PRODUCTION-PERCEPTON RELATIONSHIPS IN SENSORY DEPRIVED POPULATIONS: THE CASE OF VISUAL IMPAIRMENT

Lucie Ménard

Laboratoire de phonétique, Université du Québec à Montréal, Montréal, Canada
Center for Research on Brain, Language and Music, Montréal, Canada
menard.lucie@uqam.ca

ABSTRACT

Current understanding of speech production and perception assumes that multisensory (auditory, visual, and proprioceptive) information is linked to the actions of orofacial articulators, and these relationships are established early in life. Although many studies have examined the role of visual cues in speech *perception*, less is known about their role in speech *production*. We review studies of speech produced by congenitally blind versus sighted French speakers. They suggest that early access to visual input impacts articulatory strategies used to implement phonological targets. Implications for perceptuo-motor theories of speech production and perception are discussed.

Keywords: speech production, vision, speech development, multisensory perception.

1. MULTISENSORY PRODUCTION-PERCEPTION RELATIONSHIPS

In face-to-face conversation, speech is produced and perceived through various modalities. Movements of the lips, jaw, and tongue, for example, modulate air pressure to produce a complex waveform perceived by the listener's ears. Visually salient articulatory movements (of the lips and jaw) also contribute to speech identification in acoustically degraded conditions [45, 46] and in non-degraded conditions [1]. The seminal McGurk effect [22] shows how high-level speech perception integrates auditory and visual features. This phenomenon occurs when a listener hears a stimulus such as /pa/ while watching a discordant visual stimulus such as /ka/. For most participants, the resulting percept is a fusion of both modalities (i.e., /da/).

In recent decades, much evidence was found to support the existence of a functional link between action and perception in speech, e.g. [5, 13, 42]. This phenomenon is referred to as "sensorimotor integration." Several behavioral and neurophysiological studies have attempted to evaluate how speech perception processes are influenced by production tasks [25, 36, 39, 15, 41, 49]. Other experiments have explored the link between production and perception by measuring the effects of various perceptual tasks on speech production, e.g. [7, 8, 34]. Importantly, these studies suggest that not only heard speech, but also seen and felt speech is tightly linked to speech perception and production [9].

2. VISUAL IMPAIRMENT AND SPEECH

Since multimodal sensory input is linked to speech production, sensory deprivation can affect speech production. Many studies have shown that severe to profound hearing loss greatly affects a speaker's ability to produce intelligible speech, as reviewed in [6]. Reliance on visible articulators such as the lips and jaw is enhanced in some hearing-impaired listeners. In contrast, less is known about the influence of visual deprivation on speech perception and production. The fact that congenitally blind speakers learn to produce correct speech sounds suggests that visual cues are not needed to control speech movements.

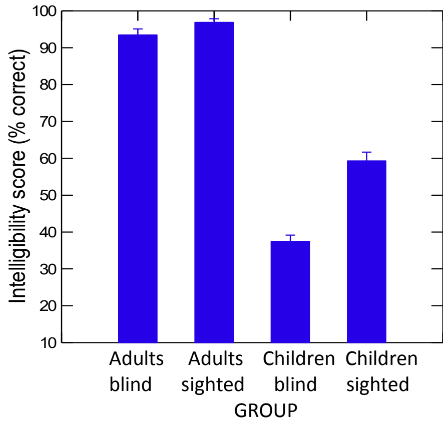
A quick examination of the perceptual saliency of French contrasts in the auditory and visual channels suggests that some features might be more affected by visual deprivation than others. French oral vowels are organized along three phonological contrasts: height, rounding, and place of articulation. The phonetic implementation of those contrasts requires both visible (jaw and lips) and invisible (tongue) articulators. For example, producing rounding contrasts (involving the lips) is more visible than producing place-of-articulation contrasts (involving the tongue), although modelling studies reveal that complementary maneuvers between the tongue and the lips can produce the same acoustic target [23]. Regarding the height dimension, however, various gestures can be recruited to achieve specific contrasts, in complementary ways [20, 32, 33]. For example, contrasts between high and low vowels can be implemented mainly through variations in jaw position, the tongue being passively carried. In contrast, jaw position can remain relatively stable while the tongue is actively elevated or depressed. These articulatory maneuvers yield various visual effects, jaw movement being visible while tongue movement is not (apart from tongue tip displacements, partially visible when the mouth is opened). It remains to be determined whether the use of visible vs. invisible articulators in the implementation of French vowels is affected by sensory input.

