



HARRY MCGURK AND THE MCGURK EFFECT

Denis Burnham

A significant part of the history of auditory-visual speech processing is the discovery of the way in which auditory and visual speech components, when in conflict, can give rise to an emergent percept. Not many of the present-day young auditory-visual speech researchers know about the history of this finding. Early this year the organisers of AVSP'98 contacted the Australian Institute of Family Studies in Melbourne, Australia, to ask its Director, Harry McGurk, to talk at AVSP'98. Unfortunately Harry died in April this year and so was unable to meet and talk to the band of researchers now working in the area of auditory-visual speech processing. However, we still wanted to inform people about the origins of this effect, and so with the kind assistance of his daughter, Rhona McGurk, we publish here for the first time Harry's inaugural lecture as Professor of Psychology at the University of Surrey. This tells of his general orientation to developmental psychology, and his discovery of the McGurk effect. Here I will give you a little extra information - some historical background before and since the discovery; and Harry's reaction to the discovery.

Harry's primary field of psychological research was human development. He conducted infant perception studies back in the early '70s amid the great sense of excitement generated by the availability of new powerful methods for tapping infants' perceptual abilities [see 1]. Harry's early work was in unimodal visual perception by infants [2]. In 1971 Aronson and Rosenbloom (with a theoretical orientation in line with the then guru of infant studies, Tom Bower) published a paper in *Science*, in which 1- to 2-month-old infants viewed auditory-visual presentations of their mother [3]. The infant's mother was behind a glass window in front of the infant, but her voice came from two audio speakers, one on each side of the infant. When the audio speakers were in balance the mother's voice appeared coincident with her face, directly in front of the infant. After a baseline period with the speakers in balance, Aronson and Rosenbloom made either one or the other speaker dominant, so that the mother's voice emanated from the left or the right of the infant, ie, spatially displaced from her face. Under such spatial dislocation conditions, Aronson and Rosenbloom found that the infants became agitated and upset in various idiosyncratic ways. Interestingly, one index that was evident in all the infants was an increase in tongue protrusions after the mother's voice became spatially separated from her face. Aronson and

Rosenbloom took their results to show that spatial perception in infancy occurs within a common auditory-visual space.

These results fit in nicely with the Gibsonian and neo-Gibsonian (Bowerian) view of the initial unity of the senses, but not so nicely with McGurk's view that unity of the senses, especially in speech, is a product of experience. In 1974 Harry McGurk at Surrey and Michael Lewis at Princeton, published a rejoinder experiment in *Science* [4]. They tested 1-, 4-, and 7-month-old infants in similar (though not identical) conditions to those in the original experiment. Their results failed to confirm Aronson and Rosenbloom's earlier findings, and thus did not provide support for the initial unity of the senses notion.

Possible reasons for the discrepancy between these results, such as the developmental course of auditory localisation in infancy [5,6], and the different viewing distance used in the two studies [7] need not concern us in the current context. What is of interest here is the difference in theoretical stance between Aronson and Rosenbloom's Gibsonian view, and McGurk's more Piagetian or perhaps associationist view that sensory information is combined cross-modally as the result of operations on the environment or more generally, experience [see 8].

This conceptual difference set the scene for McGurk's next move. He had refuted Aronson and Rosenbloom's claim about *intermodal spatial dislocation*, and now went on to investigate the unity of the senses hypothesis by way of *intermodal conflict*. As documented in the accompanying address [9], McGurk took videotapes of productions of /ba/ and /ga/ and had them dubbed to produce not just the matching ba-voice/ba-lips, and ga-voice/ga-lips sequences, but also conflicting ba-voice/ga-lips and ga-voice/ba-lips sequences. What is not documented in McGurk's address is his displeasure when the videotapes returned from the auditory-visual centre at Surrey. Harry said he thought the technician had made a mistake, and that he told the technician so in no uncertain terms. He soon realised however that there had been no mistake and that he was experiencing something quite remarkable, a "da" percept when presented with an auditory [ba] and visual [ga], and a "bga" percept from an auditory [ga] and visual [ba]. (See a reminiscence of this by McGurk quoted in Massaro's recent book [10]). This serendipitous

discovery established without question that whenever visual speech information is available we humans use it. Workers in other areas, eg, hearing impairment, had long known of the contribution of visual information to speech perception, but now it had been convincingly demonstrated in a controlled experimental study.

