Impaired Speechreading Related to Arrested Development of Face Processing

[1]Beatrice de Gelder, [1]Jean Vroomen and [2]Bruno Laeng

[1]Tilburg University, Dept of Psychology
The Netherlands
[2] University of Tromsoe, Norway

ABSTRACT

Childhood prosopagnosia is a rare disorder of face recognition consisting in the loss of one or more aspects of normal visual processing for faces as a consequence of a brain injury not present at birth. The paper reports on the case of patient, RP, a child prosopagnosic who sustained a closed head injury at age 7. Different aspects of speechreading ability are examined both with static and dynamic stimuli (recognition of spoken vowels and digits, serial recall). Moreover, two bimodal tasks were used in order to assess the combination of auditory speech with speechreading. The results show that RP is performing poorly on all of the visual tasks, while his auditory performance is entirely normal. This dissociation between auditory and visual speech skills replicates similar findings with adult prosopagnosics and underscores the importance of intact face processing for the development of normal speechreading

1. INTRODUCTION

Perceiving a face initiates a series of processes that presumably address different types of information like gender, age, personal identity, expression, or speech. In line with the standard view [1] in which separate, autonomous routes for each of these processes are assumed, earlier studies have tended to report dissociations between the different routes when a breakdown in face processing ability occurs as a consequence of braindammage. More recently, though, several studies have begun to challenge this purported autonomy. The processing of expression and that of identity may not be as autonomous as originally thought. As concerns the link between face processing and speechreading, there are now reports showing that face processing disorders acquired in adulthoood entail a partial or complete loss of speechreading ability [2, 3, 4].

The latter issue is particularly intriguing when put in a developmental perspective. One of the major reasons why faces are special is derived from the argument that face expertise develops over time and only reaches full capacity after puberty. What implications does this have for the various components of face processing ability, and more specifically, for speechreading? Is there also a developmental course for speechreading ability [e.g. 5], and if so, does it follow the same time course as the acquisition of face expertise? Or, speechreading being a linguistic skill, does it follow its own course, more related to the maturation of speech processing than anything else?

In all the patients suffering from face disorders whose speechreading abilities have been examined, brain injury occurred in adulthood. Here we present the case of RP, a densely prosopagnosic patient whose problems can be traced back to an injury when he was a 7 year old child. The interest of this case resides in the fact that it offers a window on the interdependency in the course of development between the two abilities that enter into normal speechreading, i.e. face processing and speech. The latter is fully developed by age 7 while the former continues to grow till puberty [6]. On the other hand, the only developmental evidence about impaired speechreading available so far links impaired speechreading to impaired phonological processes in dyslexics [7].

RP was presented a series of speechreading tasks requiring recognition of lipshapes from still pictures, speechreading with dynamic displays, as well as bi-modal stimuli. Tasks included recognition of lip shapes, of silently spoken single digits, and serial recall of audio-only, visual-only, and audio-visually presented digits as well as tasks examining audiovisual speech perception. As we shall see, on all of these tasks RP's performance was well below that of normal controls.

Case presentation. RP is a 49 year old male native of Massachusetts who sustained a closed head injury at age 7. Following his accident he had a normal school career and achieved a degree in engineering at a college in the Boston area. He has a Verbal IQ of 128 and a performance IQ of 98 (WAIS). Block Construction was above average. Performance on the Hooper Visual Organization test was average. In tasks like reverse counting, alphabet recitation and serial computation, RP performed at maximum level. On the Wechsler Memory scale - Revised, RP's performance was average for stories but higher than average for designs at a retention of 30 min. Similarly, in the Trailmaking test his performance was average. Performance in the Visual Search and Attention test was within normal limits, whereas it was in the superior range for the Raven Matrices. An MRI scan did not show any abnormalities.

