

Thus, when the visual signal moves away from the sound source and the bias decreases, one may obtain a good fit by decreasing the visual support for the more distant location. The crucial point, however, is that the truth values are meaningless because there is no guarantee that there is a correspondence with the perceptual mechanisms that lead to these truth values. Thus, the fact that fuzzy sets of mathematics can describe aspects of results does not enlighten us about the underlying mechanism: a good fit is therefore no criterion to accept FLMP as an adequate theory of perception.

Acknowledgement

We would like to thank Paul Bertelson for his insightful comments on a previous version of this paper.

Jean Vroomen and Beatrice de Gelder
Tilburg University, Department of
Psychology, Warandelaan 2,

PO Box 90153, 5000 Le Tilburg,
The Netherlands.
tel: +31 13 466 2394
fax: +31 13 466 2370
e-mail: j.vroomen@kub.nl

References

- 1 Massaro, D.W. (1999) Speechreading: illusion or window into pattern recognition. *Trends Cognit. Sci.* 3, 310–317
- 2 Massaro, D.W. et al. (1998) Visible language in speech perception: lipreading and reading. *Visible Lang.* 22, 9–31
- 3 Driver, J. and Spence, C. (1998) Attention and the crossmodal construction of space. *Trends Cognit. Sci.* 2, 254–262
- 4 Frost, R. et al. (1988) Can speech perception be influenced by simultaneous presentation of print? *J. Mem. Lang.* 27, 741–755
- 5 Radeau, M. and Bertelson, P. (1977) Adaptation to auditory–visual discordance and ventriloquism in semirealistic situations. *Percept. Psychophys.* 22, 137–146
- 6 Welch, R.B. and Warren, D.H. (1980) Immediate perceptual response to intersensory discrepancy. *Psychol. Bull.* 88, 638–667
- 7 Bedford, F.L. (1999) Keeping perception accurate. *Trends Cognit. Sci.* 2, 4–11
- 8 Radeau, M. (1994) Auditory–visual spatial interaction and modularity. *Cahiers de Psychologie Cognitives* 13, 3–51
- 9 Vroomen, J. Ventriloquism and the nature of the unity assumption. In *Cognitive Contributions to the Perception of Spatial and Temporal Events* (Aschersleben, G. et al., eds), Elsevier (in press)
- 10 Bertelson, P. Ventriloquism: A case of crossmodal perceptual grouping. In *Cognitive Contributions to the Perception of Spatial and Temporal Events* (Aschersleben, G. et al., eds), Elsevier (in press)
- 11 Bermant, R.I. and Welch, R.B. (1976) Effect of degree of separation of visual–auditory stimulus and eye position upon spatial interaction of vision and audition. *Percept. Mot. Skills* 43, 487–493
- 12 Bertelson, P. and Radeau, M. (1981) Crossmodal bias and perceptual fusion with auditory–visual discordance. *Percept. Psychophys.* 29, 578–584

Reply to Vroomen and de Gelder

Given that I am sympathetic to Vroomen and de Gelder's commentary¹, I can only hope that they have failed to read my lips (or my research papers) rather than misunderstood what they have read. Admittedly, my short review article² could be read out of context and the reader could easily believe that I have gone beyond the evidence given (in the same way that our perception often goes beyond the information given). We use and promote our information-processing framework primarily because it encourages the investigator to determine the stage (level in Vroomen and de Gelder's terms) of processing responsible for various behaviors.

I will show that their two main points can be easily pursued within our framework, after a short qualification of the origins of the FLMP. Vroomen and de Gelder state that, 'Originally, it (the FLMP) was developed for the situation where listeners hear tokens from a /ba-/da/ continuum while viewing a face articulating /ba/ or /da/.'¹ It is important to assure the reader that the FLMP was with us well before McGurk and MacDonald published their McGurk effect³. The model was originally developed to account for the integration of several auditory cues in speech perception and for various sources of information in sentence processing^{4–6}. In assessing the model, it is important to note that the FLMP was not derived simply to describe speech perception by ear and eye, but rather to describe pattern recognition more generally.