Instead of continuing with the planned intermodal conflict studies with infants, McGurk and McDonald published their results in the now classic *Nature* paper [11], and later went on to postulate the visual place/ auditory manner hypothesis to account for the effect, a version of which appears in the accompanying address. In an excellent example of the *Zeitgeist* phenomenon, it is interesting to note that elsewhere in Britain, about the same time if not before, Barbara Dodd had discovered a similar effect, eg, the perception of 'towel' from auditory 'tough', and visual 'hole' [12].

About the time of the McGurk address at Surrey, Barbara Dodd and Ruth Campbell published an edited volume, *Hearing by Eye*, which integrated the work of a number of researchers in auditory-visual speech perception and production [13]. In this Bible for the modern study of auditory-visual speech processing, Quentin Summerfield's Genesis chapter considered a range of possible ways in which auditory and visual information could be combined in speech perception [14]. This theoretical review did not favour the visual place/auditory manner hypothesis, but this is of little consequence here: Harry had provided the auditory-visual speech community with a dramatic phenomenon on which to hang its mortarboard.

It is a pity he could not have lived to attend this conference. But we have his address, and we also have something else. In 1988, while attending the International Congress of Psychology, Harry visited my lab at the University of NSW, and allowed me to make a copy of a videotape of his. His daughter, Rhona McGurk, has kindly allowed us to reproduce this here. There are two accompanying video files on the CD-Rom. In one you will see Harry McGurk producing a sequence of syllables in which the auditory component is always /ba/, but the visual component varies across trials from /ba/, /va/, /tha/, /da/, /ya/, /ga/, to /ha/. This corresponds to the example mentioned in the inaugural address on page 10. In the second sequence you will see and hear the example referred to on page 13 of the address. In this, Harry finally reveals who taught him to drive.

REFERENCES

1. Burnham, D. & Dodd (1998) Familiarity and novelty in infant cross-language studies: Factors, problems, and a possible solution. In C. Rovee-Collier (Ed) *Advances in Infancy Research*, 12, pages nos to come.
 2. McGurk, H. Infant discrimination of orientation. *Journal of Experimental Child Psychology*, 14, 151-164.
 3. Aronson, E. & Rosenbloom, S. (1971) Space perception in early infancy: Perception within a common auditory-visual space. *Science*, 172, 1161-3.
 4. McGurk, H., & Lewis, M. (1974) Space perception in early infancy: Perception within a common auditory-visual space? *Science*, 186, 649-650.
 5. Field, J., Muir, D., Pilon, R., Sinclair, M., & Dodwell, P. (1980) Infants' orientation to lateral sounds from birth to three months. *Child Development*, 51, 295-298.
 6. Burnham, D., Taplin, J., Henderson-Smart, D., Earnshaw-Brown, L. & O'Grady, B. (1993) Maturation of precedence effect thresholds: full-term and preterm infants. *Infant Behavior and Development*, 16, 213-232.
 7. McKenzie, B.E. & Day, R.H. (1972) Object distance as a determinant of visual fixation in early infancy. *Science*, 178, 1108-1110.
 8. McGurk, H., & MacDonald, J. (1978) Auditory-visual coordination in the first year of life. *International Journal of Behavioral Development*, 1, 229-239.
 9. McGurk, H., (1988) Developmental Psychology and the Vision of Speech. Inaugural Professorial Lecture, University of Surrey.
 10. Massaro, D. (1998) Perceiving talking faces: From speech perception to a behavioral principle. Cambridge, Massachusetts: MIT Press. (page 25)
 11. McGurk, H., & MacDonald J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746-748.
 12. Dodd, B. (1977). The role of vision in the perception of speech. *Perception*, 6, 31-40.
 13. Dodd, B. & Campbell, R. (Eds.) (1987) *Hearing by eye: The psychology of lip-reading*. Hillsdale, New Jersey: Lawrence Erlbaum.
 14. Summerfield, Q. (1987). Some preliminaries to a comprehensive account of audio-visual speech perception. In B. Dodd & R. Campbell (Eds.) *Hearing by eye: The psychology of lip-reading*. Hillsdale, New Jersey: Lawrence Erlbaum.
1. Burnham, D. & Dodd (1998) Familiarity and novelty in infant cross-language studies: Factors, problems, and a