

Sex differences in lipreading

FERN M. JOHNSON and LESLIE H. HICKS
Howard University, Washington, D.C.

and

TERRY GOLDBERG and MICHAEL S. MYSLOBODSKY
*National Institute of Mental Health, Washington, D.C.
and St. Elizabeth's Hospital, Washington, D.C.*

Three lipreading tests designed to assess recognition of syllables, words, and overlearned sentences were administered to 53 normal (17 male and 36 female) volunteers. As predicted, females were superior in lipreading ability, notably in recognizing sentences. We discuss the possibility that sex differences in gazing patterns explain the sex differences in lipreading ability.

Some writers suggest that there are gender differences in gazing, with females being more avid gazers than males (Berndl, Dewitz, Grusser, & Kiefer, 1986; Exline, 1963; LaFrance & Mayo, 1976). These data led researchers to pose the question as to why females would look longer than males at a speaker in the dyadic interaction. Kleinke's (1986) review suggests that the gains may be numerous: gazing provides information, regulates interaction, expresses intimacy, exercises social control, facilitates service and task goals, and so forth. There is yet another function not covered by Kleinke's review, that of speechreading (Campbell, 1986). We reasoned that if females' gazing is directed at monitoring another person's speech, females may well evince a better capacity for lipreading.

METHOD

Subjects

Fifty-three healthy volunteer students (17 males and 36 females) at Howard University were recruited as subjects. All participated as a part of course requirements. Their ages ranged from 17 to 27 (mean ages of 19.86 and 20.06 for females and males, respectively). All subjects were native English speakers, and all were screened for normal hearing, normal (or corrected-to-normal) vision, right-handedness, and adequate persistence.

Neuropsychological Testing

Handedness was assessed using the modified Edinburgh Inventory (Oldfield, 1971). Given the possibility of errors in self-report, the subjects were asked to perform the items of the questionnaire in the form of a pantomime.

Persistence (Joynt, Benton, & Fogel, 1962) was assessed to exclude individuals who were unable to sustain a "tonic" sensory-motor effort (i.e., to keep eyes closed, tongue protruded, or mouth open for 20 sec or to fix the gaze on a target at about 45° laterally from the examiner for 30 sec).

A battery of three lipreading tests was designed to determine the subjects' ability to recognize syllables, words, and overlearned phrases. Consequently, the procedure consisted of three speech tests, which were recorded on videotape.

Address correspondence to Leslie H. Hicks, Department of Psychology, Howard University, Washington, DC 20059.

The syllable recognition test was composed of three segments: (1) In the *sight-only* condition, the subject had to identify a set of 10 consonant-vowel combinations. The consonants /b, p, t, d, k, g, n, m, v, l/, which represent labial, alveolar, and velar articulation, were used. They were pronounced in a consonant-vowel context with the vowel [a] (as in partner). (2) In the *hearing-only* condition, the subject had to identify 10 spoken syllables, the same as those used in the sight-only condition. (3) In the nonillusory, audiovisual condition, the list of the above 10 syllables was presented visually and acoustically in the nonconflict combination. Any incorrect response in the hearing-only or the audiovisual conditions was noted as an error; however, because visual cues available through lipreading can reliably indicate the place of consonant articulation but cannot distinguish facial cues of nasality and voicing, a response was scored as erroneous in the sight-only condition only when the place of articulation was confused (e.g., [p] confused with [d]).

In the *consonant (word) recognition test*, the subject had to recognize one-or two-syllable words (e.g., *wish, will, with*) pronounced by the speaker and circle them on a multiple-choice form. There were 25 such videotaped words spoken by a female, native speaker of English whose head and shoulders were taped. (The maximal score was 25.) At all times the subjects were allowed to consult the typed list of words.

In the *overlearned phrases recognition test*, there were 24 test phrases with 2 to 6 words per sentence ($M = 4.1$), such as "What became of them," "Too good to be true," and "Sorry to hear that." All were read to the subjects in blocks of 3(trial)-10-11(test) prior to the viewing of the videotaped phrases to familiarize them with the linguistic stock. There were 1 or 2 syllables in every word ($M = 1.4$). The syllables and the phrases were spoken by a male native speaker of English. Only the lower part of his face was viewed on the screen. A correct answer was scored as 1; a partially correct response (close guess) received a score of 0.5. The maximal score was 24.

