Impaired speechreading and audio-visual speech integration in prosopagnosia

de Gelder, Beatrice; Vroomen, Jeanne; Bachoud-Levi, A.-C.

Published in:
Hearing by eye II, Advances in the psychology of speechreading and auditory-visual speech

Publication date:
1998

CHAPTER TEN

Impaired speechreading and audio-visual speech integration in prosopagnosia

Beatrice de Gelder, Jean Vroomen and Anne-Catherine Bachoud-Levi *Department of Psychology, Tilburg University, Tilburg, The Netherlands*

INTRODUCTION

Visual agnosia is the condition following cortical insult that results in an inability to recognise visually presented objects. In severe cases all classes of visual object recognition may be impaired. There is debate concerning the loss of face processing ability in this context (Farah, 1991), but the general pattern is for face recognition skills to be very impaired (prosopagnosia) in most cases of visual object agnosia. Despite such profound loss, copying and imagery skills may be relatively normal, indicating a breakdown in the recognition mechanism that can dissociate from spared visual processing and spared visual knowledge. Speech is not usually construed as a "visual object" in the sense that an orange or an automobile—or even a facial identity or expression—may be. So might speechreading be spared in patients with visual agnosia?

In this chapter we report a case of dense visual agnosia affecting the processing of all visual material, including faces and written words, who could nevertheless perform some speechreading tasks. We speculate on the implications of this for modelling speechreading in neurophysiological systems.

SPEECHREADING IN VISUAL AGNOSIA: PREVIOUS CASES

Several cases have been reported that have a bearing on this question. One study which suggested that speechreading was a different sort of visual object than most others was the report of patient Mrs D (Campbell et al., 1986). She was densely agnosic with profound prosopagnosia, yet could sort pictures of faces according to speechsound and was sensitive to the effects of seeing the speaker in reporting heard speech (McGurk effects). She could speechread silent spoken numbers and could discriminate lipspoken vowels and consonants.

By contrast, patient Mrs T was unable to perform such tasks, although she had no difficulty recognising faces or facial expressions or other visual objects. Mrs T was also unable to read any word presented to her although she could spell words perfectly, a letter at a time, when she heard the word (pure alexia). Mrs T's lesion was unilateral and affected the left hemisphere, Mrs D's only affected the right. Patient HJA (Campbell, 1992) had bilateral lesions of occipito-temporal areas. He was unable to recognise faces or words presented to him and could not classify photographs of speaking faces. However, he was completely normal in speechreading moving faces and in showing McGurk effects for incongruent auditory-visual speech stimuli. This suggests that HJA's impairment was at a lower functional level than that of Mrs T: under certain visual conditions (moving stimuli) he could speechread, while Mrs T was unable to speechread even moving faces.

HJA's visual impairment affects all stationary objects—letters, faces, common objects, drawings—but spares moving ones, when tested (see also Humphreys et al., 1993). The critical importance of visual movement pathways to speechreading is illustrated by case LM (Campbell, 1996; Campbell et al., submitted). LM's lesion affects only the cortical visual movement areas, including area V5, and sparing areas V1–V4 which are all damaged in HJA. LM can only speechread in the sense of being able to classify still photographs of seen speech. She does not show McGurk effects and cannot read silently spoken speech.

Thus, studies reported to date suggest that visual speech may be construed as a visual object—but as a rather special one. First, some aspects of its recognition and classification call on specific language processing sites, probably in the left hemisphere (Mrs T). Second, speech may be more dependent than other visual objects on the integrity of visual movement processing: dynamic aspects of speechreading may be among its defining aspects (Rosenblum & Saldaña, this volume).

THE PRESENT STUDY

In this chapter we present data on speechreading from a visual agnosic patient, BC, who was densely prosopagnosic and alexic, but with normal auditory language skills, good drawing, copying and visual imagery abilities. We tested her ability to classify speech from still and moving faces and to integrate seen and heard speech. We finally tested her ability to remember seen speech.

