Maëva Garnier, Equipe PCMD (Perception, Contrôle, Multimodalité et Dynamiques de la Parole), GIPSA-lab, Département Parole et Cognition
Lucie Ménard, UQAM, Montreal, Canada
Frédéric Elisei, Equipe CRISSP (Cognitive Robotics, Interactive Systems, Speech Processing)

Prediction of One's Own Intelligibility and Speech Adaptation to Interlocutor Feedback

Context and Questions: Speakers adapt their articulation to the communicative situation, consistent with Lindblom's H&H theory (1990)[1]: they not only adjust to the acoustic environment (e.g., Lombard speech in noise or reverberation)[2] but also to the listener's comprehension needs (e.g., infant- or foreigner-directed speech)[3]. In general, they hyper-articulate and enhance cue clarity as predicted by the probabily of perceptual confusion[4], scaling vocal intensity with ambient noise[1], raising pitch inversely with a child's age[2], and reshaping spectral characteristics to match noise type[4]. These behaviors suggest that from situational analysis, speakers are able to infer how they will be perceived by their interlocutor – i.e., their own intelligibility. We hypothesize that such predictions are developed, and then updated, using interlocutor feedback, conveyed not only verbally ("What?") but also through rich non-verbal signals (brow raises, shrugs, gaze)[5]. This project investigates how speakers perceive and weight these signals, how accurately they infer their intelligibility from them, and how quickly they revise their level of speech clarity when cues indicate misunderstanding.

Project : A face-to-face interaction experiment will be conducted in a noisy situation of communication, with around fifteen participants. An interactive game with a virtual avatar will be used to elicit controlled yet relatively natural instances of misunderstanding[1]. We will compare multiple sessions by manipulating the avatar's verbal and non-verbal signaling: no cues; verbal cues only; non-verbal cues only; graded amplitude of facial expressions; and (in)congruence between selected cues. Speaker's attention and adaptation to these signals will be characterized using occulometry (eye-tracking) and from acoustic analyses of their speech.

Perspectives : The questions raised by this project are too broad for a Master's thesis. The internship proposed here will therefore focus on designing and implementing the experimental paradigm and testing one specific manipulation of interlocutor feedback cues. Additional cues and manipulations could be pursued in a subsequent PhD project.

Required skills : Basic programming and signal-processing skills, with familiarity in speech production and behavioral psychology. Prior experience with Matlab, R, and Praat is desirable, as are strong writing skills.

Skills developped during the internship: Theoretical knowledge of social interaction and speech multimodality; Familiarization with piloting a virtual avatar ; End-to-end management of a human-subjects experiment; Skills in acoustic annotations and signal processing, Matlab or Python programming, and statistical analysis.
Supervision : The internship will take place over a period of 4–5 months under the supervision of Maëva Garnier, in collaboration with Lucie Ménard and Frédéric Elisei.

Monthly allowance : €4,35/hour (approximately €650 for a full 30-day month).

References :

[1] Lindblom, "Explaining Phonetic Variation: A Sketch of the H&H Theory," in *Speech Production and Speech Modeling*, Hardcastle&Marchal (Dordrecht, 1990).
[2] Garnier et al., "Influence of Sound Immersion and Communicative Interaction on the Lombard Effect," *Journal of Speech, Language and Hearing Research* 53, no. 3 (2010): 3.
[3] Kitamura et al., "Universality and Specificity in Infant-Directed Speech: Pitch Modifications as a Function of Infant Age and Sex in a Tonal and Non-Tonal Language," *Infant Behavior and Development* 24, no. 4 (2001): 4.
[4] Jaeger and Ferreira, "Seeking Predictions from a Predictive Framework," *The Behavioral and Brain Sciences* 36, no. 4 (2013): 4.
[5] Garnier and Henrich, "Speaking in Noise: How Does the Lombard Effect Improve Acoustic Contrasts between Speech and Ambient Noise?," *Computer Speech & Language* 28, no. 2 (2014): 2.
[6] Garnier et al., "Verbal and Nonverbal Feedback Signals in Response to Increasing Levels of Miscommunication," *INTERSPEECH 2023*, ISCA, August 20, 2023, 2698–702, https://doi.org/10.21437/Interspeech.2023-1074.
[7] Garnier and Dohen, "' I Said simPle, Not symBol!' Is Clear Speech Tailored to the Listener's Feedback?," *Speech Communication*, Elsevier, 2025, 103251.

GIPSA-lab, Département Parole et Cognition
11, rue des Mathématiques, ENSE3/DU - BP 46 - 38402 Saint Martin d'Hères cedex, France
Mail: maeva.garnier@gipsa-lab.grenoble-inp.fr Tel: (+33) 4 76 57 50 61
Web : http://www.gipsa-lab.grenoble-inp.fr/~maeva.garnier