The Influence of Familial Sinistrality on Audiovisual Speech Perception

Sandhya Vinay, Dawn M. Behne

Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway
sandhya.vinay@ntnu.no, dawn.behne@ntnu.no

Abstract

The present study investigates the influence of familial sinistrality on audiovisual speech perception for young adults. Incongruent video stimuli dubbed over dichotic and diotic audio stimuli were utilized. Participants' responses for the right incongruent, left incongruent and video incongruent audiovisual stimuli were analyzed. Results indicated significantly higher proportion of fusion responses demonstrating increased audiovisual interaction in participants with familial sinistrality compared to those without familial sinistrality. Thus, the presence of familial sinistrality influences the role of visual information on the processing of audiovisual speech.

Index terms: Familial sinistrality, audio-visual, incongruence, fusion

1. Introduction

The influence of visual speech on the perception of auditory stimuli has been demonstrated in various contexts [1, 2, 3]. The "McGurk effect" [2, 4], which produces a percept of /da/ for a presentation of auditory /ba/ with a video /ga/ is a well-documented demonstration of the interaction between the auditory and visual modalities for audiovisual (AV) speech perception [5].

Functional differences between the two cerebral hemispheres were observed early neuropsychological research in neuropsychology [6, 7, 8]. The phenomenon of cerebral dominance has long been studied in relation to the hemispheric distribution of language, music and handedness in the human brain [6, 9, 10, 11]. Right ear advantage in dichotic listening tasks has been used to demonstrate left hemispheric dominance for speech [12, 13]. Right ear advantage makes the McGurk effect harder to induce, that is, influences audiovisual integration of speech, when the audio stimulus is presented to the subject's right ear [14]. The McGurk effect is testable by implying simple behavioral methods where the participants can be instructed to report what they perceive and their responses can thus be compared to the specific characteristics of the stimuli [15].

Another aspect, familial sinistrality (FS), a notion that encompasses having one or more left-handers among one's close relatives, has also been shown to have an impact on lateralization for language [16]. Occurrence of FS is reported to be either similar between right- or left-handed individuals [17] or slightly more frequent in left-handed ones [18]. According to some theories, right-handed individuals without left-handed relatives (without FS) should be the most strongly lateralized participants for language functions, whereas left-handed individuals with left-handed relatives (with FS) should be the most ambilateral ones [19, 20].

The present study aims to investigate the possible influence of FS on AV speech perception (and cerebral dominance) using dichotic and diotic audio stimuli with incongruent video stimuli which produce a McGurk effect. The dichotic stimuli are used to throw light on cerebral dominance for speech and the incongruent video component to determine AV interaction.

2. Method

2.1. Participants

Forty right-handed young adults, 20 women and 20 men (mean = 23 years, SD = 2.24 years) with Norwegian as their native language voluntarily participated in the study. Twelve participants (5 women and 7 men) reported the presence of familial sinistrality (FS+). FS was determined based on the presence of left-handedness in immediate family members (parents and/or siblings). All participants had bilateral normal hearing sensitivity with audiometric thresholds equal to or better than 20 dB hearing level (HL) for octave frequencies from 0.25 to 8 kHz. All participants had normal or corrected to normal vision with no history of color blindness or other disorders related to visual acuity. None of the participants reported of any neurological, speech and language, attention or motor disorders. The present study was approved by the Norsk Samfunnsvitenskapelig Datatjeneste (Norwegian Social Science Data Services, NSD).

2.2 Speech material / Stimuli

A young female Norwegian speaker was audio and video recorded uttering the syllables /ba/, /pa/, /ga/ and /ka/. PDWF800 Sony Professional XDCAM HD422 camcorder and two Røde NT1-A microphones were used for recording the stimuli. The audio recording of each syllable was dubbed on the video recordings to obtain incongruent audiovisual (AV) stimuli as shown in table 1. The four auditory and four visual tokens were paired to create four AV congruent stimuli and 24 stimuli with three types of AV incongruence: 1) Eight dichotic audio stimuli with video congruent to left ear and incongruent to right ear (referred to as Right incongruent), 2) Eight dichotic audio stimuli with Video congruent to the right ear and incongruent to the left ear (referred to as Left incongruent), 3) Eight diotic audio stimuli with incongruent Video (referred to as Video incongruent).

Table 1: *Examples of AV stimuli used in the present study.*

AV incongruence	Audio segment		Video segment
	left	right	
Video incongruent	/ba/	/ba/	/ga/
Right incongruent	/ba/	/ga/	/ba/
Left incongruent	/ba/	/ga/	/ga/
Congruent	/ba/	/ba/	/ba/

The audio and video editing was performed in PRAAT and AVID media composer. The video and audio stimuli in the two channels were synchronized by temporally aligning the energy release (burst) in the consonant segment of the stimuli. Each AV stimulus, 1000ms long, started and ended with a neutral face and lips closed.