2.1. Speech perception and production in congenitally blind adults

It is well documented that auditory perception abilities of congenitally blind individuals differ from those of sighted individuals [12, 14, 19, 27, 30, 43, 44]. Many studies have also shown how cortical areas devoted to vision in sighted subjects are active in blind speakers during speech perception tasks, e.g. [3], pointing to the important cortical reorganization that occurs in sensory-deprived populations. Since, according to many, the ability to perceive speech is related to the amount of contrast produced between two sounds, e.g. [34], this between-group difference in auditory discrimination may entail differences at the production level. Furthermore, apart from differences in auditory discrimination between congenitally blind and sighted speakers, visual deprivation per se might also lead to differences in the control of the speech articulators (especially the visible ones).

The effects of congenital visual impairment on the perception and production of phonetic contrasts in Canadian French vowels were investigated in a series of studies. First, a cohort of 12 congenitally blind adults and age-matched sighted adult control participants was recruited [24]. All participants produced several repetitions of the ten French oral vowels. Auditory acuity was evaluated through AXB discrimination tests of synthesized vowels along five continua: /i/ versus /e/, /e/ versus /ɛ/, and /ɛ/ versus /a/ (representing height), /y/ versus /u/ (representing place-of-articulation), and /i/ versus /y/ (representing rounding). Blind speakers had significantly higher peak discrimination scores than sighted speakers for the /e/–/ɛ/ and the /ɛ/–/a/ contrasts. The difference in peak discrimination scores for the /i/–/y/ continuum did not reach significance ($p < .07$) but the observed pattern was similar to the significant difference noted for the other two contrasts. At the acoustic level, produced contrast distances, measured by the value of average vowel space (AVS), were significantly higher for sighted speakers than for blind speakers. Vowels were thus spaced farther apart in the acoustic formant space for sighted speakers than for blind participants, despite the higher auditory discrimination scores attained by the latter group. Next, a subset of the vowels produced by all speakers were submitted, as an auditory identification test, to a group of 20 adult sighted French listeners. As shown in the two left-most bars in Fig. 1, interestingly, the intelligibility scores did not differ significantly between blind or sighted adults. (Scores in children are discussed later).

Figure 1: Intelligibility scores (% correct responses) of vowels produced by sighted and blind adults and children.



To further investigate the effects of blindness on speech production, the contributions of upper lip protrusion and tongue shape/position in the implementation of the French phonological vowel contrasts mainly involving those articulators (rounding, place of articulation, rounding and place of articulation combined) were examined [25]. Ultrasound imaging was used with audiovisual recordings. This showed that the lips and tongue were involved in the implementation of the rounding contrast, but the magnitude of the variance in upper lip protrusion (in mm) between those vowel pairs was significantly greater for sighted participants than for blind participants. Regarding the place of articulation feature, tongue front-back position differences between those pairs were significantly greater for congenitally blind speakers than for their sighted peers. However, the contribution of upper lip protrusion was reduced for the blind speakers, suggesting a trade-off relationship. The analysis of vowel pairs involving contrasts in both rounding and place of articulation showed that sighted participants had a larger range of upper lip protrusion compared to blind participants. The reverse pattern was found for tongue curvature and front-back position of the tongue, for which the blind group produced greater variation in articulatory position.

In follow-up studies, variations of intelligibility demands were done through manipulations of prosodic focus [25] and speaking condition [26]. Contrastive focus has been reported to increase perceptual saliency and to enhance phonemic distinctiveness, whereas speaking condition, especially clear speech, is used to increase global intelligibility of the constituents. Concerning prosodic focus [25], two groups of nine speakers were recorded while producing the vowels /i/, /y/, /u/, and /a/ in three consonantal contexts (/b/, /d/, and /g/) and in two prosodic conditions (contrastive focus and neutral). Both subject groups produced acoustic correlates of focus, but the articulatory strategies they used differed. At the acoustic level, this study showed that both sighted and congenitally blind speakers used increased values of pitch (F0), intensity (RMS), and duration to signal prosodic contrastive focus in French. At the articulatory level, lip geometry was affected differently by the prosodic condition: the internal lip area values were significantly increased under focus for all consonantal contexts in sighted speakers, while they were not significantly increased for blind speakers. As for upper lip protrusion, prosodic condition was found to affect only the vowel /y/ in sighted speakers. In the case of clear speech [26], sighted and congenitally blind participants were recorded using electromagnetic articulography (EMA) while producing multiple repetitions of the ten French oral vowels in carrier sentences, in conversational and clear speaking conditions. Articulatory variables (lip, jaw, and tongue positions) as well as acoustic variables (contrasts between vowels, within-vowel dispersion, pitch, duration, and intensity) were measured. Lip movements were larger when going from conversational to clear speech in sighted speakers only. However, tongue movements were affected to a larger extent in blind speakers compared with their sighted peers. Taken together, these studies provide evidence that production-perception relationships in speech are different in congenitally blind individuals compared with sighted individuals.