CASE REPORT

BC suffered in May 1995 from a haematoma located across the left temporo-occipital sulcus, involving the middle occipital gyrus and the inferior temporal gyrus (Brodmann areas 18, 19, and 37). She presented with a right homonymous hemianopsia and showed a mild anomia, without any comprehension or repetition deficit, that subsided after some weeks. No other linguistic deficits were present, apart from a pure alexia. After some weeks, Goldmann perimetry showed a residual right para-central scotoma, which disappeared with IV/4 test. In December 1995 she suffered from a second, right-sided haematoma, almost symmetrical to the first. The lesion was centred on the middle occipital gyrus, just posteriorly to the temporo-occipital sulcus, involving area 19 and the white matter underlying area 18. After the occurrence of the second stroke, BC found herself unable to recognise familiar faces and common objects by sight, and complained of seeing the world in shades of grey. Goldmann perimetry showed a central scotoma with II/4 test. Visual evoked responses with black and white pattern were normal for latency and amplitude. She named about half of a set of black and white realistic drawings without any systematicity in subsets (Snodgrass & Vanderwart, 1980). When asked to name real objects presented by the examiner, she was 13/35 correct on visual presentation, claiming that she was unable to recognise the other items. Tactile naming of the same objects was flawless.

Since her injury, BC has suffered from visual agnosia, prosopagnosia, alexia, and achromatopsia. We have also reported some aspects of her object and face recognition impairments (de Gelder et al., 1996). Notwithstanding all these perceptual impairments, mental imagery was intact in all these domains (Bartolomeo et al., in press).

Early visual processing (i.e. from the retina to V1) is unimpaired when assessed by MRI, by normal visual evoked potentials and by her good performance on copying drawings and line orientation judgment. Her perfect performance on imagery tasks and tactile naming indicates that her knowledge of objects, letters, colours, and faces is preserved.

FACE PROCESSING

A number of aspects of face processing have been examined which we briefly review before presenting the evidence on speechreading.

Examination of face processing in BC started in March 1996 (three months after the second CVA) with subtests from the face recognition battery (Bruyer & Schweich, 1991). Structural encoding, as examined by the facial decision subtest, was clearly impaired. Her face/nonface decisions were based on recognition of individual features and noticing their incorrect location in the face. Gender and age decisions were at chance levels. Direction of gaze, examined with still pictures, was well below normal.

There was no indication of preserved face recognition in overt or covert tasks (de Gelder et al., in preparation). In contrast, facial imagery was preserved equally well for overall configurational properties of familiar faces and for specific facial features. BC could produce clearly recognisable drawings of faces as well as of parts of faces.

Expression recognition was severely impaired when tested with static pictures but was near normal when tested with short video clips showing the full transition from a neutral resting state to a full display of the emotion achieved after 5sec. Neither the initial states nor the end states presented in isolation could however be recognised. Categorisation of facial expressions was clearly impaired even for a continuum of face images from happy to sad. There was no indication of any language-related impairment except for reading.

SPEECHREADING

Speechreading from still photographs, in normal facial context, and in isolation

Testing started by administering the subtest "facial speech" from the prosopagnosia test battery (Bruyer & Schweich, 1991). Performance was within chance levels. Given this poor performance, we re-examined her with new materials. These consisted of 16 black and white photographs (4 actors, 4 mouth positions, i.e. saying /a/, /i/, /o/, or making a face). BC was given the pictures one by one and asked to put each one on the table under one of the three written labels for the vowel (which were repeatedly pronounced by the experimenter) or put it aside as a grimace. Her total score was 6/16 correct (chance is 4/16, þ2(1) = 1.33 p > .10). None of the subcategories was dealt with better than any other (6 photographs were assigned to "grimace" with 3/6 correct; 3 to /o/ with 1/3 correct; 5 to /i/ with 1/5 correct; 2 to /a/ with 1/2 correct).