2.3. Procedure

The experiment was set up and data were collected using Superlab (v. 4) run on a iMac11.3 in the Speech Lab, Department of Psychology at the Norwegian University of Science and Technology (NTNU), Trondheim, Norway. The auditory stimuli were presented through K271 studio headphones at approximately 64 dBA. Participants were seated facing the monitor (1920 x 1200 pixels) at a distance of approximately 70 cm. The experiment was run with one participant at a time. The experiment commenced with example trials, followed by experimental trials where three repetitions of each AV stimulus were randomly presented to each participant. Control stimuli were random presentations of Congruent AV stimuli. Videos were always presented in the center of the monitor. The task of the participants was to press keys on a Cedrus RB-730 response box by choosing from the options /pa/, /ta/, /ka/, /ba/, /da/ and /ga/, based on the syllables perceived.

3. Results

Participants´ responses for each stimulus were categorized as matching place of articulation and voicing with respect to the audio and video components of the stimuli. For the Video incongruent stimuli, the responses were categorized as audio place match (Apm), video place match (Vpm) and fusion (Fp). For the Right incongruent and Left incongruent stimuli, the responses were categorized as left audio place match (Lpm), right audio place match (Rpm) and fusion (Fp). With these dependent variables, repeated measures analyses of variance (ANOVA) were performed on the responses of the participants with FS and AV incongruence as independent factors.

As shown in Figure 1, for the Video incongruent stimuli with bilabial audio and velar video ($A_{bil}V_{vel}$), results (figure 1) reveal a significantly higher proportion of fusion responses for participants with FS than for participants without FS [$F(1,36)=6.52$, $p<0.05$)]. Correspondingly, participants with FS had significantly fewer auditory place match responses than participants without FS [$F(1,36)=5.06$, $p<0.05$)], demonstrating the role of FS on audio-visual interaction in speech perception. However, participants with FS and without FS did not differ from each other for the Video place match responses. The effect of FS on audio place match and fusion responses was not observed for the responses to Video incongruent stimuli with velar audio and bilabial video ($A_{vel}V_{bil}$).

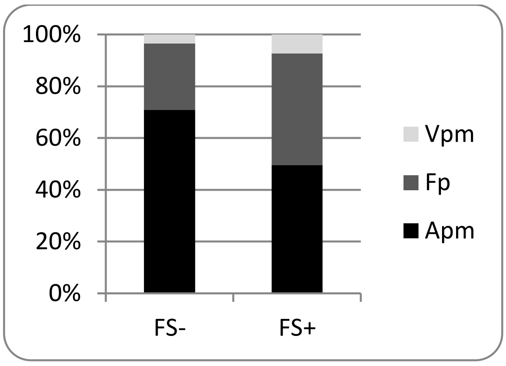

Figure 1: *The proportions of audio place match (Apm), video place match (Vpm) and fusion (Fp) responses for Video incongruent stimuli in participants with familial sinistrality (FS+) and without familial sinistrality (FS-).*

Participants responses for the video incongruent stimuli were also analyzed to examine the effect of voicing (audio segment of the AV stimuli) on the audio place match and fusion responses. As shown in Figure 2, a significant effect of voicing (audio segment of the AV stimuli) on audio place match and fusion responses was observed for the Video incongruent ($A_{bil}V_{vel}$) stimuli. Voiceless audio stimuli resulted in higher auditory place match responses than voiced audio stimuli [$F(1,36) = 37.34$, $p<0.05$)]. Fusion responses were higher for Video incongruent stimuli with voiced audio than voiceless audio [$F(1,36) = 17.77$, $p<0.05$)].

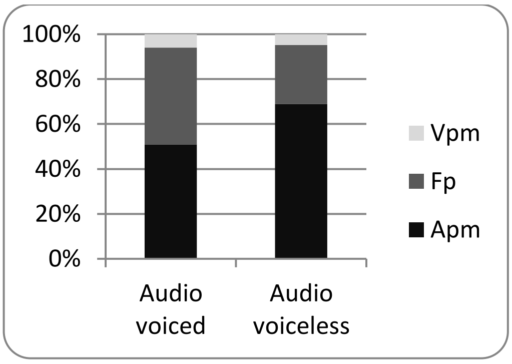

Figure 2: *Proportions of audio place match (Apm), video place match (Vpm) and fusion (Fp) responses for audio voiced and audio voiceless Video incongruent stimuli.*

Analysis of responses for Right incongruent and Left incongruent stimuli with velar video segment, illustrated in Figure 3, showed a significant main effect of AV incongruence on the left audio place match [$F(1,36)=76.56$, $p<0.05$)] and right audio place match responses [$F(1,36) = 49.87$, $p<0.05$)]. The fusion responses were not significantly different between the Right incongruent and Left incongruent stimuli.

