Speechreading for information gathering:

A survey of scientific sources[1]

Ruth Campbell Ph.D, with Tara-Jane Ellis Mohammed, Ph.D

Division of Psychology and Language Sciences

University College London

Spring 2010

Contents

1 Introduction

1.1 This report aims to clarify what is and is not possible in relation to speechreading, and to the development of speechreading skills. It has been designed to be used by agencies which may wish to make use of speechreading for a variety of reasons, but it focuses on requirements in relation to understanding silent speech for information gathering purposes. It provides the main evidence base for the report : ***Guidance for organizations planning to use lipreading for information gathering*** (Ruth Campbell) - a further outcome of this project.

1.2 The report is based on published, peer-reviewed findings wherever possible. There are many gaps in the evidence base. Research to date has focussed on watching a single talker's speech actions. The skills of lipreaders have been scrutinised primarily to help improve communication between the lipreader (typically a deaf or deafened person) and the speaking hearing population. Tests have been developed to assess individual differences in speechreading skill. Many of these are tabulated below (section 3). It should be noted however that:

- There is no reliable scientific research data related to lipreading conversations between different talkers.
- There are no published studies of expert forensic lipreaders' skills in relation to information gathering requirements (transcript preparation, accuracy and confidence).
- There are no published studies relating individual performance on a specific test of lipreading (see section 3) with performance in relation to information gathering requirements. Studies exploring this, using a variety of test materials closer to those used by agencies requiring speechreading interpretation, are underway.

Because inferences have to be drawn from the published and reported studies which arose in a different context, caution and an empirical approach should be applied in considering how lipreading may be used for information gathering purposes.

2 Questions and answers

2.1 Lipreading or speechreading? Lipreading/Speechreading is the skill of understanding speech by watching the talker's actions. **Lipreading** is the historical term used to describe this. **Speechreading** (a more recent term) is often preferred, since understanding speech makes use not just of watching the talker's mouth movements, but of her head, face, eye and torso actions as well.[2] The terms are used interchangeably in this report.

2.2 Who lipreads/speechreads? Lipreading is practised by hard-of-hearing and by deaf people in order to understand speech which they cannot hear, or hear clearly.[3] Every hearing person has some capacity for speechreading, since they are better at understanding speech when they can see the talker, and are worse when the talker's mouth movements do not correspond to what is heard (i.e. bad dubbing).[4] Some people become so good at speechreading that their comprehension of silent (i.e. lipread) speech, in a face-to-face setting, can seem indistinguishable from that of a hearing person listening to clear speech.[5]

2.3 What makes for good speechreading? These are the main factors that are likely to lead to good speechreading, as determined by performance on silent speechreading tasks.

2.31 Lipreading **experience** – either as a deaf or deafened person, or as someone close to deaf people, such as a spouse of a deaf person.[6] Historically, people who worked in noisy environments (machine weavers for example), became skilled lipreaders.[7]

2.32 Good **language** knowledge, especially vocabulary skills.[8] Good speechreaders tend to be good readers, too.[9] People are better at speechreading when they are familiar with the language being spoken.[10]

2.33 Normal vision, with acute sensitivity to visual movement.[11]

2.34 Good **verbal short term memory.** This contributes especially to sentence and discourse comprehension.[12]

2.35 **Familiarity** with the talker, and the talker's accent and speech style.[13]

2.36 Risk-taking **personality** traits (willingness to 'have a go', guessing).[14]

2.37 **Age**: younger speechreaders (20-50 years) are better than older (>65 years), generally.[15]

2.38 Gender, schooling, sign-language experience, scientific training have not been shown to affect speechreading skill reliably. People with higher IQ scores tend to be better speechreaders on some tests.[16] Other personality traits (e.g. tenacity, self-confidence) cognitive traits (e.g. executive function, attention switching, visual memory) are underexplored.

2.39 Many of these factors also characterise a listener's abilities in speech perception in noise (SpIN), where cognitive factors have been explored to a greater extent than they have for lipreading.[17]

2.4 Why is lipreading difficult? Why does it show so much variation from person to person?

Speech is designed primarily to be heard, not viewed. Blind children develop good speech although they cannot see a talker's vocal actions.[18] Many of the critical aspects of speech are hidden from view. Thus, most consonants are produced by actions of the tongue inside the oral cavity (g,d,t,y,r,s,n,k,sh,s,j,z) and not by visible actions of tongue, lips, teeth (m/p, f/v, th) . Also, lipshape does not always reflect the speech sound being made, but can anticipate or follow it: mouthshape for the final 'th' varies for saying 'too**th**' and 'tee**th**' because of the preceding vowel (watch yourself say these in the mirror).[19]

2.41 For these reasons, it is often concluded that only around 30% of spoken English is lip readable.[20] However, this is misleading as it is too general. Firstly, people vary in their lipreading ability, and better lipreaders can see more distinctive speech elements than poor lipreaders[21]. Secondly, some distinctions between speech actions are highly visible to good and poor lipreaders alike. For example, when presented with a full range of seen **syllables ,** up to 60% accuracy can be achieved by naïve hearing perceivers, as long as the dynamic characteristics of the speaking face are well captured (i.e. the talker's face shows natural movement) and the critical parts of the mouth can be seen[22].