Since BC had difficulties with facial decision and part/whole recognition tasks, there was reason to suspect that the full facial context might have a detrimental effect on the recognition of the lip shapes. We therefore administered the same task in a later session, this time presenting each of the photographs with a white mask over the face and a window made into it that left only the lip shapes and the area immediately around it visible. However, this appeared to make the task even more difficult. We must conclude that speechreading from still pictures is entirely inaccessible to BC. Her face recognition impairment seems to be so dense that mouth shapes cannot be identified. Even simple forced choice among limited alternatives fails to benefit her. But BC shows some preserved visual form abilities with face-like material. In the face decision task, non-face stimuli consisting of a facial contour with jumbled facial parts inside it were often correctly rejected on the basis of comments about the location of the mouth (e.g. when the mouth was drawn on the forehead). One possibility might be that it is difficult for BC to focus on the mouth region alone. This suggestion receives some support from the comments she gave when isolated mouth parts were shown for speech classification. BC indicated that she saw nothing in such displays, while with full faces at least some effort was made to decode the form of the lips. In general terms, BC resembles HJA who could not speechread from still faces. HJA did reasonably well with dynamic stimuli; would this be true for BC?

Speechreading from movement

BC shows greatly impaired perception of visual forms in relation to faces with very few spared discrimination abilities. However, some speechreading skill might still be present in BC because there was no indication that area V5 was damaged by her strokes and there was no clinical indication of movement perception disorder. For this purpose a videotape was made showing a female speaker (BdG) pronouncing a lists of vowel-varied or consonant-varied items. For the consonant-varied list, five visually distinct consonants embedded in VCV cluster were used: /apa, ata, aka, afa, awa/. Each of them was repeated six times in random order. The vowel-varied list required speechreading of vowels embedded in a CV syllable: /bi, ba, bu/. Each CV sequence was repeated six times in random order. The videotape was presented (no sound) for repetition by BC. Her result on the consonant-varied list was above chance level. The overall score was 14 out of 30 correct (chance level is 6 out of 30, þ2(1) = 13.33, $p < .001$). Most mistakes were made on speechreading /apa/(perceived as /awa/), and /aka/ (also perceived as /awa/). In speechreading /bi, ba, or bu/, 14 out of 18 trials were correct (chance level is 6 out of 18, þ2(1) = 16.00, $p < .001$). The errors were evenly distributed across the response categories: visual /bi/ was perceived once as / bu/, visual /ba/ was perceived once as /bu/, and there were two misses in speechreading /bu/. Compared to the recognition of static lip shapes, the dynamic speechreading task yields reasonable scores. Dynamic facial actions carry useable speech information, partly available to BC.

Audio-visual speech

In the following tasks, speechreading skill is examined in the context of auditory input. It remains to be seen then if the performance with unimodal dynamic stimuli can be maintained in bimodal conditions. The following reported tasks were given to BC a few weeks later. We administered three tasks designed to appreciate the separate processing of auditory and visual input as well as their combination, the conflicts, and the audio-visual conflicts.

Visual bias in audio-visual token identification

Massaro and Cohen (see Massaro & Cohen, 1983, experiment 2) have developed a videotape of dubbed, synchronised speech that measures the influence of vision on auditory categorisation. The auditory stimuli were a nine-step /ba/ to /da/ continuum. These nine auditory tokens were factorially combined with two possible visual articulations, /ba/ or /da/. These 18 trials represent the bimodal condition. There was also an auditory-only and a visual-only condition. In the auditory-only condition, one of the nine auditory stimuli was presented, but the speaker did not move his mouth. In the visual-only condition, the speaker articulated either /ba/ or /da/, but no auditory speech was presented. In this case, the subject had to rely entirely on lip-reading. In every block of 54 trials, there were 18 bimodal conditions, 18 auditory-only conditions, and 18 visual-only conditions, in random order. The experiment consisted of 5½ blocks of trials preceded by 10 practice trials, producing a total of 307 trials. Participants report whether the speaker said /ba/ or /da/. Previous studies have shown that visual information systematically affects the categorisation of the speech sound. Seeing "ba" moves the categorical boundary for the auditory series towards the "ba" end of the continuum, seeing "da" towards the "da" end. Unimodal presentations (i.e. auditory token with a still face and visual token without sound) are judged in their own right and offer a second metric for assessing the influence of the mixed modality. In control subjects, visual unimodal inputs tend to be correctly reported, while the auditory-only input constitutes a baseline for establishing the effect of the mixed-mode condition.

