

**Tilburg University**

## **The lateralization of lipreading**

Campbell, R.; de Gelder, B.; de Haan, E.H.F.

*Published in:*  
Neuropsychologia

*Publication date:*  
1996

[Link to publication in Tilburg University Research Portal](#)

*Citation for published version (APA):*  
Campbell, R., de Gelder, B., & de Haan, E. H. F. (1996). The lateralization of lipreading: A second look. *Neuropsychologia*, 34(12), 1235-1240.

### **General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

### **Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.



# Note

## The lateralization of lip-reading\*: A second look

RUTH CAMPBELL,† BEATRICE DE GELDER‡§ and EDWARD DE HAAN||

†Department of Psychology, Goldsmiths College, University of London, London, U.K.; ‡Department of Psychology, University of Tilburg, Tilburg, The Netherlands; §Université Libre de Bruxelles, Brussels, Belgium; ||Department of Psychonomics, University of Utrecht, Utrecht, The Netherlands

(Received 18 October 1995; accepted 19 March 1996)

**Abstract**—Photographs of unfamiliar speaking faces were matched by normal right-handed subjects on the basis of perceived mouth-shape (i.e. visible speech sound) across different face-views. A clear left-hemisphere (RVF) processing advantage emerged, which was absent when the task was that of identity matching. In contrast to earlier proposals, the extraction of lip-shape from face photographs may be better managed by left-hemisphere- than right-hemisphere mechanisms even at its initial stages. This may contribute to the observed patterns of dissociations in speech-reading and in audiovisual speech-processing in neurological patients. Copyright © 1996 Elsevier Science Ltd

**Key Words:** speech-reading; lip-reading; laterality; mouth-shape; face-processing.

### Introduction

Left hemisphere lateralization for language is well established from neurological, neuropsychological, experimental and, now, brain-imaging studies [10, 11, 18, 22]. Even when language is not speech-based, as in sign-languages of the deaf, its pattern of localization replicates that for spoken language in most essential respects [17, 25]. In complementary fashion, the cortical localization of face-processing is largely based in the right hemisphere [8, 22, 23, 26]. Questions of interest, then, are these: is speech-reading, which is a face-based skill, yet which requires language to be processed from mouth shapes and face actions, essentially a right-hemisphere or a left-hemisphere function? Alternatively, is it managed equally well in either hemisphere, or does it require both? The answers to these questions should help to establish more clearly the modular geography of cortical function in relation to language on the one hand, and visual processing on the other.

Conflicting answers have been given. Neuropsychological evidence [4, 5, 6] suggests that the left hemisphere is critically

implicated in lip-reading. Patient T, with a circumscribed left parieto-occipito-temporal medial lesion was slightly impaired in sorting face pictures according to speech sound, then failed to classify them properly or show visual influences on heard speech when seen and heard monosyllables were incongruent when dubbed (the audiovisual fusion or McGurk illusion; [5, 9, 13, 14, 15]). Patient D, with dense face-processing problems consequent to a right-sided lesion in an analogous site, showed no impairment in any lip-reading task. By contrast, in normal subjects, it has been reported (a) that matching a still photograph of a lip-shape to an aftercoming speechsound was faster in the LVF/right hemisphere than the RVF/left hemisphere [3], and (b) that very slight asymmetries in favour of the right hemisphere can sometimes be found for audiovisual dubbed speech sounds [1, 7]. Furthermore, two commissurotomy patients, NG and JW, do not show a clear left hemisphere advantage in the few speech-reading tasks which have been explored [1, 4].

With speech-reading as with other higher cognitive functions, it may be useful to consider that functional sub-components of the task utilize cortical structures differentially. Thus, since the right hemisphere appears to play a major role in the analysis of the face as a face [23], it might be implicated in the initial extraction of lip-shape from the face, while the subsequent mapping of this information to speech-analysis itself may require access to effective left posterior hemisphere receptive speech sites. Patterns of lateralization may then vary or switch, depending on precise task demands.