2.2. A developmental perspective

Early in life, sighted infants demonstrate strong capacities to associate sounds with corresponding visual representations of the lips [16, 17, 38]. Babies also imitate labial movements of sounds that are visually presented. It is therefore clear that at the language acquisition stage, infants establish relationships between auditory parameters and visual events. Fine-tuning of speech perceptual processing abilities continues in childhood. For instance, the weight given to multiple sensory input (such as visual and auditory cues) is not the same at 10 years of age as in adulthood [10, 22]. Furthermore, sensory modalities interact with each other during development [11]. When one modality is missing, such as in cases of congenital visual deprivation, a complex reorganization of sensory processing occurs. Such deprivation could affect the strategies used to develop language [21, 37] and more specifically to produce phonological targets. Lewis [18] reported that at the pre-babbling stage, there was less imitation of lip gestures by a blind baby than by sighted babies. Blind babies also show longer babbling phases, as well as delays in the production of the first words [2, 48]. Elstner [4] and Mills [29] presented several studies showing phonological delays and phonetic-phonological disorders in older children. In a study of syllables produced by a congenitally blind 2-year-old German child, Mills [28] reported a higher number of phonological confusions between groups of visually dissimilar consonants (labial /b/ vs. velar /k/) for the blind child compared to two English-speaking sighted children. A few studies of phonological awareness, however, reported contradictory results. Lucas [19] reported a similar percentage of correct responses in blind and sighted children in an imitation task. Thomas, Prost, Espesser, and Rey [47], in contrast, found differences in responses from eight visually impaired children aged 6.5 to 9.5 years and eight age-matched control subjects. In a non-word repetition task, the visually impaired children had significantly more errors on phoneme contrasts based on visible place of articulation such as /p/ and /k/. Prost, Espesser, Sabater, Thomas-Bartalucci, and Rey [35] further studied access to phonological targets and found similar results. These mixed results might be ascribed to additional variables. Indeed, it is difficult to study homogeneous populations of blind speakers because observed differences in speech production abilities between blind and sighted groups might be related to the presence of uncontrolled variables such as additional motor control disorders or language disorders unrelated to the visual impairment.

A recent study investigated the acoustic and articulatory strategies used by 12 congenitally blind French speaking children aged 5 to 10 years old and 12 age-matched, sighted control children. They were instructed to produce multiple repetitions of a few sentences in a neutral condition and in a focused condition (similar to the elicitation method used in [25]). The vowels /i y u a/ were examined. Articulatory data were recorded using ultrasound imaging (for the tongue) and audiovisual recording (for the lips). Apart from the fact that blind children had larger acoustic and kinematic variability than their sighted peers, a result suggesting less mature motor control in the latter than in the former, blind children had reduced acoustic and articulatory contrasts (in the lip and tongue dimensions) between vowels. Moreover, the effect of focus was observed only in sighted children, who increased lip and tongue contrasts between the neutral and the focused condition. Blind children did not alter their lip and tongue positions under contrastive focus compared to the neutral condition. Of note, both groups used increased F0, intensity, and duration on the focused constituent. Unlike adults, intelligibility scores of vowels produced by blind children were significantly lower than those produced by their sighted peers (see Fig. 1), suggesting that their articulatory strategies were less efficient perceptually.

3. CONCLUSION

According to the Perception-For-Action-Control Theory (PACT) described in [40], speech goals correspond to multisensory perceptuo-motor units. In the course of speech development, perception and action are tightly linked, and speech perception necessarily involves procedural knowledge of speech production mechanisms. Furthermore, perceptual mechanisms provide gestures with auditory, visual, and somatosensory templates that guide and maintain their development. The fact that visual deprivation triggers different production strategies strongly supports the view that perception and production are co-structured. In the course of speech development, blind speakers do not integrate lip movements as a component of the speech task for some phonological features as strongly as sighted speakers do. Indeed, for the latter, seeing the lips might act as a constraint on lip movements; since this articulator has auditory and visual correlates (among others), its weight during speech development could be more important than that of less visible articulators such as the tongue. Blind speakers, in contrast, would not be affected by such constraints and articulatory movements would have comparable perceptual correlates. In early childhood, the effects of blindness likely reduce speech intelligibility. This pattern is no longer observed in adulthood. Congenitally blind individuals thus seem to find alternate paths to intelligible speech [31]. Further studies are currently underway to examine the interplay of higher cognitive factors and sensory input in both populations.

