Modeling multimodal processing in Simultaneous Interpreting

Seeber, Kilian

Modeling multimodal processing in Simultaneous Interpreting

Kilian G. Seeber
FTI – University of Geneva

Summary

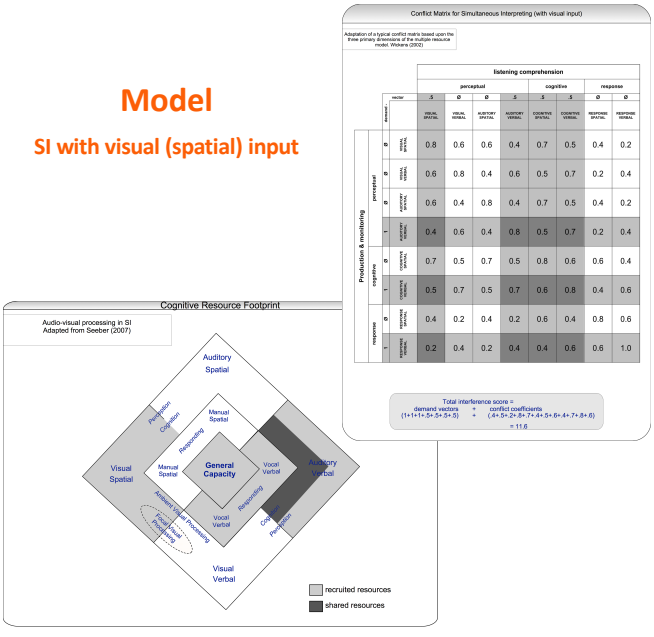
Simultaneous conference interpreting is a complex language-processing task combining language comprehension and language production in real time. One of the factors with the potential to compound the complexity of this task is the integration of visual information in the process. Whether it is facial expressions, gestures, images or text, the processing of visual information has been shown to interact with auditory information when processed simultaneously. The question about the extent to which the availability of such visual information helps or hinders comprehension in simultaneous interpreting is as yet unclear. This is an attempt to model these parameters.

Background and Theory

- In natural conversation settings auditory speech is normally accompanied by visual speech information provided by the speaker
- Visual information supports and facilitates the perception of auditory speech
- Speech reading is robust both in the comprehension of spoken and sung text
- The human processor seems capable of integrating different cues within one and the same modality (->multicue integration) as well as between or among several modalities (-> multimodal integration) without much effort.
- Single-resource theory predicts early overload during multi*cue* and multi*modal* integration
- Multiple-resource theory predicts early overload only during multi*cue* integration

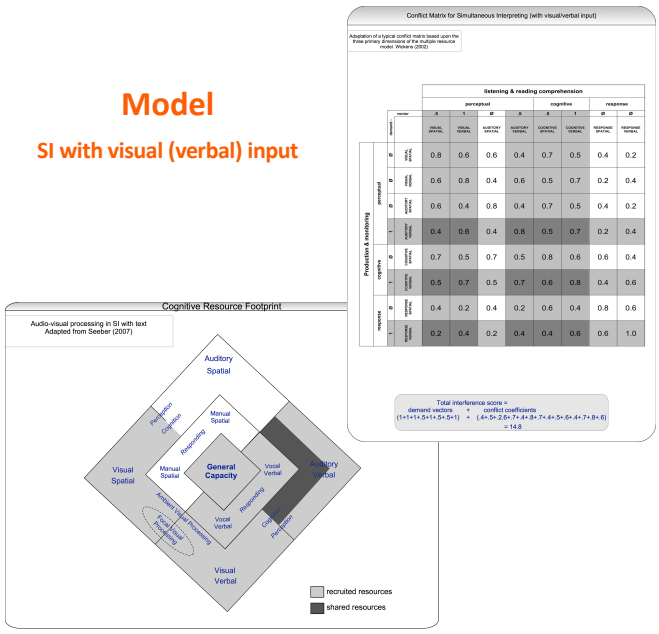
Model

SI with visual (spatial) input



Model

SI with visual (verbal) input



Discussion and Conclusion

Audio-visual information can be integrated even when the information coming from the two channels is not perfectly synchronous.

Audio-visual information presented between a 40ms audio lead and a 240ms audio lag, in other words within a window of just under 300ms.

Leading auditory information hinders integration while leading visual information enhances integration.

- ➢ A multimodal facilitation effect during SI with text is contingent on the synchronicity of the two signals (e.g., text and speech)
- ➢ Asynchrony of signals on different channels is expected to increase cognitive load
- ➢ Interpreter's inability to keep this synchrony during SI with text would explain the perceived increase in task difficulty as compared to SI without text

References

Groh, J.M. & Werner-Reiss, U. (2002). Visual and auditory integration. In *Encyclopedia of the human brain* (Vol. 4., pp. 739-752). San Diego: Academic Press.
Jesse A., Vrignaud N., Cohen M.M., & Massaro, D.W. (2000/01). The processing of information from multiple sources in simultaneous interpreting. *Interpreting 5*(2), 95-115.
Lu, S.A., Wickens, C.D., Sarter, N.B., Thomas, L.C., Nicolic, M.I., & Sebok A. (2012). Redundancy gains in communication tasks: a comparison of auditory, visual, and redundant auditory-visual information presentation on NextGen flight decks. Proceedings of the human factors and ergonomics society. 56th annual meeting. 1476 – 1480.
Maragos, P., Gros, P., Katsamanis, A. & Papandreau G. (2008). Cross-modal integration for performance improving in multimedia: A review. In Maragos P., Potamianos A. & Gros P. (Eds.), *Multimodal processing and interaction: Audio, video, text* (pp.3-48). NY: Springer.
Massaro, D.W. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. Cambridge, MA: MIT Press.
Seeber, K.G. (2007). Thinking outside the cube: Modeling language processing tasks in a multiple resource paradigm. Interspeech 2007, Antwep, Belgium, (pp. 1382-1385).
Seeber, K.G. (2011). Cognitive load in simultaneous interpreting: Existing theories – new models. *Interpreting 13*(2), 176-204.
Seeber, K.G. (2015c). Cognitive load. In Pöchhacker, F., Grbic, N., Mead, P. & Setton, R. (eds.), *Encyclopedia of Interpreting Studies* (pp. 60-61). New York: Routledge.
Van Wassenhove, V., Grant, K. W., & Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia 45*, 598-607.
Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science 3*(2), 159-177.