BC showed an unusual pattern. In speechreading visual-only /ba/ or /da/, she reported almost exclusively /ba/ (visual /ba/ was 96% of the time correctly perceived as /ba/, but visual /da/ was 86% of the time perceived as incorrect /ba/). So the total percentage of correct responses on the visual-only part of the test was only 55% (þ2(1) = 1.02, $p > .10$). A similar strange pattern was observed with the auditory-only and the audio-visual trials. On the auditory-only (with still silent face) trials, she reported on 93% of the trials /da/, independent of which stimulus from the auditory /ba/ to /da/ continuum was presented. On the audio-visual trials, she reported /ba/, 88% of the time, independent of what was heard or seen. It thus appeared that her response strategy was based on whether or not the lips moved. If the lips moved, as in the visual-only and audio-visual trials, then /ba/ was given as response. This response was given independent of whether /ba/ or /da/ was said by the speaker. If the lips did not move, as in the auditory-only trails, she responded /da/, independent of whether /ba/ or /da/ was said. There was no indication that this behaviour was in relation to a perseveration strategy.

In order to confirm this surprising behaviour, BC was tested two weeks later with the same videotape. Her response pattern on this occasion was exactly the same, except that the response labels were switched: in the visual-only and audio-visual trials she reported hearing /da/ (78% for visual-only, and 86% for the audio-visual trials), and in the auditory-only trials she reported hearing /ba/ on 90% of the trials. Both kinds of unimodal stimuli were judged very poorly. For the visual-only stimuli her responses appear not to be related to the phonemic information but to whether or not a cue to movement was present. More surprisingly even, a similar overall strategy seems to have generated the auditory responses (unimodal auditory tokens were dubbed to a still face). Findings from other testing sessions may throw some light on this particular result. First, there is no reason to suspect a deficit in auditory processing (timed semantic decisions on auditory stimuli were 100% correct and RTs well within the normal range). Second, in a task examining categorical perception of expressions presented bimodally, performance was less erratic but the categorisation of unimodal stimuli presented in the experiment was, as here, different from those obtained in a testing session where only visual stimuli were presented. The interesting finding is that where performance is tested on two modalities, one of which is impaired, there is interference with processing in the *good* modality. We keep this point in mind while testing it in a somewhat different way.

Audio-visual speech with an artificial synthesized face

The next task was administered a few weeks later and consisted of a videotape showing an artificially created synthesized face (Massaro & Cohen, 1990). One possibility in "visual bias in audio-visual token identification" is that BC was, for some reason, particularly affected by some idiosyncrasy of the speaker's face. Another, of course, is that the still face interfered with her auditory processing in an abnormal way. The material used in the present test first established how general these effects might be across different speaking faces. Because the audio-only token occurs without a visual display in this videotape, her auditory report may be less affected than in the previous test.

The synthetic face is controlled by 11 display parameters which determine jaw rotation, lip protrusion, upper lip raise, etc. By varying these parameters, a mobile image of a face is created that articulates "ba", "da" or any intermediate position between these two syllables. In the test, five levels of audible speech varying between "ba" and "da" were crossed with five levels of visible speech varying between "ba" and "da". These 25 stimuli comprise the audio-visual condition. The auditory and visual stimuli were also presented alone, so that there was a total of 25 + 5 + 5 = 35 independent stimulus events. The whole test consisted of 6 sets of these 35 trials in which the order of items was randomised.

The performance of BC was compared with four French-speaking control subjects of the same age. They were instructed to listen and to watch the videotape and to identify each token as "ba", "da", "bda", "dba", "va", "tha", "ga", or "other". There were thus 8 response possibilities × 35 trial types = 280 categories. In order to decrease this number, we scored the number of "ba"—and "bda"—responses as one category, and "da"—and "tha"—responses as another category, because these categories are visually very similar and they accounted for more than 91% of BC's judgments. We then computed four different performance measures: the visual and auditory influence in the bimodal condition, and the percentage correct in visual-only and auditory-only trials.

