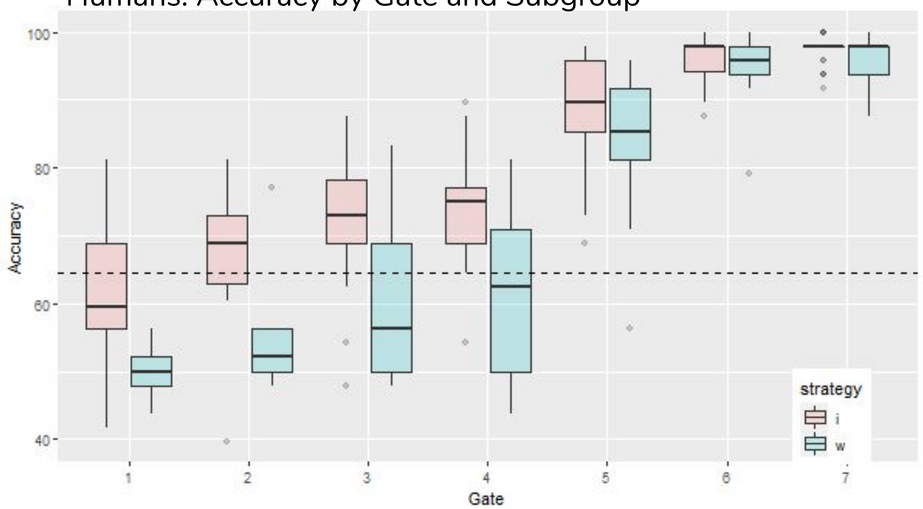
ML models:

- Linear classification models with increasing number of variables (prosodic cues), 6 models in total
- Logistic regression [6]
- Repeated *k*-fold cross-validation

Stimulus	Le	ni			und	Mim	mi			und	Ma	nu
Prosodic cues	leng	f0	leng	pau		leng	f0	leng	pau			
Gates	1	2			3	4	5			6	7	
ML models	1	2		3		4	5		6			

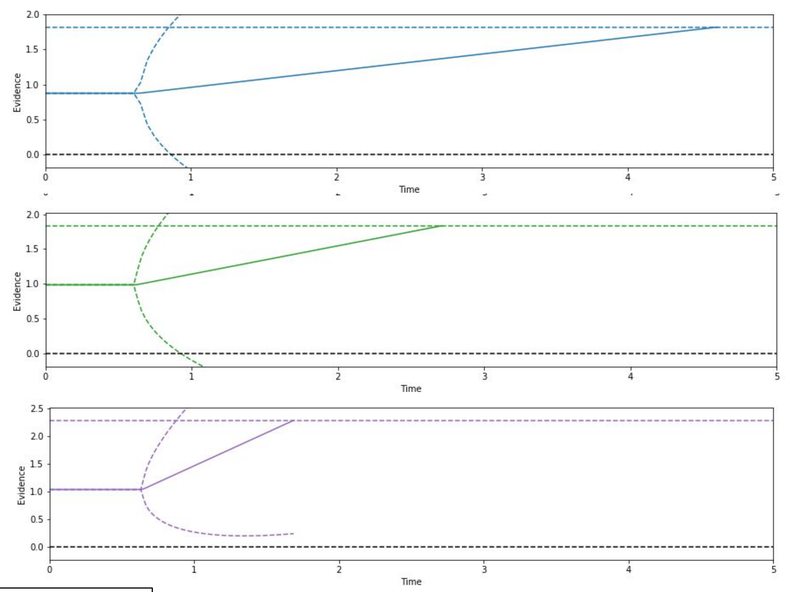
RESULTS & DISCUSSION

Humans: Accuracy by Gate and Subgroup

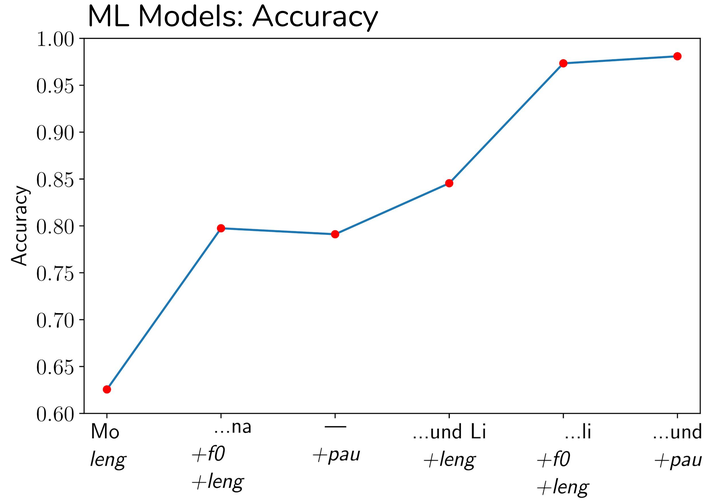


Two **subgroups** (Interrater Agreement) with different performance: Identification Strategy (i) > Waiting Strategy (w); Group level: **accuracy > chance** in gate 3; major increase in gate 5; some high accuracy scores (79%) already in gate 1, especially in i.

Humans: Drift rates over time (gate 1, gate 3, gate 5) from **Diffusion Models** [5]: higher drift = more correct & faster decisions



ML Models: Accuracy



Results from ML models: The last model, which uses all the available cues, achieved 98% accuracy in the classification task. This corresponds to gate 5 in the human study.

Deutsche Forschungsgemeinschaft
Funded by DFG, Project number 317633480, Collaborative Research Group SFB 1287, Project B01

Contact: aivazian@uni-potsdam.de
marie.hansen@uni-potsdam.de

1. Cole, J. (2015). Prosody in context: A review. Language, Cognition and Neuroscience, 30, 1–31.
2. Kentner, G., & Féry, C. (2013). A new approach to prosodic grouping. The Linguistic Review, 30 (2), 277–311.
3. Frazier, L., Carlson, K., & Clifton, C. Jr. (2006). Prosodic phrasing is central to language comprehension. Trends in Cognitive Science, 10, 244 - 249.
4. Huttenlauch, C., de Beer, C., Hanne, S., & Wartenburger, I. (submitted). Production of prosodic cues in coordinate name sequences addressing varying interlocutors.
5. Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. Neural Computation, 20(4), 873–922.
6. Pedregosa, F., Varoquaux, G., et al. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.