2.42 Focussing on **consonants**, while a few consonants may be reliably distinguished from each other 'by eye'('m' compared with 'n', or 'th' compared with 'f'), lipreading of most consonants is liable to error. For example, 'd','t','k' all look alike on the lips.[23] Some consonant sounds

that are easily lost in noise ('f' and 'th', for example) can be readily perceived by eye. That is, lipreading can *compensate* to some degree for hearing loss.[24]

2.43 In contrast to consonants, many **vowels** in English have distinctive mouth shapes ('a','i','u') and can be readily identified by sight. Since many English words are distinguished by their vowel patterns these can afford some clues to word understanding, particularly for words that have few 'lexical neighbours'.[25]

- The complex pattern of speech elements (consonants, vowels, and their combination in syllables and words) in the continuous speech stream means that the speechreadability of a particular utterance will vary. Some utterances are easier to see than others. These are likely to contain visible consonants and clearly articulated vowel sequences, and contain words that are distinctive in relation to other words.
- Some speechreading can occur despite the impoverishment of the signal. Context provides many useful cues, whether from information outside the visual record, or from the visual record itself[26]. Because such 'top-down' processes constrain the set of likely interpretations, some lipreading tasks may become more feasible. For instance, deciding whether a *particular* given word was or was not spoken may be achieved more accurately than simply identifying a word when one does not know what it might be.
- Accent, speech style, and even language, can sometimes be 'seen', even though individual words may be hard to identify (see 2.32 above).

2.44 **Exposure and training** In principle, *anyone* who has natural audiovisual experience of a spoken language can lipread to some extent. For instance, people cannot *avoid* making use of seen speech actions when interpreting heard speech.[27] However, this does not mean that everyone can understand speech simply by watching it. Because lipreading is intrinsically harder than hearing, most people with good hearing do not consciously develop their lipreading skills, and rely more on what they hear. Nevertheless, when the environment is noisy or when hearing loss occurs, lipreading skill comes to the fore – though to a variable extent from person to person. Lipreading training can help people develop confidence in this ability, and can offer useful tips and clues. It is not clear that any specific lipreading training regime makes a reliable difference to actual performance, and some training schemes can depress performance accuracy.[28]

2.5 Can lipreading be used as an expert professional skill?

Media organisations use lipreaders to interpret video clips of celebrities/sports people talking.[29] Expert lipreaders also interpret speech from patients who can no longer vocalise, in order to 'translate' the silent speech into an audible or written form.[30] The main forensic use of speechreading is to identify the content of a spoken conversation recorded under surveillance. But it can also be used to inform about the language being spoken ('is he speaking English?'), or speech style, including accent ('Is he from Scotland?'). These aspects can be visible to the speechreader familiar with the language(s) or speech style.[31] Transcripts based on a lipread interpretation may be used evidentially, but only with great care, given the inherent ambiguity of most silent speechread utterances. [32] Transcripts are very time-consuming to prepare, often requiring multiple viewings of video clips.[33]

2.6 Which aspects of visual image capture are important for speechreading?

2.61 Lighting Top illumination is best for most face-processing tasks, as it most closely resembles most natural (daylight) illumination.[34] However, faces where overhead lighting produces shadow in the mouth area are hard to lipread. Thus, for speechreading, frontal illumination is preferred. Under these conditions, changes in luminance (within range 0.3-30 footlamberts) do not seem to have a marked effect. [35]

2.62 View Full or ¾ face views are generally better, but profiles can be speechread, too. [36] Live lipreading between a lipreader and a talker allows conversational interaction, which is helpful in following discourse since the speechreader can interrogate the talker. To date, with current image technology, there is no indication that a 3D image is significantly easier to speechread than a 2D image.[37]

- *2.*621 Surveillance cameras for speech capture should be positioned at a height to obtain a reasonable view of the mouth, avoiding shadowing of the mouth by top or back illumination. This is likely to be slightly in front of and above head height for close view, although for distant views a higher camera position can still capture mouth actions.[38]

2.63 Colour. No published studies report an advantage for colour over greyscale video speechreading. What is important is to maintain appropriate contrast and brightness relations across the skin, lips, teeth, tongue and inner part of the mouth.[39]

2.64 Distance There is a fall-off in accuracy with apparent distance between camera/viewer and talker[40]. However, as long as the talker's face can be seen, positive effects of seeing the speaker are apparent and some level of intelligibility is likely.[41] Because camera/screen size is limited, there is a trade-off between image size and distance of talker from camera.