This study aimed to examine directly a logical first stage of processing the face for speech—the ability to extract information about mouth-shape from a facial image. Sergent [20, 21] among several others, showed that hemispheric lateral-

\* The term “lip-reading” has been generally superceded by the term “speech-reading”, which more clearly indicates that reading speech from the face is not confined to the lips alone, but to the perceived shapes and actions of teeth, tongue and jaw. We use the earlier term here in direct reference to an earlier paper entitled “The lateralization of lipreading: a first look”

†Address for correspondence: Department of Human Communication Science, University College London, Chandler House, Wakefield Street, London WC1 N1PG, U.K.; e-mails: R.Campbell@ucl.ac.uk.

ization for face discrimination was predicted by a number of factors. Right-hemisphere localization is more likely for stimuli of low spatial frequency, of reduced luminance, for stimuli in noise or presented for very short durations. Left-hemisphere lateralization is more likely to occur when the tasks required compositional, rather than configural analysis of the face (parts rather than wholes), and can be induced by changes in rate of presentation [21, 23, 24]. The right hemisphere seems to be more efficient at matching faces across different views and lighting conditions [2, 26]. Neuropsychological evidence generally, though by no means conclusively, supports these generalizations [16, 26].

Part-face matching (in this case, matching whole faces on the basis of the lower face-part) is often believed to be better performed by the left hemisphere. However, we can find no published reports that demonstrate this clearly. The matching of schematic face pictures across identical views and using face-parts directly as stimuli produce the most reliable evidence for a role of the left hemisphere [10]. We are aware of no study that has shown a RVF (left hemisphere) advantage when subjects are asked to match naturalistic whole face pictures on the basis of a part of the face (e.g. "is the mouth making a smile or a frown?" or "are the eyes looking ahead or to the side?"). Indeed, in our earlier study [3], we found a marked LVF advantage for speeded matching of a whole face photograph to an after-coming speech sound, implicating the right hemisphere quite strongly. Moreover, Patient T, whose lip-reading was abnormally poor although other face-processing skills were intact, was only slightly impaired at matching different views of face photographs for speech sound, contrasting with a more marked impairment in other speech-reading tasks. One interpretation of T's functional loss, therefore, is that an intact right hemisphere may support the analysis of mouth-shape quite well, and that the LH is required only when matching to a speech event is required directly, as in speaking the face-shape ("what speech sound could this face be saying?") or processing audiovisual inputs.

Our experimental hypothesis, therefore, was that matching different views (full and three-quarter) of the photographed face of different individuals on the basis of perceived speech sound (vowel shape) should generate a right hemisphere advantage. Two vowel shapes (seen face saying /I:/ and seen face saying /u:/) were presented for matching ("are the faces saying the same sound?"). Vowels may be less lateralized to the left hemisphere than are consonantal phonemes; moreover, unlike the earlier study, no overt matching to a speech sound was required, but only matching for mouth-shape. Our aim was to establish, at this simplest level of extracting mouth-shape from well-defined images of faces, whether, and to what extent, processing may be lateralized. This could then form a basis for further explorations of more complex speech-reading abilities and inform a fuller explanation of the various neuropsychological patterns of speech-reading localization.

## Method

### *Subjects*

A group of 24 participants, all students at the University of Tilburg, Netherlands, was tested. They were assessed for hand-preference using a standard questionnaire (Dutch version of the Oldfield-Marshall questionnaire). Six right-handed males and 18 right-handed females formed the final experimental group. Subjects were paid a small amount for participation.

### *Materials*

Stimuli were derived from four different male sitters (also students at the University of Tilburg). The sitters were not known to any of the subjects. They were of similar age (around 24 years), ethnicity and general appearance. All were clean-shaven and without spectacles. Images were captured directly by videocamera on computer (IBM-PC). These were 8-bit greytone images of head-and-shoulders of each sitter, whose head position in relation to the camera was initially fixed using a chin-rest and clamp (removed for image capture). The following shots were chosen from each of the four sitters: full-face saying "ee" (/I:/), three-quarter face, facing left (30° lateral deviation of the head) saying "ee", three-quarter face facing right, saying "ee", and the corresponding set of images of these same speakers saying "oo" (/u:/). These sounds were embedded in Dutch words to enhance the "naturalness" of the speech images. Samples of the images used, for two different sitters, are shown in Fig. 1.