BC's visual influence in bimodal trials was very small: it was only 4% vs 26% (range 8%–43%) for control subjects. Her auditory influence in bimodal trials was large: 75% for BC vs 26% (range 8%–43%) for control subjects. Surprisingly, in the visual-only trials BC performed extremely well: 88% correct for BC vs 55% (range 46%–67%) for the controls, but in the auditory-only trials, BC performed poorly (29% correct for BC vs 64% correct for control). BC thus can discriminate visual-only "ba" from "da", but nevertheless, her visual influence in bimodal trials is almost non-existent.

This performance suggests that BC's anomalous results with auditory-plus-still-face displays may not have accounted fully for her impaired auditory performance in the previous tests. Even when, as in this study, auditory tokens were not accompanied by a visual display, she was poor at identifying them. In most people this would be a good indication for reliance on visual input for audio-visual displays—but for BC this was not the case. As in the previous test, BC, while able to identify silent speech tokens accurately, failed to integrate them in her reports of audio-visual events.

Audio-visual conflict

The previous tests used material originally designed for testing American speakers. It would be more appropriate to test BC with material for French speakers, while further probing this surprising and hitherto unreported failure to integrate seen and heard speech. For instance, audio-visual speechreading is to some extent sensitive to phonetic constraints of the listener's native language as shown in de Gelder et al. (1995; see also Burnham, this volume). Unimodal visual classification of, for example, "ba" and "da" might make use of language-general mechanisms that are less sensitive. On this sort of explanation, BC's integration difficulties could reflect an impairment in the processing of visual speech that may be language-specific. In this test a videotape was made of a female speaker pronouncing a series of VCV sequences. 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 videotape 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, unimodal presentation conditions were produced. For the auditory-only condition, the original auditory signal was dubbed onto the speaker's non-moving image. For the visual-only condition, the auditory channel was deleted from the recording, so the subject had to rely entirely on speechreading. Each presentation condition comprised three replications of the six possible stimuli. BC was instructed to watch the speaker and repeat what she said. In the audio-visual conflict condition, there were only 3 fusions out of 18 trials (16%) while normal performance is about 50% (see de Gelder, Vroomen & van der Heide, 1991). The fusion response occurred when auditory /ana/ was combined with visual /ama/, whereupon BC reported hearing /ama/. In all other trials she reported the audio-part of the audio-visual stimulus. Auditory-only reports were accurate, except that auditory /ana/ was perceived as /ama/ in two out of three cases. This qualifies her fusion responses in the audio-visual case, because it might well be that her /ama/ responses in the auditory /ana/–visual /ama/ trials were exclusively based on the auditory part of the stimulus that was misperceived as /ama/.

Performance on the 18 visual-only trials was quite good. For the visual-only trials, two response categories were made, based on two broad viseme classes: lingual (d, t, n) or bi-labial (b, p, m). She never confused a lingual (d, t, n) with a bi-labial (b, p, m), and she confused a bi-labial with a lingual just once. Six out of nine linguals were perceived as lingual, there was one non-response, and /awa/ was reported twice. On the nine bi-labial trials, she reported six times a bi-labial, once an /ata/ and twice /awa/. So 12 out 17 trials (excluding the non-response) fell within the same viseme class and there was only one confusion between the visually distinct viseme classes.

On this test BC, once more, showed good unimodal visual discrimination, adequate unimodal auditory discrimination, but a lack of normal auditory-visual integration, despite attending fully to all stimuli and despite the fusion stimuli being phonotactically acceptable in French. BC's performance on previous bimodal tasks (in "Visual bias in audio-visual token identification" and "Audio-visual speech with an artificial synthesized face") was not due to unfamiliarity (different language) of the stimulus materials.

Short-term memory for heard and speechread digits

One possible reason for BC's failure to show normal integration is that while viseme discrimination is adequate, driven by knowledge of facial speech (top-down information), nevertheless the achieved representations from the stimulus display (bottom-up information) are insufficiently specified *phonetically* to integrate with heard speech elements (see Green, this volume). It has been well established that in normal speaker-hearers, silently speechread material is remembered as if it has been heard rather than if it has been read (e.g. de Gelder & Vroomen, 1992, 1994). If BC remembers silently lipread material well then we should search for another explanation of her "failure to integrate".