RP is specifically impaired in all aspects of face processing without any evidence of abnormal object recognition. He has no recognition of any familiar face, neither from real photographs nor from caricatures. He can make gender decisions more or less correctly and presumably bases his judgement on external facial cues. He does not recognize cartoon figures but does sometimes make a right guess about the animal species on which the cartoon is based (rabbit, pig). His performance in a delayed matching task using unfamiliar faces is at chance but with simultaneous presentation he performs near ceiling, level though very slowly. He does not appear to show either the normal face inversion or face context effects suggesting that his processing of the facial configuration may not be intact. RP has no recognition of facial expressions nor has he any mental imagery in this domain. Nevertheless, RP recognizes a face as being a face. For the time being we can only conclude that the procedures on which he relies in making a facial decision may not be those used by normal subjects. It is worth noting that RP's language production is intact but his reading is slower than normal.

2. METHOD AND RESULTS

Recognition of speech sounds from still photographs.
Testing of RP's speechreading abilities started with the simplest but also the most contentious task, that of speechreading from still photographs. RP was presented with a series of 16 black and white photographs. They represented four different actors each with four different mouth positions (saying /a/, /i/, /o/ or making a grimace). He was given the pictures one by one and told about the four response choices which were written down on cards in front of him. He was asked to put each photograph down next to what he deemed to be the correct response. After each choice the picture was taken away in order to avoid confusion from previous (mistaken) category assignments.

RP completely failed this task and could not reliably recognize the speech sounds in any of the photographs. Such total inability to process any facial gestures whether corresponding to emotional expressions or to speech has been observed in other cases of prosopagnosia [3, 4]. Testing with dynamic stimuli can improve performance for speechreading as well as for recognition of facial expressions. Will RP perform better when presented with dynamic information provided by short video clips?

Recognizing spoken digits
RP was shown short video clips of a female speaker articulating one by one the digits 1, 2, 4, 5, 6, 7, 8, or 9 in random order with each digit presented twice. Tests with normal controls had shown that these digits were clearly speech-readable. RP's performance was surprisingly poor and performed at chance (2 out of 16 trials correct)

Serial recall of audio-visual, auditory and speech-read digits.
A videotape was constructed of the same female speaker used in the single digit task, but this time she pronounced digit lists. The video was then edited so that the digit lists were presented in an audio-visual mode, an audio-only mode in which just the sound was heard, but with a blank screen, and a speechreading mode in which the speaker was visible, but without sound. Each list was composed by drawing without replacement from the numbers 1, 2, 4, 5, 6, 7, 8, and 9 in pseudo-

random order. Each list consisted of eight digits. Each presentation mode consisted of one block of eight digit lists. The audio-visual block was presented first, followed by the audio-only block, and then the speech-read block. Instructions specified to attend to the auditory, the audiovisual and the visual input respectively, and to report the digits in the order as presented. With the auditory and audio-visual presentation, he was correct, in both cases, at 80%, but with the speechread digits, RP's performance was at 0%.

Auditory processing, speechreading and audio-visual conflict.

We used a video recording of a female speaker pronouncing a series of VCV sequences [3,7]. Each sequence consisted of one of the four plosive stops /p, b, t, d/ or a nasal /m, n/ in between the vowel /a/ (e.g., /aba/ or /ana/). There were three presentation conditions: an audio-visual, an auditory-only, and a visual-only presentation. In the audio-visual presentation, dubbing operations were performed on the recordings so as to produce a new video-film comprising six different auditory-visual combinations: auditory /p, b, t, d, m, n/ were combined with visual /t, d, p, b, n, m/, respectively. The visual place of articulation feature thus never matched the auditory place feature. Appropriate dubbing ensured that there was auditory-visual coincidence of the release of the consonant in each utterance. In addition, there were unimodal presentation conditions. For the audio-only condition, the original auditory signal was dubbed on a blank screen. For the visual-only condition, the auditory channel was deleted from the recording, so the subject had to rely entirely on speech reading. Each presentation condition comprised three replications of the six possible stimuli. RP was instructed to watch the speaker and repeat what she said. In the auditory-only condition, he repeated 13 out of 18 trials exactly, but auditory /m/ was heard as /n/. In the speechread condition, his answer matched the place of articulation of the visual stimulus in 7 out 18 trials, but mainly because he reported to hear a bilabial (/m/) most of the time. With the audio-visual presentation, the visual input influenced his response on 5 out 18 trials, but mainly because auditory /m/ combined with visual /n/ was perceived as /n/. Since auditory-only /m/ was also heard as /n/, it is difficult to determine whether this effect is due to a genuine cross-modal bias.