The final image size on the monitor display was 8 × 8.4 cm. The distance from central fixation to the inner edge of the laterally-placed picture was 4.5 cm, while the distance to the outer edge of the picture frame was 12.5 cm. The luminance of the image, measured at the viewer, was 80 cd/m<sup>2</sup>. Stimuli were shown on a monotone monitor attached to a PC (Commodore 486SX-25). Subjects were seated in front of the monitor at a distance of 74 cm from the centre of the display, with their head resting on a fixed headrest. The visual angle subtended by the image of the head was under 2° in central vision, and lateral displacement to the inner edge of the image, for the laterally placed images, was approximately 5.5°.

The experimental sequence was as follows: each experimental trial was preceded by a 200-msec audible warning signal (tone). Two hundred milliseconds following the offset of the warning signal, the first face image appeared in the centre of the screen for 250 msec. Two hundred milliseconds following the offset of this image, a central fixation mark (a cross 0.5 × 0.5 cm) was presented for 250 msec and replaced by a 100-msec exposure of a lateralized image, appearing unpredictably left or right of the fixation. The intertrial interval was 2000 msec, and the time-out for the total sequence was set at 4000 msec.

### *Procedure*

Subjects were instructed to respond by pressing one of two buttons depending on whether the second stimulus was "saying the same speech sound" as the first image. The positions of the two buttons (left and right on keyboard) was balanced across participants for 'same' and 'different' decisions. A total of 160 experimental trials per subject were run. These comprised 20 trials for each of the following conditions: same-person/same lip-shape; same-person/different lip-shape; different person/same lip-shape and different person/different lip-shape for both LVF and RVF presentation. Stimuli for each trial were randomly selected by the computer, without replacement, for each subject. Each stimulus sequence had been coded in advance with respect to identity of sitter, angle of view and mouth shape, and the selection program ensured that every stimulus series occurred with the same probability from trial series to trial series. In the 'same' condition, no completely identical pairs of images were seen; all matches were made across different views of the facial image. In addition, 20 practice trials, comprising stimulus sequences which did not occur on the experimental trials, were given. Practice trials were followed by experimental trials without a break. No feedback was given. Subject time on the task was around 30 min, and subjects were debriefed following the experiment, when handedness measures were taken by questionnaire and familiarity with the sitter was



Fig. 1. Examples of the stimuli used in the experiment. Top, face saying "oo"; bottom, face saying "ee".

checked. All subjects were native Dutch speakers, and all instructions and test procedures were in Dutch.

## Results

### Reaction times

Table 1 shows the means of median RTs for each of the relevant conditions for all right-handed subjects. (Five left-handed subjects were identified following the debriefing. Their data are not included in this report). These data were subject to repeated-measures analysis of variance (MANOVA) procedure in which the experimental conditions (all within-subjects) were field of presentation, lip-pattern (same/different) and person (same/different). The effect of field was significant [ $F(1,23)=17.53$ ,  $P<0.001$ ], as was the effect of lip-pattern [ $F(1,23)=4.77$ ,  $P<0.05$ ] and that of person [ $F(1,23)=12.22$ ,  $P<0.01$ ]. Responses were faster in the RVF, for same-lip judgements and for same persons. These main effects were not moderated by any significant interactions. *T-*

tests confirmed that, for each experimental condition, the effect of the field was significant ( $P<0.01$  for each comparison)

### Errors

Error rates varied from 5 to 13% depending on the condition (errors per condition are shown in Table 1). Errors were analysed exactly as for reaction times. There was no effect of visual field, or of lip-pattern, but matching across different facial identities was significantly harder than matching across the same identities [ $F(1,23)=5.00$ ,  $P<0.05$ ]. Since this pattern was also observed for reaction times (different decisions were slower), there is no evidence here of a trade-off between speed and accuracy.

