The integration of Cued-Speech (CS) cues to speech sounds and the effect of multimodal CS cues perception on speech-in-noise understanding in profound hearing-impaired adults

Cora Jirschik Caron[1], *Jean-Luc Schwartz*[2], *Coriandre Vilain*[2], *Jacqueline Leybaert*[1], *Cécile Colin*[1]

[1]Université Libre de Bruxelles, Belgium

[2]GIPSA-lab, Université Grenoble Alpes, CNRS, Grenoble INP

cora.jirschik.caron@ulb.be,jean-luc.schwartz@gipsa-lab.grenoble-inp.fr, coriandre.vilain@gipsa-lab.grenoble-inp.fr, jacqueline.leybaert@ulb.be, cecile.colin@ulb.be

The Cued-Speech (CS) is a visual communication system composed by co-speech gestures designed to facilitate oral speech decoding in hearing-impaired people. Differently from other co-speech gestural system based on semantics, CS acts at the phonemic level of the language. The manual code was designed in fine-grained complementarity with lipreading information: hand configurations distinguish consonants and positions represent vowels. Specifically, phonemes that share a same labial image are coded by different CS cues whereas those which are visually contrastive are coded by the same CS cues. Moreover, the effectiveness of the system depends on the integration between CS cues and lipreading into a unified percept. Children who are exposed to CS early in their lives effortlessly learn to integrate CS cues to lipreading throughout their linguistic experience at home and at school (Alegria et al. 1999). For these children, the combination of CS cues to lipreading enables clear speech understanding regardless the presence of sound. In addition, previous study showed that adults who are fitted with cochlear implants can use their CS decoding skills to understand speech in the presence of noise (Bayard et al. 2019). In this communication context, deaf CS users outperformed typically hearing adults on the score of correct words-in-sentence identification. Although results confirmed that CS perception improve speech-in-noise understanding, it remains unclear whether the gain was related to audio-visual integration of acoustic information to CS cues associated to lipreading or to visual speech decoding. The fact that CS information associate to lipreading enables speech perception without sound prevents the exclusion of the last hypothesis. In this sense, the present study aimed at shedding light on the integration of CS cues to speech sounds in the absence of lipreading information. Moreover, we investigated whether speech-in-noise understanding is improved in the presence CS cues compared to auditory-only condition. Thirty adults fitted with cochlear implants or hearing aids and exposed to CS in childhood participated to the study. We tested the score of correct word-in-sentence identification in the presence of noise in eleven different conditions. Precisely, we used the adaptive tracking method to individually determine the signal-to-noise ratio (SNR) ensuring comprehension of 40% of words presented in sentences using an audio-only condition. The individually determined SNR was used to select the set of pre-prepared stimuli embedded in noise (-2dB to +6dB). Each participant was exposed to 110 trials which were presented in eleven conditions randomly mixed : Audio-only (10); CS cues-only (10); Lipreading-only (10); Lipreading + CS cues (10); Audio + CS cues (10); Audio + Lipreading (10); Audio + Lipreading + CS cues (10); Audio-only + noise (10); Audio + CS cues + noise (10); Audio+ Lipreading + noise (10); Audio + Lipreading + CS cues + noise (10). Moreover, to control for item effect, we created two orders counterbalancing sentences among conditions. To investigate whether deaf CS users integrate cues to speech sounds, we compared the mean score of correct word-in-sentence identification in Audio-only condition to Audio + CS cues condition. Secondly, we compared the score obtained in Audio-only + noise condition to Audio + CS cues + noise condition. The presence of CS cues significantly increased the mean score of correct word identification from 39% obtained in Audio-only condition to 67% in Audio + CS cues condition. Moreover, in the presence of noise, adding CS cues enabled participants to increase the mean score of 16% to attain 57%. Hence, results support the hypothesis that CS cues can be integrated to speech sound and that multimodal CS perception significantly improve speech understanding in the presence of noise. Importantly, the association of CS cues to speech sound without lipreading information is an artificial condition created to validate audio-visual integration in CS perception. When lipreading information were combined to CS cues, the mean of corrected responses achieved 87%. Adding speech sounds to CS cues associated to lipreading information slightly increased the score to 91%. Therefore, results suggest that the presence of at least two sources of information is enough to assure optimal speech understanding. Finally, the present study provided data to confirm that deaf adults who were exposed to CS at young ages are capable of integrating manual gestures to speech sounds. Moreover, the ability to decode CS information confers an advantage in adverse listening conditions.

Index Terms: Cued-Speech, Speech-in-noise perception, multimodality

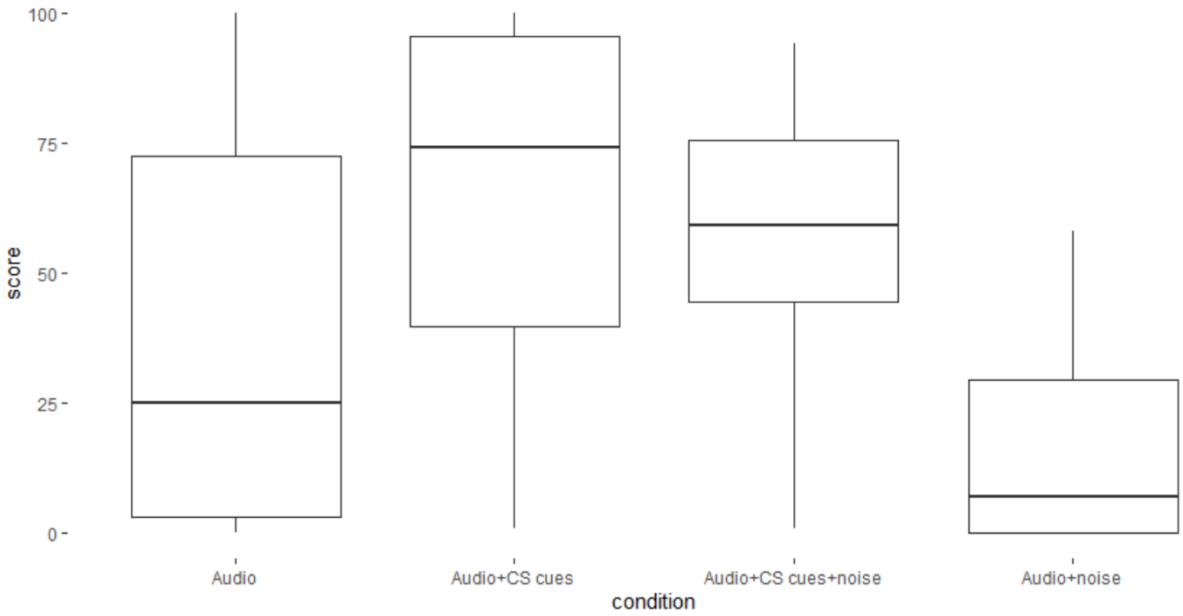

Figure 1: Percentage of correct word-in-sentence identification by condition.

References

Alegria J., Charlier B. L., Mattys S. (1999). The role of lip-reading and cued speech in the processing of phonological information in French-educated deaf children. Eur. J. Cogn. Psychol. 11 451–472. 10.1080/095414499382255

Bayard, C., Machart, L., Strauß, A., Gerber, S., Aubanel, V., & Schwartz, J. L. (2019). Cued speech enhances speech-in-noise perception. The Journal of Deaf Studies and Deaf Education, 24(3), 223-233.