Perception of anticipatory labial coarticulation by Belgian French blind listeners: A comparison with sighted listeners in audio-only, visual-only and audiovisual conditions

V. Delvaux[*,+], K. Huet[*], M. Piccaluga[*], B. Harmegnies[*]
[*]Institut de Recherche en Sciences et Technologie du Langage, UMONS, [+]FNRS, Belgium
Veronique.Delvaux@umons.ac.be

Speech perception is multimodal. Information from the auditory and visual modalities is integrated in speech processing, as evidenced by the classical McGurk effect (for a review, see Tiipana, 2014). In particular, audio-visual integration exploits the time-varying properties shared by the acoustic and visual signals as a result of their structural coupling within the talking individual (e.g., Chandrasekaran *et al*., 2009). Auditory and visual modalities typically convey complementary information, so that audio-visual integration, when possible, improves speech perception. Indeed, visual information enhances overall speech intelligibility and perceptual learning, especially in adverse conditions, such as in noisy environments (Grant & Seitz, 2000; Schwartz *et al*., 2004; Wayne & Johnsrude, 2012; Winn *et al*., 2013).

By definition, visual information is not available for speech perception in blind people. Actually, when blindness is congenital or early acquired, there is substantial evidence of cross-modal plasticity, with visual cortex activity observed for various tasks involving sound processing (e.g. Bedny *et al*., 2012). Behavioral studies have documented enhanced performance of blind listeners over sighted controls in a wide range of auditory tasks, such as auditory spatial tuning, echolocation, processing of simple sounds like tones, pitch detection, absolute pitch, and identification of voices (for a review, see Kupers & Ptito, 2014). Concerning speech sounds specifically, there is a large body of evidence demonstrating improved processing of (synthetic or time-compressed) fast speech in blind people (Gordon-Salant & Friedman, 2011; Dietrich *et al*., 2013), and scarcer evidence of enhanced intelligibility of speech material (words and sentences) in a noisy environment (Niemeyer & Starlinger, 1981; Muchnik *et al*., 1991; Chen and Chen, 2014).

There is very little phonetic literature reporting on the speech perception of segmental contrasts in blind people using the most classical tasks in speech perception, namely identification and discrimination tasks. One exception is the work conducted by Ménard and colleagues, who investigated the production and perception of synthetic vowels by 12 congenitally blind and 12 sighted native speakers of Canadian French (Ménard *et al*., 2009). Using an AXB discrimination task, they found that blind speakers had significantly higher peak discrimination scores than sighted speakers for the /e-ɛ/ and /ɛ-a/ contrasts (the /i-y/ contrast almost reaching significance), but not for the /i-e/ and /y-u/ contrasts. These results are consistent with a "compensatory" account of the gain in auditory acuity in blind listeners since visual information susbstantially contributes to the perception of height and (even more) of rounding in French vowels, either in static configurations (i.e. for sustained vowels: Robert-Ribes *et al*., 1998) or in dynamic configurations (i.e. in anticipatory labial coarticulation: Troille *et al*., 2010; Roy, 2012).

Little is known of the perception of anticipatory labial coarticulation in blind people. Hirsch and colleagues reported that a group of 8 blind and 2 visually-impaired French people correctly identified rounded vowels /y, u, ø, œ/ earlier than sighted controls in gated [$V_iC_sV_{rounded}$] audio stimuli (Hirsch *et al*., 2011). Even if the differences in methodology (natural vs. synthetic stimuli; AXB discrimination vs. identification in gating; etc.) render difficult a detailed comparison, the pioneer work carried out by both Hirsch's and Ménard's teams suggests that even with visual deprivation, or maybe in compensation for it, the perception of rounding is preserved in blind people, and may even outperform that of sighted listeners in an auditory-only condition.

The aims of the present study were to complement this previous work by (i) fully apprehending the dynamics of the perception of the /i-y/ contrast by adult Belgian French blind listeners, and (ii) positioning precisely the performances of the blind listeners with respect to the range of performances in audiovisual speech perception demonstrated by sighted controls (matched for age and gender). To achieve these goals, 16 participants completed two tasks on pairs of stimuli gated from original [agi] and [agy]

(reference gate at burst), a two-alternative forced choice identification task and an AX discrimination task. Both tasks were performed on stimuli which were presented in quiet vs. (acoustically-) noisy conditions, combined with three sensory modalities: audio-only (for blind and sighted listeners), audio-visual and visual-only (for sighted listeners only). Performances were assessed using metrics based on proportion of correct answers (to pairs of different stimuli), measured as a function of gate.