4. ACKNOWLEDGEMENTS

We thank all participants and members of the Laboratoire de phonétique for their precious contribution to this work. Thanks to Marlene Busko for copy-editing the paper.

5. REFERENCES

[1] Arnold, P., Hill, F. 2001. Bisensory augmentation: A speechreading advantage when speech is clearly audible and intact. *British Journal of Psychology*, 92, 339–355.

[2] Burlingham, D. 1961. Some notes on the development of the blind. *Psychoanalytic Study of the Child*, 16, 121–145.

[3] Collignon, O., Vandewalle, G., Voss, P., Albouy, G., Charbonneau, G., Lassonde, M., et al. 2011. Functional specialization for auditory-spatial processing in the occipital cortex of congenitally blind humans. *PNAS USA*, 108, 4435–4440.

[4] Elstner, W. 1983. Abnormalities in the verbal communication of visually-impaired children. In A.E. Mills (Ed.) *Language acquisition in the blind child* (pp.18–41). London: Croom Helm.

[5] Galantucci, B., Fowler, C. A., Turvey, M. T. 2006. The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review*, 13, 361–377.

[6] Gautam, A., Naples, J. G., Eliades, S. J. 2019. Control of speech and voice in cochlear implant patients. *The Laryngoscope*, DOI: 10.1002/lary.27787

[7] Gentilucci, M., Bernardis, P. 2007. Imitation during phoneme production. *Neuropsychologica*, 45(3):608–15.

[8] Ghosh, S., Matthies, M., Maas, E., Hanson, A., Tiede, M., Ménard, L., et al. 2010. An investigation of the relation between sibilant production and somatosensory and auditory acuity. *J. Acoust. Soc. Am.* 128, 3079–3087.

[9] Gick, B. and Derrick, D. 2009. Aero-tactile integration in speech perception. *Nature*, 462, 502-504

[10] Gori, M., Del Viva, M., Sandini, G., Burr, D. C. 2008. Young Children Do Not Integrate Visual and Haptic Form Information. *Current Biology*, 18, 694-698.

[11] Gori, M., Sandini, G., Martinoli, C., Burr, D. 2010. Poor Haptic Orientation Discrimination in Nonsighted Children May Reflect Disruption of Cross-Sensory Calibration, *Current Biology*, 20, 223-225.

[12] Gougoux, F., Lepore, F., Lassonde, M., Voss, P., Zatorre, R. J., & Belin, P. 2004. Pitch discrimination in the early blind. *Nature*, 430, 309-310.

[13] Hicock, G., Poeppel, D. 2007. The Cortical Organization of Speech Processing, *Nat. Rev. Neurosc.*8, 393-402.

[14] Hugdahl, K., Ek M., Takio F., Rintee T., Tuomainen J., Haarala C., et al. 2004. Blind individuals show enhanced perceptual and attentional sensitivity for identification of speech sounds. *Cognitive Brain Research*, 19, 28–32.

[15] Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. 2002. Hearing sounds, understanding actions: Action representation in mirror neurons. *Science*, 297, 846–848.

[16] Kuhl, P. K., Meltzoff, A. N. 1982. The bimodal perception of speech in infancy. *Science*, 218 (4577), 1138–1141.

[17] Legerstee, M. 1990. Infants use multimodal information to imitate speech sounds. *Infant Behavior*