Materials were constructed along the general lines of previous studies where three presentation conditions were contrasted (de Gelder & Vroomen, 1992, 1994). Memory lists contained five French digits (pseudo-randomly selected from the digits 2, 4, 5, 6, 8) spoken by a female speaker and recorded on videotape. The items were presented in three different formats: hearing-plus-speechreading, hearing-only, speechreading-only.

Each condition consisted of 10 experimental trials. The presentation was blocked, and each block was preceded by one warming-up trial. The test started with a practice session of six trials, two of each type. An item was scored correct only if recalled in the correct serial position. BC's overall score for the heard-plus-speechread, heard-only, and speechread-only conditions was .76, .77, and .07, respectively. Her memory for speechread items was thus impaired, but in the heard-plus-speechread and heard-only condition BC had a normal primacy and recency effect. BC has normal verbal memory function when tested with auditory or auditory-visual material.

Recognition of silently spoken digits

One possibility is that speechreading the digits is more effortful for BC than for normals, and that during her effortful speechreading fewer mental resources are allocated for storage and rehearsal. To test this hypothesis BC was presented some weeks later with a new test—this time not involving memory but simple recognition of the same mouthed digits. When tested on speechreading single digits (same digits spoken by the same speaker) only, she scored 17 out of 30 correct (þ2(1) = 25.2, $p < .001$). This is well above chance. BC can recognise and report single speechread items but cannot retain these phonetic forms for rehearsal and recall. Speechread representations can be generated but not sustained.

Together, these findings suggest that BC has *some* capacity to achieve a phonetic representation of seen lexical items by speechreading alone, but that this is weak and, either for structural or for strategic reasons, is insufficient to support verbal memory processes.

GENERAL DISCUSSION

Can seen speech be perceived by a visual agnosic patient with no observable auditory language difficulties and with very dense face-perception impairments but good face knowledge? BC, like HJA (another dense visual agnosic), but unlike the pure prosopagnosic Mrs D (see Campbell, 1992), was impaired at *some* speechreading tasks. Like HJA, her speechreading was much better when the stimulus was a moving face, rather than a still one, and this is in line with her demonstrated lesion sites (bilateral medial occipito-temporal areas, including areas V1 to V4, but sparing V5). However, unlike HJA, and unlike other cases reported to date, facial movement did not afford *full* speechreading capabilities in BC.

BC was able to discriminate silent seen speech quite well and even to identify speechread segments when required, but this ability did not support the integration of seen and heard speech under audio-visual presentation conditions and failed to support immediate memory for seen speech. This pattern of performance has not been reported in patients or in other speechreaders before: auditory-visual fusion effects appear to be mandatory under the sort of presentation conditions used here and are predicted by visual and by auditory identification skills (Massaro & Cohen, 1983). Moreover, auditory-visual fusions have often been considered an index of "implicit" effects of vision on audition: people are usually sensitive to the effects of vision on audition even when they are unwilling to report the classification of the visual stimulus alone (i.e. explicit processing of seen speech alone (e.g. Jordan & Bevan, 1996) Silent speech, moreover, is consistently remembered well by hearing viewers (de Gelder & Vroomen, 1992, 1994).

It would appear from these findings that there is a route to identifying and classifying seen speech that BC can use, but which does not mesh with auditory speech effectively at the phonetic level required to support audio-visual integration and immediate memory for speech that is seen but not heard (Campbell, 1990). There are at least three ways to conceive of this difficulty:

1. The representation derived from silent seen speech is simply too poor or too transient to integrate effectively with heard speech.
2. The temporal processing characteristics of silent seen speech have become desynchronised from those for heard speech in processing, thus losing the temporal-binding feature necessary to effective integration.
3. The way that BC speechreads is qualitatively very different from the way that other people speechread because of over-reliance on a dynamic processing route.

This third possibility is unlikely given the finding that purely dynamic (point-light) displays can generate McGurk effects (Rosenblum & Saldaña, this volume)—at least at first sight. It is possible, however, that point-light displays may access speech-form representations indirectly, and that this is denied to BC. The second possibility, that of some phenomenological desynchronisation of seen and heard speech may be worth further investigation: it could help explain why, under some circumstances (see "Audio-visual speech"), BC is actually *worse* at reporting the auditory component of an audio-visual display than at auditory-only report. This phenomenon is harder to explain on the intuitively appealing idea that BC lacks a robust, stable representation of seen speech because of her brain damage.