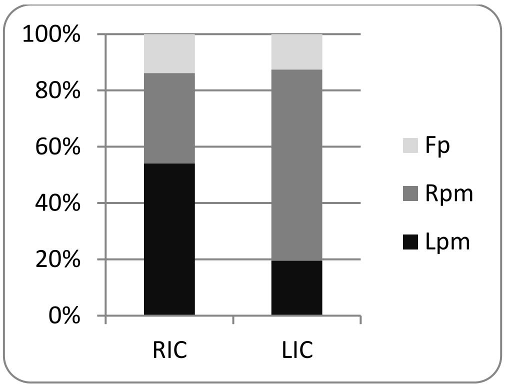

Figure 3: *The proportion of audio place match (Apm), video place match (Vpm) and fusion (Fp) responses for Right incongruent and Left incongruent stimuli*

The responses for Right incongruent and Left incongruent stimuli also revealed a significant interaction between FS and AV incongruence. As is shown in Figure 4, participants with FS had a significantly higher proportion of fusion responses than participants without FS for Right incongruent stimuli [$F (1, 36) = 6.53$, $p<0.05$)], demonstrating more fusion between the Video and Audio segment presented to the right ear. Participants with and without FS had similar fusion responses for Left incongruent stimuli.

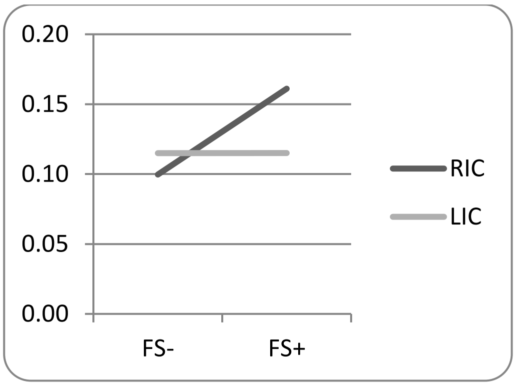

Figure 4: *Interaction between familial sinistrality (FS) and Right incongruent - Left incongruent stimuli for fusion responses in participants with familial sinistrality (FS+) and without familial sinistrality (FS-).*

For the Right incongruent and Left incongruent stimuli with bilabial video segments, an effect of FS on left place match responses was observed [$F(1,36) = 6.19$, $p<0.05$)]. Participants with FS had higher left place match responses than participants without FS. The proportion of right place match and Fusion responses were similar for participants with and without FS.

4. **Discussion**

The present study discusses the role of FS and its interaction with stimuli related incongruence factors on the perception of AV speech stimuli. The McGurk stimuli (diotic bilabial audio with incongruent velar video) stimuli resulted in significantly higher fusion responses for participants with familial sinistrality than participants without familial sinistrality. Audio place match responses were significantly lower for participants with familial sinistrality compared to those without familian sinistrality. .In other words, the influence of the visual signal on AV speech perception was more apparent for individuals with FS, even though the video place match response did not significantly differ between the participant groups.

Fusion responses were higher for video incongruent stimuli with voiced audio segment than those with voiceless audio segment, which demonstrate the participation of the right hemisphere in the perception of voicing of the audio stimuli. These findings are consistent with the previous research showing that the right hemisphere is more involved in the processing of voicing than place of articulation for stop consonants [21].

For AV incongruent stimuli with dichotic audio presentation, perception of place of articulation of the audio segment was more accurate when the audio was concordant with the video, irrespective of ear. The results are in agreement with previous studies which have shown that visual stimuli congruent to one ear in a dichotic presentation reduce the right ear advantage [14], otherwise reported for the dichotic stimuli [10, 13]. General equal distribution of fusion responses for Right incongruent and Left incongruent stimuli suggests that concordant video enhances the audio response irrespective of the hemisphere receiving the congruent AV stimuli, thereby not resulting in a difference in audiovisual interaction between the two ears. However, participants with FS had higher fusion responses when the AV incongruent stimuli were received by the dominant left hemisphere compared to the right hemisphere, a difference that was not observed in participants without FS. These findings suggest lesser left hemisphere dominance for audiovisual speech perception in participants with FS [16].