### Control experiment

Since the lip-matching task generated a strong and unexpected RVF advantage, the experimental conditions may have

Table 1. Lip-shape-matching task means of median reaction times and errors per condition ( $n=24$ )

	LVF same lip- shape same person	LVF same lip- shape different person	LVF different lip-shape same person	LVF different lip-shape different person	RVF same lip-shape same person	RVF same lip-shape different person	RVF different lip-shape same person	RVF different lip-shape different person
Mean RT (msec)	1069.5	1113.9	1102.0	1170.0	1019.6	1062.7	1056.4	1075.1
S.D.	304.6	269.0	222.0	256.6	243.5	238.2	194.8	235.6
Mean errors (%)	2.27	1.45	2.18	1.64	1.95	1.31	2.63	1.1

lent themselves to a general RVF bias under these presentation and response conditions. For example, directional artefacts of the refresh rate of the computer monitor, or of scanning direction speed might have been present. A control task was therefore administered, which, while using identical materials and procedures, changed the task demands from matching for lip-pattern to matching for identity. We hypothesized that identity-matching across different face views and lip-shapes should generate a reduced RVF advantage or even a LVF advantage, in line with neuropsychological findings and earlier experimental reports [2].

Subjects for this study were 22 further right-handed students, drawn from the same pool as that for the lip-shape matching experiment. Sixteen were female. The study was identical in all respects, except that subjects were asked to indicate, by pressing the appropriate key, whether the sequence of two face pictures was of the same person or two different people.

## Results

The means of median reaction times are shown in Table 2. These were subjected to analysis of variance (MANOVA procedure, SPSS) with three within-subject factors: field, lip-shape (same or different) and identity (same or different). The main effect of field failed to approach significance [ $F(1,21)=0.09$ ], while that of lip-pattern was significant [ $F(1,21)=7.75$ ,  $P=0.01$ ]; 'same' lip-pattern displays were faster, as was that of person (decision type) [ $F(1,21)=13.95$ ,  $P<0.01$ ]; 'different' decisions were faster. No interactions reached significance. Error rates were only around 2% overall, and were not analysed since they did not differ significantly from 0. Correlations of errors with reaction times for each condition failed to reach significance.

### Further analyses: difficulty of task effects

Since the identity-match task generated faster response times than the lip-shape matching task, it may be argued that it was not an adequate control for task-laterality sensitivity. However, some internal evidence from the identity matching task suggests that this is not so. The six slowest subjects on this task were analysed separately. These subjects' RT scores overlapped with the scores obtained on the lip-matching tasks, yet none showed an advantage to the RVF for any condition: indeed, three showed an overall LVF advantage for identity matching. Then, in none of the subjects in the identity task, was there any correlation between mean response speed and an overall laterality advantage (LVF performance–RVF performance, calculated separately for each of the four trial types). Similar correlations within the lip-shape matching task were also insignificant.

### Comparing experimental and control conditions

In an overall ANOVA (MANOVA, SPSS procedure) in which experimental condition (lip-matching or identity-matching) was the between-subjects factor, with field, lip-shape and decision-type as within-subject factors, significant effects were obtained that confirmed those indicated by the separate analyses reported above. There was a main effect of experimental condition (identity faster) [ $F(1,44)=16.21$  ( $P<0.001$ )]. There was also a significant overall effect of visual field [ $F(1,44)=7.22$ ,  $P<0.02$ ], that interacted with experimental condition [ $F(1,44)=7.59$ ,  $P<0.01$ ]. This interaction confirms the robustness of the main finding that, whereas identity decisions were not lateralized in this study, lip-shape matching was faster in the RVF. Additionally, interactions of experimental condition with the stimulus (same or different lip-shape) and decision (same or different decision) factors were significant. No other main effects and no further interactions were significant.

### Methodology: unequal gender distributions

In both the control and experimental conditions, women subjects outnumbered men by about 3:1. Could this have implications for the pattern of lateralization found? We believe that it does not. While consistent sex differences in brain organization have been reported, they tend to be less marked for receptive speech and for visual processing than for other tasks [12]. We found no main effects of gender, or any interactions with task, in any of the analyses reported here.

