Applications of a Python implementation of the Generalized Context Model: In honor of an exemplary advisor

Emily Remirez, Anna Björklund, Raksit Lau-Preechathammarach, Matthew Faytak, Sarah Bakst, Auburn Barron-Lutzross

The Generalized Context Model (GCM) of speech perception, a flavor of *exemplar theory*, holds that novel speech stimuli are recognized based on their similarity to stored, categorized examples (an exemplar cloud) [1]. Simulating speech perception allows us to generate and test behavioral hypotheses and operationalize differences between theories [2]. We apply ExemPy, a user-friendly Python implementation of the GCM, to 6 phonetic datasets. ExemPy was written with an additional purpose: Accessible tools diversify both the language data and the researchers represented in phonetics. While details vary, the analyses advance three central, overlapping goals:

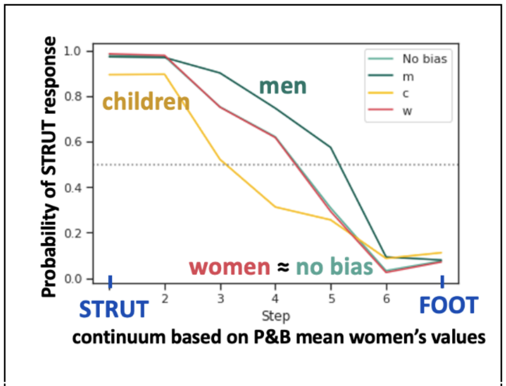

Fig 1: Results of a simulated forced choice task. Priming exemplar activation based on speaker type produces a classic "boundary shifting" effect [*]

1. Matching human behavior. ExemPy was developed using Peterson and Barney [3] and approximates their experimental confusion matrices (r = 0.95). Further simulations support sociolinguistic priming accounts (Fig 1). Modeling a wider range of behavioral data produces a better and less biased understanding of perception as a language-general process. To this end, we use production data from **Kuy** (Austroasiatic, Thailand) as an exemplar cloud to categorize continua of stimuli. We compare the simulation results to previously collected perception data [4]. The continua manipulate cues related to *register*, including f0 and voice quality. The simulation allows us to validate ExemPy on a multidimensional contrast not found in English.

category \ "response"	b	d	k	p	q	t	tS
b	0.45	0.00	0.0	0.55	0.00	0.00	0.00
d	0.00	0.98	0.0	0.00	0.00	0.02	0.00
k	0.00	0.00	1.0	0.00	0.00	0.00	0.00
p	0.05	0.00	0.0	0.94	0.00	0.00	0.01
q	0.00	0.00	0.0	0.00	0.95	0.05	0.00
t	0.00	0.19	0.0	0.00	0.00	0.73	0.08
tS	0.00	0.03	0.0	0.00	0.00	0.76	0.21

diagonal = veridical "perception"

Fig 2: Confusion matrix of Nomlaki stops, results of 327-trial simulated identification task

2. Applying simulated perception. Perceptual modeling can generate hypotheses about perception in cases where behavioral data does not yet exist or cannot be collected. This stands to facilitate experimental design and data analysis for work on low-resource languages. For instance, ExemPy offers a way to model speech perception for languages without living native speakers. We use archival recordings to classify stops and affricates in **Nomlaki** (Wintuan), an indigenous language of California with limited phonetic description [5-6]. Position in word/utterance, VOT, and burst center of gravity are examined. The resulting confusion matrices (Fig 2) shed light on how Nomlaki speakers may have historically perceived the language. Further, it may inform pedagogy in ongoing language revitalization efforts by suggesting how modern Nomlaki learners' perception differs from historical speakers'. Perceptual simulation can also complement production data from exploratory fieldwork by identifying dimensions of interest for later hypothesis-oriented work. We demonstrate this approach using data from **Kom** (Grassfields Bantu, Cameroon), in which the precise acoustic dimensions of contrast between high vowels and *fricative vowels* are poorly understood [7]. Preliminary results suggest that listeners attend to aperiodicity and formant bandwidths in addition to formant frequencies, though these cues vary in importance by speaker.

3. Modeling theories of perception. Perceivers draw on complex networks of social features (personae) associated with language use [8]. Acoustic measurements and demographics were collected from North American English-using women, alongside perception of their voices with respect to **female sexual identity** [9]. Simulating this rich dataset can systematically investigate the role of social information within the GCM, for example through cyclic categorization (*resonance* [1]), and could further be extended to model a "lexicon of talkers" [10]. Finally, **Tamil** (Dravidian) speakers' adaptation to a gradual and persistent real-time formant shift in their auditory feedback, coupled with lingual ultrasound [11], opens a window into categorization in self-perception. The inclusion of timing as a variable enables exploration of the timecourse of exemplar activation and decay; articulatory measures allow for multimodal considerations.

References [1] Johnson 2006. Resonance in an exemplar-based lexicon. *JPhon*. [2] Remirez & Johnson 2021. Activation as social encoding. LSA Meeting. [3] Peterson & Barney 1952. Control methods used in a study of the vowels. *JASA*. [4] Lau-Preechathammarach 2022. The expanding influence of Thai and its effects on cue redistribution in Kuy. *Ling Vanguard*. [5] Shepherd 2005. Proto-Wintun. *UC Publ in Ling*. [6] Björklund 2021. Nomlaki vowel quality: An acoustical and archival examination. *JASA*. [7] Faytak 2017. Sonority in some languages of the Cameroon Grassfields. Bell & Müller (eds.), *Challenging Sonority*. [8] D'Onofrio 2020. Personae in socio- linguistic variation. *WIRES: CogSci*. [9] Barron-Lutzross 2018. Speech stereotypes of female sexuality. UC Berkeley diss. [10] Munson 2010. Levels of phonological abstraction and knowledge of socially motivated speech-sound variation. *LabPhon*. [11] Bakst 2017. Articulation and altered auditory feedback. UC Berkeley diss. [*] Strand & Johnson 1996. Gradient and visual speaker normalization in the perception of fricatives. KONVENS.