Procedure

The videotape was shown on a 13-in. color Panasonic CT-1330v TV monitor via a Panasonic NV-8200 video tape recorder. The screen subtended 10.68° of the visual field for the standard viewing distance of about 1.5 m. The room was dimly lit for a comfortable viewing condition. Each utterance appeared for approximately 0.5 sec and was followed by a pause that lasted until the subject responded. The experimenter sat next to the monitor to ensure that the subject watched the screen during speech presentation. The subject was repeatedly reminded to look attentively at the screen and to repeat every syllable seen or heard. No time limit was set for the response, and a new stimulus was presented when the subject either made a guess or acknowledged his/her inability to recognize the stimulus. All responses were written down as the subject said them. No subject had difficulty in grasping the requirements of the lipreading tests.

RESULTS

The maximal scores in the word and phrase recognition tests and the number of errors in syllable recognition were compared separately using the Data-text statistical package (Armor & Couch, 1972).

Syllable Recognition Test

A 2×3 analysis of variance performed on the proportion of the erroneous identification of visemes, phonemes, and audiovisual syllables in males and females showed significant differences in the number of errors that occurred in different segments of the test [$F(2,102) = 3.31, p < .04$]. Post hoc t tests showed that males were superior on the identification of phonemes [$t(16) = 2.21, p < .025$], whereas females did better in identifying the visemes [$t(35) = 3.13, p < .005$].

Word and Phrase Recognition Tests

The word recognition test was relatively easy; almost all subjects were 100% correct (maximum score was 25). No differences between males and females were obtained on the word test. However, females were superior in recognizing the sentences; the mean score was 15.94 (± 4.64) for females and 12.59 (± 4.26) for males. A 2×2 analysis of variance (sex \times test) revealed significant main effects for both the sex factor [$F(1,51) = 6.80, p < .01$] and the test factor [$F(1,51) = 55.71, p < .001$]. This significance was associated with a superior performance by females in phrase recognition. In addition, the results revealed a significant interaction of the main effects (sex \times test) [$F(1,51) = 4.73, p < .04$]. This result suggests that females might generally have attended more to the visual than to the audio input. In fact, this suggestion is supported by the results of the syllable recognition test.

DISCUSSION

To the extent that lipreading is a perceptual task, the superiority of females in all lipreading tests would seem consistent with findings of their more active gazing. Moreover, the present study suggests that the more persistent gazing by females not only functions to express intimacy or communicate superiority but is used very effectively to monitor speech. Symptomatically, the gender differences emerged both in recognizing the visemes and in decoding the overlearned sentences (clichés). The sentences have a high level of word predictability provided that the lipreader deploys a strategy of identification of the contextual cues rather than attempting to identify each viseme separately. In this respect, clichés provide a better and more realistic measure of speech recognition than do words. It is of interest that the ability to use contextual information is a feature that differentiates good from poor readers (Williams, 1982).

One might attribute females' superiority in lipreading to their general capacity to recognize faces and grasp the emotional signals they emit (Berndl et al., 1986). Since lipreading depends on contextual information (Campbell, 1986; Williams, 1982), facial scanning might have provided some guidance in specifying emotional context of lexical messages. However, the sentences represent fragments of small talk utterance with hardly any emotional change. If they were, there was nothing in the visual appearance of the speaker that would hint at such a possibility. The intonational aspect of phrases, or rather prosodic expression was impossible to derive from the lower facial region shown in the phrase recognition test. Parenthetically, one can add that difficulties

experienced by individuals with the face recognition disorder (prosopagnosia) have long been attributed to a selective deficit in identification and description of the region around the eyes (Gloring, Gloring, Hoff, & Tschabitscher, 1966) rather than the lower part of the face. Moreover, such explanation is unlikely to account for a superior recognition of visemes by females. Nonsense syllables do not require syntactic and semantic cues crucial for more complex tasks of word or sentence recognition.

The foregoing findings might raise a question of whether females' superiority in lipreading in the tests is anchored in their higher sensitivity to linguistic stimuli. As a group, females are believed to lean toward "verbal strategies" whereas males prefer "nonverbal" ones (Bryden, 1978). Consequently, one might escape the superiority-inferiority dichotomy by rephrasing this conjecture in terms of gazing strategy. Such conjecture would suggest that females and males have dissimilar gazing patterns; when scanning a face they look for dissimilar information and rely on dissimilar cues. Gazing for lipreading and scanning the face for recognition seem to represent different cognitive functions controlled by dissimilar brain circuits. On the basis of left visual field superiority in matching briefly displayed face photographs with different lip configuration to a heard speech sound, Campbell (1986) originally proposed right-hemisphere control of lipreading. This was a tempting conjecture considering the contribution of the right hemisphere to the analysis of shape, the recognition of visual objects (Srebro, 1985), and the processing of language at a sentence level (Caramazza, Gordon, Zurif, & DeLuca, 1976; Grossman & Haberman, 1987). However, Campbell (1986) pointed out that her task was not based on the interpretation of serially presented visemes, as is required in the more realistic speechreading setting.