## Discussion

Contrary to earlier findings with normal young right-handed subjects, this study gave marked evidence of a left hemisphere (RVF) advantage in matching silent face pictures for speech-sounds. A control experiment in which identical materials and procedure were used showed no hemisphere advantage for identity matching under these conditions. This clear dissociation suggests that it is unlikely that factors extrinsic to the task (e.g. subjects' hand or visual field bias in overall terms, reading scan habits, or possibly uncontrolled variations in display settings) produced the RVF advantage for the experimental task. It also seems unlikely that task difficulty, *per se*, was responsible for the advantage, since internal analyses failed to suggest that the more difficult task was more lateralized.

LH advantages in face processing have been reported for a number of part-face-matching tasks [20, 21]. Matching photographic images for lip-shape may be such a task, since it requires that only part of the face need be judged, not the face as a whole. Nevertheless, we are aware of no published study that has found reliable RVF advantages when whole-face photographs are matched across different views for a decision that is

Table 2. Identity-matching task means of median reaction times and errors per condition ( $n=22$ )

	LVF same		LVF different		RVF same		RVF different	
	LVF same person same lip-shape	LVF different person different lip-shape	LVF same person same lip-shape	LVF different person different lip-shape	RVF same person same lip-shape	RVF different person different lip-shape	RVF same person same lip-shape	RVF different person different lip-shape
Mean RT (msec)	894.0	817.2	889.1	826.9	897.7	900.75	846.7	798.6
S.D.	235.4	170.7	205.3	183.9	231.0	237.7	210.7	200.7

not based on identity. Furthermore, if part-face matching (i.e. subjects attending solely to the lower face) were the important aspect of lateralization for speech matching, then it should have been found in the earlier study [3]. While it has often been argued that the LH is specialized for rapid temporal processing of categorical stimuli, especially consonantal phonemes, we have demonstrated a clear LH advantage for the perception of seen speech under conditions where this characterization does not apply. It is possible that the LH advantage for silent lip-shape matching may utilize processing sites and systems that overlap with, or are adjacent to, those structures that have been identified, through cortical imaging techniques, for silent reading of words (left middle and superior temporal gyri) [11, 18] and of names (left fusiform gyrus and more lateral left-hemisphere structures) [23]. An alternative, however, is that subjects (silently) articulate the lip-shapes as they see them: for such silent articulation of read material, activation is typically more anterior (involving left SMA and inferior and middle temporal regions).

One study with normal subjects has reported a RVF (left hemisphere) advantage in silent lip-reading. Smeele [24] explored the lateralization of computer-animated faces with a central dot-counting task and varied retinal eccentricities of the animated stimulus face which could be saying /ba/, /va/, /da/ or /θa/. She found a slight RVF advantage in the accuracy of reporting for relatively fast rates of speech and for more eccentric (peripheral) presentations. There was no visual field advantage at all for smaller eccentricities and more central locations.

How can the present finding be accommodated with those of previous studies? The visual characteristics of the display were very different in the various studies. In our earlier study [3], tachistoscopic displays of low illumination and low contrast were used; more different mouth shapes—all produced by a single speaker—were seen. In Baynes *et al.*'s study [1], moving facial stimuli were seen on a video display. All these conditions may foster a reduced RVF advantage. A rather different possibility is that the right hemisphere advantage may occur only when there is concurrent or recent auditory presentation. In Campbell's study [3], matching was to a heard stimulus presented for matching before the visual display. In the studies of Baynes *et al.* [1] and of Diesch [7], audiovisual stimuli were presented. A processing shift to the right hemisphere may occur for faces when the left hemisphere is engaged by auditory speech†.