3. CONCLUSION

Our study of RP's speechreading shows his severe deficit in this domain. RP's total inability to gather information about speech from watching lip movements stands in stark contrast to his normal ability to process auditory speech and to maintain auditory representations in memory. RP's memory span for spoken digits is in the top range compared to that of normal adults (see [4]). This suggests that his phonological memory for spoken input has developed normally post trauma.

Assuming that before his accident RP performed at a level normal for his age, RP's problems with faces suggest that there are clearly limits to the plasticity of brain function, more specifically to the extent to which the effects of a childhood trauma can be overcome. The fact that the problems are confined to facial information testifies to the domain specificity of his disorder and suggests that this domain specificity was already present at the time of the head injury. Moreover, there are severe limits to the degree to which such patients can acquire compensation strategies. Our results show that notwithstanding a normal development of spoken language and phonological memory, no recovery of function has taken place for face related skills. In other words, normal phonological processes for spoken language are not sufficient to drive even a partial development of speechreading skills.

Only a few cases of childhood visual recognition impairments have so far been documented in the literature. The only one that looked into speechreading skills in a patient suffering from childhood prosopagnosia is in [9] which is a follow up of the original report by [10]. Like RP, patient AB was not subject to the McGurk illusion and thus only reported the auditory information when presented with a conflicting audiovisual input.

4. REFERENCES

1. Bruce, V. and Young, A. "Understanding face recognition", *Brit. J. Psych., Vol. 77, 305-327, 1986.*

2. Campbell, R. "How brains see speech: The cortical localisation of speechreading in hearing people", In R. Campbell, B. Dodd, and D. Burnham (Eds.), *Hearing by eye II, Hove: Psychology Press, 177-194, 1998.*

3. de Gelder, B., and Vroomen, J. and Bachoud-Levi, A-C. "Impaired speechreading and audio-visual integration in prosopagnosia", In R. Campbell, B. Dodd, and D. Burnham (Eds.), *Hearing by eye II, Hove: Psychology Press, 195-210, 1998.*

4. de Gelder, B., and Vroomen, J. "Impaired speechreading in prosopagnosia", *Speech Communication, in press.*

5. Massaro, D., Thompson, L. A., Barron. B., and Laren, E. "Developmental changes in visual and auditory contributions to speech perception", *J. Exp. Child. Psy, Vol. 41, 93-113, 1986.*

6. Carey, S., and Diamond, R. "Are faces perceived as configurations more by adults than by children?", In V. Bruce, and G. Humphreys, (Ed.), *Object and face recognition. Special issue of Visual Cognition, Vol. 1, No. 2/3. 253-274, 1994.*

7. de Gelder, B., and Vroomen, J. "Impaired speech perception in poor readers: Evidence from hearing and speechreading", *Brain and Language, in press.*

8. de Gelder, B., Vroomen, J. and van der Heide, L. "Face recognition and lipreading in autism", *Eur. J. Cogn. Psy, Vol. 3, 69-86, 1991.*

9. de Haan, E. H. F. and Campbell, R. "A fifteen year follow-up of a case of developmental prosopagnosia", *Cortex, Vol. 27, 489-509, 1991.*

10. McConachie, H. R. "Developmental prosopagnosia: A single case", *Cortex, Vol. 12, 76-82, 1976.*