Available online at www.sciencedirect.com

SciVerse ScienceDirect

Speech Communication 55 (2013) 600–605

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

# Investigating the impact of lip visibility and talking style on speechreading performance

Karen Lander *, Cheryl Capek

*School of Psychological Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK*

Received 26 October 2012; received in revised form 18 December 2012; accepted 15 January 2013
Available online 30 January 2013

**Abstract**

It has long been known that visual information from a talker's mouth and face plays an important role in the perception and understanding of spoken language. The reported experiments explore the impact of lip visibility (Experiments 1 & 2) and speaking style (Experiment 2) on talker speechreadability. Specifically we compare speechreading performance (words in Experiment 1; sentences in Experiment 2 with low level auditory input) from talkers with natural lips, with brightly coloured lips and with concealed lips. Results reveal that highlighting the lip area by the application of lipstick or concealer improves speechreading, relative to natural lips. Furthermore, speaking in a clear (rather than conversational) manner improves speechreading performance, with no interaction between lip visibility and speaking style. Results are discussed in relation to practical methods of improving speechreading and in relation to attention and movement parameters.


*Keywords:* Speechreading, Speechreadability; Lip visibility, Speaking style

## 1. Introduction

Visual information from a talker's mouth and face plays an important role in the perception and understanding of spoken language (Massaro, 1998; Summerfield, 1992). Under noisy conditions, viewing the talking face supplements the auditory signal, increasing perceptual accuracy (Sumby and Pollack, 1954). Furthermore, we make use of visual information from the face even during the understanding of clear and unambiguous speech (Reisberg et al., 1987). Finally, there is evidence that access to visual speech is necessary for normal speech development (Mills, 1987).

The use of the visual speech signal, or speechreading is an important aid to communication, both within the hearing population and for the estimated nine million deaf and hearing-impaired individuals living in the UK (Valentine and Skelton, 2008). For speechreading, the lip and mouth area is of huge importance. Indeed automatic speechreading by machines is typically based on a robust lip image analysis (Vogt, 1997), often including modelled information about the teeth and tongue. Marassa and Lansing (1995) and Ijsseldijk (1992) found that information from the lips and mouth region alone is sufficient for word recognition, and that adding facial motion in other areas did not increase speech perception significantly. Conversely, considerable research has found that facial movements outside the oral area (for example, of eyes, cheeks) also provides useful information for speechreading (see Massaro, 1998; Preminger et al., 1998 ). Movement of the extraoral areas of the face are known to be highly correlated with movement of the oral articulators (Munhall and Vatikiotis-Bateson, 1998). Indeed it is now established that the mouth region is the primary source of visual speech information, additional and correlated information is available elsewhere on the face (see Lansing and McConkie, 2003; Rosenblum et al., 1996).

* Corresponding author. Tel.: +44 (0) 161 275 2598; fax: +44 (0) 161 275 2588.
*E-mail address:* karen.lander@manchester.ac.uk (K. Lander).

0167-6393/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.specom.2013.01.003

Given the importance of communicative importance of speechreading, a number of researchers have aimed to look why some talkers are easier to speechread than others. This is an interesting issue as (Demorest and Bernstein, 1992) estimated that 4.9% of speechreading variability can be accounted for by the talker. Factors such as lip shape, speech rate (Massaro et al., 1993) and talker familiarity (Lander and Davies, 2008) have all been proposed to influence talker 'speechreadability'. Furthermore, talkers naturally vary in the number of distinct visemes (visible speech token) that can be identified from their speech. This too can have an impact on intelligibility (Kricos and Lesner, 1985), with talkers who have a larger number of distinct visemes tending to produce speech that is easier to speechread. Finally, more recent work by Irwin et al. (2011) has proposed that accent type may have an impact on visual speech intelligibility.

Despite this research, relatively little experimental work has directly tested the importance of specific lip variations on speechreadability in a naturalistic setting. One practical change that talkers can readily make to their lips is the application of lipstick. Early work by McGrath (1985) used luminous lipstick and ultraviolet light to restrict displays, so that just the lips or lips and teeth were visible. McGrath (1985) aimed to compare human performance to that of a synthesized talking head, rather than to look at the impact of lip visibility per se. Thus, the aim of Experiment 1 is to investigate the impact of lip visibility on speechreading performance, by asking participants to speechread single words from three talkers wearing bright lipstick, with natural lips or wearing concealer on their lips.

## 2. Experiment 1: The effect of talker identity and lip condition on talker speechreadability (words)

### 2.1. Materials and methods

#### 2.1.1. Participants

Sixty undergraduate students (30 male; 30 female) aged between 18 and 31 years old (mean 20 years) at the University of Manchester took part in the experiment. None had taken part in any other speechreading experiments of this kind.

#### 2.1.2. Materials

Stimuli for the speechreading task were recorded from three unfamiliar talkers. The talkers were 22, 23 and 25-year old females, none of whom had a strong regional accent. The accents of the participants were not formally recorded but most did not have strong regional accents either. All speakers spoke the words in a normal clear, but not exaggerated manner and were not aware that the stimuli were to be used in a speechreading study. The words were recorded in three conditions; (1) with a bright red lipstick applied; (2) without any lipstick applied (natural); (3) with concealer applied (see Fig. 1). The order of recording by lip conditions was counterbalanced across talkers and talkers were not informed about the specific purpose of the filmed clips. All clips displayed the head and shoulders of the person from a frontal viewpoint. Filmed individuals maintained a fixed distance of approximately 1.1 m from the camcorder. Lighting during filming consisted of one overhead fluorescent strip light and one 45 W light positioned below the face. The resulting colour full face clips were displayed on computer (Power Mac G4; refresh rate 60 Hz), showed the speaker's head and shoulders and were 12 × 9 cm in size. However, the size of the face on the screen varied slightly in width due to the nature of the footage.

Each talker was recorded saying 63 single simple words (three practise), chosen from the University of Western Australia MRC Psycholinguistic Database, (http://websites.psychology.uwa.edu.au/school/MRCDatabase/mrc2.html). The 60 target words were split into three groups of 20, matched for age of acquisition and concreteness. Each participant was presented with all three word lists from a single talker. Thus, listeners heard one list of

Fig. 1. Examples of images used in Experiments 1 and 2 (colour images used in experiments). The top image shows the bright lips, middle image shows natural lips and the bottom image shows concealed lips.