Whatever the ultimate explanation of BC's speechreading difficulties, this detailed case study shows there is more to "hearing by eye" than might be predicted from previous case studies, and suggests that the nature and processes of auditory-visual integration in patients with circumscribed visual processing disorders will be worth examining in more depth.

REFERENCES

Bartolomeo, P., Bachoud-Levi, A.-C., & Denes, G. (in press). Preserved imagery for colours in a patient with cerebral achromatopsia. *Cortex.*

Bruyer, R., & Schweich, M. (1991). A clinical test battery of face processing. *International Journal of Neuroscience, 61,* 19–30.

Campbell, R. (1990). Lipreading, Neuropsychology and Immediate Memory. In G. Vallar & T. Shallice (Eds), *Neuropsychological Impairments of Short-term Memory.* Cambridge, UK: Cambridge University Press.

Campbell, R. (1992). The Neuropsychology of Lipreading. *Phil Trans Roy Soc London, B, 335,* 39–45.

Campbell, R. (1996). Seeing Speech in Space and Time. *Proceedings of the 4th International Conference on Spoken Language Processing.* Philadelphia, October.

Campbell, R., de Gelder, B. & de Haan, E. (1996). Lateralization of lipreading: A second look. *Neuropsychologia, 34,* 1235–40.

Campbell, R., Garwood, J., Franklin, S., Howard, D., Landis, T., & Regard, M. (1990). Neuropsychological studies of auditory-visual fusion illusions: Four case studies and their implications. *Neuropsychologia, 28,* 787–802.

Campbell, R., Landis, T. & Regard, M. (1986). Face recognition and lipreading: A neurological dissociation. *Brain, 109,* 509–521.

Campbell, R., Zihl, J., Massaro, D., Munhall, K. & Cohen, M.M. (submitted). *Speechreading in a patient with severe impairment in visual motion perception (akinetopsia).*

de Gelder, B. & Vroomen, J. (1992). Abstract versus modality-specific memory representations. *Memory and Cognition, 20,* 533–538.

de Gelder, B. & Vroomen, J. (1994). Memory for consonants versus vowels in heard and lipread speech. *Journal of Memory and Language, 31,* 737–756.

de Gelder, B., Bertelson, P., Vroomen, J., & Chen, H.C. (1995). Inter-language differences in the McGurk effect for Dutch and Cantonese listeners. *Proceedings of the Fourth European Conference on Speech Communication and Technology* (pp. 1699–1702). Madrid.

de Gelder, B., Vroomen, J. & Van der Heide, L. (1991). Face recognition and lip-reading in autism. *European Journal of Cognitive Psychology, 3,* 69–86.

Farah, M.J. (1991). Patterns of co-occurrence among the associative agnosias: implications for visual object representation. *Cognitive Neuropsychology*, *8*, 1–19.

Humphreys, G.W., Donnelly, N. & Riddoch, J. (1993). Expression is computed separately from facial identity and it is computed separately for moving and static faces. *Neuropsychologia*, *31*, 173–181.

Jordan, T.R. & Bevan, K.M. (1996). Seeing and hearing rotated faces: influences of facial orientation on visual and audio-visual speech recognition. *Journal of Experimental Psychology: Human Perception and Performance*.

Massaro, D.W. & Cohen, M.M. (1983). Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 753–751.

Massaro, D.W. & Cohen, M.M. (1990). Perception of synthesized audible and visible speech. *Psychological Science*, *1*, 1–9.

Milner, A.D., Perrett, D.I., Johnston, R.S., Benson, P.J., Jordan, T.R., Heeley, D.W., Bettucci, D., Mortara, F., Mutani, R., Terazzi, E. & Davidson, D.L.W. (1991). Perception and action in visual form agnosia. *Brain*, *114*, 405–408.

Snodgrass, J.G. & Vanderwart, M. (1980). A standardized set of picture norms. *Journal of Experimental Psychology, Human Learning and Memory*, *6*, 174–215.