Lip-reading versus reading

First, Vroomen and de Gelder question whether written text is operating at the same stage as visible speech when these sources are separately combined

with auditory speech. It is intuitive, somehow, to believe that the influence of visible speech is more real than the influence of written text. However, it is worth noting a couple of caveats. First, Vroomen and de Gelder should not dismiss the positive finding of Frost et al.⁷ as simply a bias because we now know that biases can be truly perceptual. This possibility was pointed out long ago by Paul Bertelson⁸ when investigators tried to dismiss his ventriloquism effect as a response bias when analysed within the context of signal detection theory (for further discussion of the important distinction between perceptual bias and decision bias, see Ref. 9). How would one test whether the two types of visual input operate differently? Our experiment is simply a first step along that road. Contrasting different models of performance should then follow. It is straightforward to formulate a model based on the interpretation proposed by Vroomen and de Gelder in which the visual input has its influence on decision rather than perception (Ref. 10, Chapter 2). The outcome of these tests would speak to the issue of analogous processes in reading and lip-reading.

Vroomen and de Gelder¹ state that 'the issue is not whether reading and lip-reading interact at all with speech, but whether they interact at the same processing level, and whether FLMP allows one to distinguish between the various forms that this interaction may take'. This question has always been of central interest to us and is why I argue for the formalization and testing of alternative models. The post-perceptual guessing model, the auditory dominance model, and the 'Race' model have all been tested as alternatives to the FLMP, primarily because they assume different

forms of interaction of the two sources of information.

Additional experiments can be generated to distinguish between various theoretical explanations. One important source of evidence comes from the nature of the judgments that are given. Specifically, Vroomen and de Gelder are interested in combination responses, such as /bdi/. Before discussing the story of /bdi/, it should be noted that when and how often these combinations occur is highly variable and unpredictable. In an early study with open-ended alternatives, Repp et al. found no combination responses (Ref. 11, and see Ref. 12, pp. 52–54). In our studies with /bd/ as one of the specified response alternatives, we have found up to 80% (Ref. 13) and as low as 10% combination responses (Ref. 10, p. 146) when a visual /ba/ is paired with an auditory /da/. In order to understand whether these combination responses should be equivalent in the reading and lip-reading conditions, however, it is first necessary to understand why they occur in lip-reading. Our interpretation has been that a visual /b/ paired with an auditory /d/ provide two sources of information that are consistent with /bd/. A visual /b/ looks a lot like a visual /bd/, and an auditory /d/ is somewhat similar to auditory /bd/. Thus, /bd/ can be a reasonable percept given these two sources. This explanation also predicts very few /db/ judgments when a visual /d/ is paired with an auditory /b/. In this case, a visual /d/ is very different from a visual /db/. These types of constraints probably do not occur in the reading situation, however, because the written letter activates some speech-like representation without actually providing a speech stimulus. This situation is more analogous to the

phonemic restoration effect, as I indicated in my review article².

Being aware of this issue, we had already replicated the experiment reported in Ref. 2 (Fig. 1), in which the participants were given eight response alternatives, including /bdi/ and /dbi/. Much to our amazement, the number of /bdi/ responses was about equivalent in the letter and speech conditions. This result supports the idea of analogous processes in reading and lip-reading, but the issue is by no means resolved. It is necessary to carry out several additional manipulations, such as attention instructions and rating judgments (Ref. 10, Chapters 8 and 9). An important issue to pursue is whether any differences between visible speech and written text can be explained in terms of information differences, or require differences in information processing.

Ventriloquism

Vroomen and de Gelder's second concern addresses the reasonable finding that increasing the separation between the auditory and visual inputs can decrease the ventriloquism effect. A direct application of the FLMP would predict just the opposite finding – as stated by Vroomen and de Gelder '...because in FLMP there is always integration'¹. Every theory includes boundary conditions, however, and the FLMP is no different in this regard. We have constantly included these boundary conditions as part of our description of the model. To quote from Massaro, 1998:

'Central to our model of pattern recognition is the notion of a meaningful object or event. In order to integrate multiple sources of information as specified by the FLMP, it is necessary to relate them to the same event.' (Ref. 10, p. 72)

The research presented in that chapter revealed that:

'...the natural tendency to integrate multiple sources of information can be disrupted when the sources of information are separated in time sufficiently to signify two separate events.' (Ref. 10, p. 80)

It cannot be much of a stretch of the imagination to replace time with space: the natural tendency to integrate two sources of information can be disrupted when their spatial separation is sufficient to signify two separate events. Again, formal models of performance can be tested to determine how perceivers treat the auditory and visual signals when their spatial separation becomes large. I could speculate on several alternatives but will leave this for future research projects.