More recently, Campbell, Landis, and Regard (1986) examined lipreading and the susceptibility to the McGurk's blend illusion (McGurk & MacDonald, 1976) in two nonaphasic patients and with occipitotemporal ischemic infarction in the territory of the posterior cerebral artery. The patient with the right-sided lesion was unable to recognize familiar faces but able to lipread and was fully susceptible to the blend illusion. The patient with the left-hemisphere lesion showed the converse impairment, being able to match expressions across lower-face pictures but unable to read lips. These findings, implicating the left hemisphere, raise an additional question as to whether sex differences in lipreading are related to some aspects of cerebral organization. A tendency to imagine the female brain as more bilaterally organized (i.e., less lateralized) than the male brain (McGlone, 1986; see, however, Fairweather, 1982) necessitates an explanation as to why such brain organization has any advantage in lipreading. Until further investigations are completed, any attempt to find an explanation of the present result in terms of brain laterality remain speculative.

REFERENCES

- ARMOR, D. J., & COUCH, A. J. (1972). *Data-text primer: An introduction to computerized social data analysis*. New York: Free Press.
- BERNDL, K., DEWITZ, W., GRUSSER, O. J., & KIEFER, R. H. (1986). A test movie to study elementary abilities in perception and recognition of mimic and gestural expression. *European Archives of Psychiatry & Neurological Science*, 235, 276-281.
- BRYDEN, M. P. (1978). Strategy effects in the assessment of hemispheric asymmetry. In G. Underwood (Ed.), *Strategies of information processing* (pp. 117-149). New York: Academic Press.
- CAMPBELL, R. (1986). The lateralization of lipread sounds: A first look. *Brain & Cognition*, 5, 1-21.
- CAMPBELL, R., LANDIS, T., & REGARD, M. (1986). Face recognition and lipreading: A neurological dissociation. *Brain*, 109, 509-521.
- CARAMAZZA, A., GORDON, J., ZURIF, E., & DELUCA, D. (1976). Right hemispheric damage and problem solving behavior. *Brain & Language*, 3, 41-46.
- EXLINE, R. V. (1963). Explorations in the process of person perception: Visual interaction in relation to competition, sex, and need for affiliation. *Journal of Personality*, 31, 1-20.
- FAIRWEATHER, H. (1982). Sex differences: Little reason for female to play midfield. In J. G. Beaumont (Ed.), *Divided visual field strategies of cerebral organization* (pp. 147-194). London: Academic Press.

- GLONING, I., GLONING, K., HOFF, K., & TSCHABITSCHER, H. (1966). Zur Prosopagnosie. *Neuropsychologia*, **4**, 113-131.
- GROSSMAN, M., & HABERMAN, S. (1987). The detection of errors in sentences after right hemisphere brain damage. *Neuropsychologia*, **25**, 163-172.
- JOYNT, R. J., BENTON, A. L., & FOGEL, M. L. (1962). Behavioral and pathological correlates of motor impersistence. *Neurology*, **12**, 876-881.
- KLEINKE, C. L. (1986). Gaze and eye contact: A research review. *Psychological Bulletin*, **100**, 78-100.
- LAFRANCE, M., & MAYO, C. (1976). Racial differences in gaze behavior during conversations: Two systematic observational studies. *Journal of Personality & Social Psychology*, **33**, 547-552.
- MCGLONE, J. (1986). Sex differences in human brain asymmetry. *Behavior & Brain Science*, **3**, 122-128.
- MCGURK, H., & MACDONALD, J. (1976). Hearing lips and seeing voices. *Nature*, **264**, 746-748.
- OLDFIELD, R. C. (1971). The assessment and analysis of handedness. *Neuropsychologia*, **9**, 97-113.
- SREBRO, R. (1985). Localization of cortical activity associated with visual recognition in humans. *Journal of Physiology*, **360**, 247-259.
- WILLIAMS, A. (1982). The relationship between two visual communication systems: Reading and lipreading. *Journal of Speech & Hearing Research*, **25**, 500-503.

(Manuscript received for publication July 29, 1987.)