5. Conclusions

The higher proportion of fusion responses, that is, the percept generated by the interaction of audio and visual segments in participants with familial sinistrality compared to participants without familial sinistrality demonstrates the increased influence of visual information on audiovisual speech perception. The present study is an attempt to understand the complex interaction between familial sinistrality, hemispheric dominance and acoustical characteristics of speech stimuli that needs further investigation.

6. References

[1] R. N. Desjardins, J. Rogers, and J. F. Werker, "An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks,".*Journal of Experimental Child Psychology*, vol. 66, no. 1, pp. 85-110, 1997.

[2] H. McGurk, and J. MacDonald, "Hearing lips and seeing voices," *Nature*, vol. 264, 746-748, 1976.

[3] W. Sumby, and I. Pollack, "Visual contribution to speech intelligibility in noise," *Journal of Acoustical Society of America*, vol. 26, pp. 212-215, 1954.

[4] J. MacDonald, and H. McGurk, "Visual influences on speech perception processes," *Attention, Perception, & Psychophysics*, vol. 24, no. 3, pp. 253-257, 1978.

[5] Q. Summerfield, *Audio-visual speech perception, lipreading and artificial stimulation.* In: Lutman ME, Haggard MP (eds) *Hearing Science and Hearing Disorders*. London: Academic press, pp. 131-182, 1983.

[6] N. Geschwind, and W. Levitsky, "Human brain: left-right asymmetries in temporal speech region," *Science,* vol. 161, no. 3837, pp. 186-187, 1968.

[7] R. C. Gur, and R. E. Gur, "Handedness, sex, and eyedness as moderating variables in the relation between hypnotic susceptibility and functional brain asymmetry," *Journal of Abnormal Psychology*, vol. 83, no.6, pp. 635-643, 1974.

[8] J. McGlone, "Sex differences in functional brain asymmetry," *Cortex*, vol. 14, no. 1, 122-128, 1978.

[9] T. G. Bever, and R. J. Chiarello, "Cerebral dominance in musicians and non-musicians," *Science,* vol. 185, pp. 537-539, 1975.

[10] D. Kimura, "Cerebral dominance and the perception of verbal stimuli," *Canadian Journal of Psychology*, vol. 15, pp. 166-171, 1961.

[11] M. Annett, "Handedness and cerebral dominance: the right shift theory", *The Journal of Neuropsychiatry and Clinical Neurosciences*, vol. 10, no. 4, pp. 459-469, 1998.

[12] D. Kimura, "Functional asymmetry of the brain in dichotic listening," *Cortex,* vol. 3, pp. 163-178, 1967.

[13] K. Hugdahl, "What can be learned about brain function from dichotic listening," *Rev Esp Neuropsicol*, vol. 2, no. 3, pp. 62-84, 2000.

[14] M. Scott, "The McGurk effect affected by the right ear advantage," *Canadian Acoustics*, vol. 36, no. 3, 156-157, 2008.

[15] N. Bedoin, E. Ferragne, and E. Marsico, "Hemispheric asymmetries depend on the phonetic feature: A dichotic study of place of articulation and voicing in French stops," *Brain and language*, vol. 115, no. 2, pp. 133-140, 2010.

[16] N. Tzourio-Mazoyer, L. Petit, A. Razafimandimby, F. Crivello, L. Zago, G. Jobard,... and B. Mazoyer, "Left hemisphere lateralization for language in right-handers is controlled in part by familial sinistrality, manual preference strength, and head size," *Journal of Neuroscience,* vol. 30, no. 40, pp. 13314-13318, 2010.

[17] B. J. Spiegler, and G. H. Yeni-Komshian, "Incidence of left-handed writing in a college population with reference to family patterns of hand preference," *Neuropsychologia*, vol. *21, no.*6, pp. 651-659, 1983.

[18] D. L. Orsini, P. Satz, H. V. Soper, and R. K. Light, "The role of familial sinistrality in cerebral organization," *Neuropsychologia*, vol. 23, no. 2, pp. 223-232, 1985.

[19] C. Hardyck, and L. F. Petrinovich, "Left-handedness," *Psychological bulletin*, *84*(3), 385, 1977.

[20] H. Hécaen, M. De Agostini, and A. Monzon-Montes, "Cerebral organization in left-handers," *Brain and language*, vol. 12, no. 2, pp. 261-284, 1981.

[21] N. Bedoin, E. Ferragne, and E. Marsico, "Hemispheric asymmetries depend on the phonetic feature: A dichotic study of place of articulation and voicing in French stops," *Brain and language*, vol. 115, no. 2, pp. 133-140, 2010.