Conclusion

I appreciated Vroomen and de Gelder's effort to communicate their reticence in accepting the FLMP. Their comments, however, substantiate my belief in the tremendous value of the interdisciplinary framework we have helped develop. The approach of cognitive science offers the potential to uncover general principles that would not be easily revealed in specialized studies of different modules of the mind. In the latter approach, there would be little communication among researchers studying spatial localization, speech perception, and reading. By confronting all of these fields with a common framework for inquiry, we discover more general principles that help us understand how perceivers solve nature's many and varied puzzles.

Dominic W. Massaro
University of California, Santa Cruz,
CA 95060, USA.
tel: +1 831 459 2330
fax: +1 831 459 3519
e-mail: massaro@fuzzy.ucsc.edu

References

- 1 Vroomen, J. and de Gelder, B. (1999) Crossmodal integration: a good fit is no criterion. *Trends Cognit. Sci.* 4, 37–38
- 2 Massaro, D.W. (1999) Speechreading: illusion or window into pattern recognition. *Trends Cognit. Sci.* 3, 310–317
- 3 McGurk, H. and MacDonald, J. (1976) Hearing lips and seeing voices. *Nature* 264, 746–748
- 4 Massaro, D.W. and Cohen, M.M. (1976) The contribution of fundamental frequency and voice onset times to the /zɪl-/sɪl/ distinction. *J. Acoust. Soc. Am.* 60, 704–717
- 5 Oden, G. C. (1977) Integration of fuzzy logical information. *J. Exp. Psychol. Hum. Percept. Perform.* 3, 565–575
- 6 Oden, G.C. and Massaro, D.W. (1978) Integration of featural information in speech perception. *Psychol. Rev.* 85, 172–191
- 7 Frost, R. et al. (1988) Can speech perception be influenced by simultaneous presentation of print? *J. Mem. Lang.* 27, 741–755
- 8 Bertelson, P. and Radeau, M. (1976) Ventriloquism, sensory interaction, and response bias: remarks on the paper by Choe, Welch, Gilford and Juola. *Percept. Psychophys.* 19, 531–535
- 9 Massaro, D.W. and Cowan, N., (1993) Information processing models: microscopes of the mind. *Annu. Rev. Psychol.* 44, 383–425
- 10 Massaro, D.W. (1998) *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle*, MIT Press
- 11 Repp, B.H. et al. (1983) Exploring the 'McGurk Effect'. *Bull. Psychonomic Soc.* p. 358 (abstract)
- 12 Massaro, D.W. (1987) *Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry*, Erlbaum
- 13 Massaro, D.W. and Cohan, M.M. (1983) Evaluation and integration of visual and auditory information in speech perception. *J. Exp. Psychol. Hum. Percept. Perform.* 9, 753–771

Trends in Cognitive Sciences online – tics.trends.com

The online version of Trends in Cognitive Sciences contains a simple search and browse facility that will enable you to view full-text versions of all articles published in the journal since January 1998. Browsing and searching the articles is straightforward and designed around the need for simple navigation. Articles can be viewed in three display formats: text-only, text with pop-up figures, or as PDF files for high-quality printing.

Personal print subscribers benefit from online access to Trends in Cognitive Sciences online by accessing the site using a unique subscription key. For information about how to obtain your key please visit <http://www.trends.com> The subscription key is important and should be kept safe as it is your unique access code to your online journal for the remainder of the year. Institutional access is currently available through Elsevier's ScienceDirect (<http://www.sciencedirect.com>) services.

Subscribers and non-subscribers alike will continue to have free access to full content lists, abstracts, search facilities and the e-mail contents alerting service. It is also possible to download full-text articles on a pay-per-view basis.

The following recent articles are just some of those available in the fully searchable archive:

- Thornhill, R. and Gangestad, S.G. (1999) Facial attractiveness. *Trends Cognit. Sci.* 3, 452–460
- Perner, J. and Lang, B. (1999) Development of theory of mind and executive control. *Trends Cognit. Sci.* 3, 337–344
- Gallagher, S. (2000) Philosophical conceptions of the self: implications for cognitive science. *Trends Cognit. Sci.* 4, 14–21