Results may be summarized as follows: (i) in the audio-only condition, blind listeners overall outperformed sighted listeners, even more so for a lower signal-to-noise ratio, which mostly conforms with the relevant literature; (ii) overall, sighted listeners exhibited strong visual enhancement, i.e. better performances in the audio-visual modality than in the visual-only and audio-only conditions (in that order, presumably because in our gated stimuli information was incomplete in both channels but available earlier in the visual stream), and stronger visual enhancement when acoustical noise was added; (iii) to some extent, the enhanced performances of blind listeners were mediated by the perceptual task to be performed, i.e. discrimination vs. identification of speech stimuli ; (iv) complex interactions were observed between groups of listeners (blind vs. sighted) and perceptual conditions (noise as well as sensory modalities) in terms of relative timing between perceptual scores and the date of the earliest available information in the audio and visual streams. We will discuss at the conference how these results may be relevant for current theories of speech perception as a - multimodal - dynamic process (e.g. Jesse & Massaro, 2010).

References

Bedny, M., Pascual-Leone, A., Dravida, S., and Saxe, R. (2012). A sensitive period for language in the visual cortex: Distinct patterns of plasticity in congenitally versus late blind adults. *Brain Lang,* 122, 162–170.

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A.A. (2009). The natural statistics of audiovisual speech. *PLoS Comput Biol,* 5, e1000436.

Chen, M.-S., Liu, J.-S., & Chen, W.-R. (2014). Different performance on auditory discrimination between sighted and blind individuals. *Proc. 5th IEMI*, 133-137.

Dietrich, S., Hertrich, I., & Ackermann, H. (2013). Training of ultra-fast speech comprehension induces functional reorganization of the central-visual system in late-blind humans. *Front Hum Neuro* 7, 701.

Gordon-Salant, S., & Friedman, S.A. (2011). Recognition of rapid speech by blind and sighted older adults. *J Speech Lang Hear R,* 54, 622–631.

Grant, K.W., & Seitz, P.F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *J Acoust Soc Am,* 108, 1197-1208.

Hirsch, F., Sock, R., Vaxelaire, B., Fauth, C., Bouarourou, F., Béchet, M., & Barkat-Defradas, M. (2011). A preliminary study on speech perception and production by blind people. *Proc ISSP 2011, Montreal, Canada*.

Jesse, A., & Massaro, D.W. (2010). The temporal distribution of information in audiovisual spoken-word identification. *Atten Percept Psychophys,* 72, 209-25.

Kupers, R., & Ptito, M. (2014). Compensatory plasticity and cross-modal reorganization following early visual deprivation. *Neurosci Biobehav R,* 41, 36-52.

Ménard, L., Dupont, S., Baum, S.R., & Aubin, J. (2009). Production and perception of French vowels by congenitally blind adults and sighted adults. *J Acoust Soc Am*, 126, 1406-1414.

Muchnik, C., Efrati, M., Nemeth, E., Malin, M., & Hildesheimer, M. (1991). Central auditory skills in blind and sighted subjects. *Scand Audiol,* 20, 19–23.

Niemeyer, W. & Starlinger, I. (1981). Do the blind hear better? Investigations on audi- tory processing in congenital or early acquired blindness. II. Central functions. *Audiology*, 20, 510–515.

Robert-Ribes, J., Schwartz, J.L., Lallouache, T., & Escudier, P. (1998). Complementarity and synergy in bimodal speech: Auditory, visual, and audio-visual identification of French oral vowels in noise. *J Acoust Soc Am,* 103, 3677-3689.

Roy, J.-P. (2012). Anticipation labiale et perception visuelle. *Can J Ling*, 57, 109–132.

Schwartz, J.L., Berthommier, F., Savariaux, C. (2004). Seeing to hear better: evidence for early audio-visual interactions in speech identification. *Cognition*, 93, B69-78.

Tiippana, K. (2014). What is the McGurk effect? *Front Psy*, 5, 725.

Troille, E., Cathiard, M.-A., & Abry, C. (2010). Speech face perception is locked to anticipation in speech production. *Speech Commun*, 52, 513–524.

Wayne, R.V., & Johnsrude, I.S. (2012). The role of visual speech information in supporting perceptual learning of degraded speech. *J Exp Psy: App*, 18, 419-435.

Winn, M.B., Rhone, A.E., Chatterjee, M., & Idsardi, W.J. (2013). The use of auditory and visual context in speech perception by listeners with normal hearing and listeners with cochlear implants. Front. Psy. 4, 824.