The present study shows that a (required) stage in the analysis of speech reading, that of identifying mouth-shape from a photographic-quality image, may be effected more readily by the left than the right hemisphere. We are not claiming that only such speech-related facial processing tasks will show this pattern, but we have shown that the assumption, based on earlier findings, that the right hemisphere must be necessarily prepotent at early stages of analysis of faces for seen speech, is wrong. An important task for future research is to establish the extent to which other tasks requiring the analysis of parts (still and moving) in a naturalistic whole face display may moderate patterns of lateralization, and the extent to which subject-

related factors, including practice [19], may be implicated in these patterns. The aim of such explorations is to build up, piece by piece, an understanding of the ways in which different specifically lateralized processes—obligatory or preferential—can build into the observed patterns of localization for speech-reading in relation to both language and visual object processing.

*Acknowledgements*—This work was supported by a project grant from the Economic and Social Research Council of Great Britain to RC and EdH, a research fellowship to RC from the Leverhulme Foundation and by travel grants from the British Council to all three authors. We also thank Menno Hensen, and Jan-Pieter Teunisse for their help in running the experiment, and Maurice Douglas for his technical support.

## References

1. Baynes, K., Funnell, M. G. and Fowler, C. A. Hemispheric contributions to the integration of visual and auditory information in speech perception. *Perception and Psychophysics* **55**, 633–641, 1994.
2. Bertelson, P., Van Haelen, A. and Morais, J. Left hemifield superiority and the extraction of physiognomic information. In *Structure and Function of the Cerebral Commissures*, I. Steele-Russell, M. W. Van Hof and G. Berlucchi (Editors), pp. 400–410. Macmillan, London, 1979.
3. Campbell, R. The lateralisation of lipreading: A first look. *Brain and Cognition* **5**, 1–21, 1986.
4. Campbell, R. The neuropsychology of lipreading. *Philosophical Transactions of the Royal Society* **B**, 298–308, 1992.
5. Campbell, R., Garwood, J., Franklin, S., Howard, D., Landis, T. and Regard, M. Neuropsychological studies of auditory–visual fusion illusions. *Neuropsychologia* **28**, 787–802, 1990.
6. Campbell, R., Landis, T. and Regard, M. Face recognition and lipreading: A neurological dissociation. *Brain* **109**, 509–521, 1986.
7. Diesch, E. Left and right hemifield advantages of fusions and combinations in audio-visual speech perception. *Quarterly Journal of Experimental Psychology* **48a**, 320–333, 1995.
8. Ellis, H. D. The role of the right hemisphere in face perception. In *Functions of the Right Hemisphere*, A. Young (Editor), pp. 33–64. Academic Press, London, 1983.
9. Green, K. and Kuhl, P. Integral processing of visual place and auditory voicing during phonetic perception. *Journal of Experimental Psychology: Human Perception and Performance* **17**, 278–288, 1991.
10. Hellige, J. *Hemispheric Asymmetry: What is Left and What is Right?* Harvard University Press, Cambridge, 1993.
11. Howard, D., Patterson, K., Wise, R., Brown, W. D., Friston, K. and Weiller, C. The cortical localisation of the lexicons. *Brain* **115**, 1769–1782, 1992.
12. Kimura, D. Sex differences in the brain. In *Readings from Scientific American: Mind and Brain*, pp. 78–90. W. H. Freeman, 1993.
13. MacDonald, J. and McGurk, H. Visual influences on speech perception. *Perception and Psychophysics* **24**, 253–257, 1978.
14. McGurk, H. and MacDonald, J. Hearing lips and seeing voices. *Nature* **264**, 746–748, 1976.
15. Massaro, D. and Cohen, M. Evaluation and integration of visual and auditory information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance* **9**, 753–771, 1983.

† However, this suggestion requires expansion: in Diesch's (1995) [7] study, bilaterally presented audiovisual speakers generated different visual field advantages as a function of the type of stimulus presented, while 'fusion' type stimuli were more likely to be perceived as such when the visual form was in the LVF (in line with the suggestions proposed here); 'mixed' or 'blend' responses such as "bga" or "bda" tended to occur more frequently with RVF responses. It is possible that the phonetic output requirements of this task (producing items that are not common in the phonological repertoire) may have biased processing.