and Development, 13, 343–354.
[18] Lewis, M.M. 1975. *Infant speech: A study of the beginnings of language* New York: Arno Press.
[19] Lucas, S. A. 1984. Auditory discrimination and speech production in the blind child. *International Journal of Rehabilitation Research*, 7(1), 74–76.
[20] Maeda, S. 1990. Compensatory articulation during speech: Evidence from the analysis and synthesis of vocal-tract shapes using an articulatory model. In W. J. Hardcastle & A. Marchal (Eds.), *Speech Production and Speech Modelling* Dordrecht: Kluwer Academic, 131–149.
[21] Mcconachie, H. 1990. Early language development and severe visual impairment. *Child: Care, Health, and Development*, 16(1): 55–61.
[22] McGurk, H., MacDonald, J. 1976. Hearing lips and seeing voices. *Nature*, 264, 746–748.
[23] Ménard, L., Schwartz, J.-L., Boë, L.-J., Aubin, J. 2007. Production-perception relationships during vocal tract growth for French vowels: analysis of real data and simulations with an articulatory model, *Journal of Phonetics*, 35 (1), 1-19.
[24] 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.
[25] Ménard, L., Toupin, C., Baum, S. R., Drouin, S., Aubin, J., Tiede, M. 2013. Acoustic and articulatory analysis of French vowels produced by congenitally blind adults and sighted adults, *J. Acoust. Soc. Am.*, 134 (4), 2975-2987.
[26] Ménard, L., Trudeau-Fisette, P., Côté, D. 2016. Speaking clearly for the blind: Acoustic and articulatory correlates of speaking condition, *PlosOne.*
[27] Miller, L. 1992. Diderot reconsidered: Visual impairment and auditory compensation. *Journal of Visual Impairment & Blindness*, 86(5), 206–210.
[28] Mills, A. E. 1983. *Language acquisition in the blind child: Normal and deficient*. San Diego, CA: College-Hill Press.
[29] Mills, A. E. 1987. The development of phonology in the blind child. In B. Dodd *&* R. Campbell (eds.), *Hearing by eye: The psychology of lip-reading* London: Lawrence Erlbaum Associates, 145–163.
[30] Niemeyer, W., & Starlinger I. (1981). Do the blind hear better? Investigations on auditory processing in congenital or early acquired blindness. II. Central functions. *Audiology*, *20*, 510–515.
[31] Pérez-Pereira, M., Conti-Ramsden, G. 2015. *Language Development and Social Interaction in Blind Children*, New York: Psychology Press.
[32] Perkell, J. S., Nelson, W. L. 1985. Variability in production of the vowels /i/ and /a/. *J. Acoust. Soc. Am.,* 77, 1889–1895.
[33] Perkell, J. S., Matthies, M. L., Svirsky, M. A., Jordan, M. I. 1993. Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot "motor equivalence" study. *J. Acoust. Soc. Am.*, 93, 2948–2961.
[34] Perkell, J. S., Matthies, M. L., Tiede, M., Lane, H., Zandipour, M., Marrone, N., et al. 2004. The distinctness of speakers' /s/–/s/ contrast is related to their auditory discrimination and use of an articulatory saturation effect. *J Speech Lang Hear Res*, 47, 1259–1269.
[35] Prost, V., Espesser R., Sabater C., Thomas-Bartalucci K., Rey V. 2002. *Entraînement de la conscience phonologique d'enfants déficients visuels: Quel support temporophonologique?* XXIVèmes Journées d'Étude sur la Parole, Nancy, France.
[36] Pulvermuller, F., Huss, M., Kherif, F., Moscoso del Prado Martin, F., Hauk, O., Shtyrov, Y. 2006. Motor cortex maps articulatory features of speech sounds. *PNAS*, 103, 7865–7870.
[37] Reynell J. 1978 Development patterns of visually handicapped children. *Child: Care, health and development* 4, 291-303.
[38] Rosemblum, L.D., Schmuckler M.A., Johnson, J. A. 1997. The McGurk in infants. *Perception & Psychophysics*, 59(3), 347–357.
[39] Sams, M., Mottonen, R., Sihvonen, T. 2005. Seeing and hearing others and oneself talk. *Cognitive Brain Research*, 23, 429–435.
[40] Schwartz, J.-L., Basirat, A., Ménard, L., Sato, M. 2012. The Perception for Action Control Theory (PACT): A perceptuo-motor theory of speech perception. *J. Neurolinguistics*, 25, 336-354.
[41] Skipper, J. I., Nusbaum, H. C., Small, S. L. 2005. Listening to talking faces: Motor cortical activation during speech perception. *NeuroImage*, 25, 76–89.
[42] Skipper, J. I., Van Wassenhove, V., Nusbaum, H. C., Small, S. L. 2007. Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, 17, 2387–2399.
[43] Stankov, L., Spilsbury, G. 1978. The measurement of auditory abilities of blind, partially sighted, and sighted children. *Applied Psychological Measurement*, 2(4), 491–503.
[44] Starlinger, I., Niemeyer, W. 1981. Do the blind hear better? Investigations on auditory processing in congenital or early acquired blindness. I. Peripheral Functions. *Audiology*, 20, 503–509.
[45] Sumby, W.H., Pollack I. 1954. Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.*, 26(2), 212–215
[46] Summerfield, Q. 1979. Use of visual information for phonetic perception. *Phonetica*, 36, 314–331.
[47] Thomas, K., Prost V., Espesser R., Rey V. 2000. *Capacités phonologiques implicite et explicite chez les malvoyants.* Presented at the 13e Journées d'Étude sur la Parole, Aussois, France.
[48] Warren, D.H. 1977. *Blindness and early childhood development*. New York: American Foundation for the Blind.
[49] Wilson, S. M., Iacoboni, M. 2006. Neural responses to non-native phonemes varying in producibility: Evidence for the sensorimotor nature of speech perception. *NeuroImage*, 33, 316–325.