- 2.641 Which is better: a close-up of the mouth/head that fills the screen, or a smaller view of head and shoulders, and possibly torso, in a scene? On the whole, it is better to use camera settings to capture torso and head of the talker, rather than just the face or mouth. Not only

is this less prone to loss of information due to the talker's movements, but it allows non-oral movements that are associated with speech to be taken into account. These can help interpret mouth actions.[42]

2.65 Looking Speechreaders rest their gaze on the talker's face in the eye region, not on the talker's lips, whether they are viewing a live or recorded talker.[43] They can glance (saccade) to the mouth region for specific information that might be predicted from the interpretation up to that point. Thus, a good deal of the information used in speechreading is not dependent on *foveal* (central vision) processes, but includes *parafoveal* (peripheral vision) processing too.

- 2.651 This dynamic natural attentional process cannot be simulated by zooming a camera plane of vision to bring it 'closer'. This simply interrupts human interpretation. The lipreader's anticipation of upcoming speech drives their shifts of attention.[44]

2.66 Image compression Optimal image and frame rate characteristics are thought to be 352*288 pixels and 25 fps.[45] However, it is possible to distinguish lipspoken utterances at reduced frame rates/ image size. Reduced frame rate (to around 12 fps) and reduced image quality, when applied together, make for extremely difficult lipreading.[46]

2.661 High-speed photography (around 120 fps) can, when examined at a slower rate, indicate some distinctive features of speech that may not be apparent in 'natural' speechreading by non-experts. For instance, the 'cheek-puff' that accompanies a plosive bilabial /p/ (' a **p**ath') can be distinguished from the non-plosive /b/ ('a **b**ath'). Good speechreaders can pick up these 'micro-actions', invisible to the naïve observer.[47]

2.662 In attempting to understand spoken words or sentences, applying distortions such as visual amplitude changes to the image of the mouth (to attempt to exaggerate the lip patterns), or reducing frame rate (to try to capture just the salient mouth patterns) are harmful. They upset the critical time-varying aspects of speech actions, which are essential for visible speech comprehension.[48]

2.7 Is every talker speech readable?

Not all talkers are equally easy to speechread. Most tests of speechreading tend to use talkers who are clean-shaven, speaking using normal voice (neither whispering nor shouting). These are usually judged easiest to speechread.[49] There is no compelling evidence for *systematic* talker differences based on age, gender, skin colour or perceived ethnicity. As previously indicated, the best indicator of a talker's speech readability is *familiarity* – with the talker, their speech style, and accent.

3 Chronologically organised survey of tests of Speechreading In English [50]

AUTHOR(S); DATE FIRST REPORTED; TEST NAME	COUNTRY OF ORIGIN	TEST MATERIAL	NUMBER OF TALKERS	MANNER OF PRESENTATION	MANNER OF RESPONSE & SCORING	TARGET POPULATION	NB
Nitchie 1913	USA	3 proverbs: ''Tis love that makes the world go round', 'Spare the rod and spoil the child' & 'Fine feathers make fine birds'.	One (male)	Filmed (moving-picture camera operated at a speed of 16 pictures per second)	--	Lipreading students	Developed as a teaching / learning aid to illustrate points about lipreading.
Kitson 1915	USA	No formal test constructed – teachers' judgments of speechreading aptitude were used to rank subjects.	--	--	Subjects ranked	Adult lipreading students	This was not a constructed test, and ranking is not an ideal way to measure speechreading ability, but this study has been described as 'the first experimental investigation of lipreading' (O'Neill & Oyer, 1961, pg. 36).
Conklin 1917	USA	8 consonants, 52 familiar words (selected to present all the sounds of English) & 20 simple sentences (10 from Nitchie's 1912 manual [see revised ed.: Nitchie 1930], 10 used regularly in the classroom by teachers for the deaf)	One	Live, each item repeated 3 times	Written; 1 point per consonant, 1 per word & up to 5 per sentence giving a possible total of 160 points	Deaf school pupils	This test helped stimulate general interest in speechreading tests. (O'Neill & Oyer, 1961, pg. 23).
Göpfert 1923	Germany	Lists of vowels, consonants, words & sentences		Live		Children & adults	
Day, Funsfeld & Pintner 1928	USA	? 2 or 4 lists, each of 10 sentences		Live	Written		Described by Reid (1946) as the 'first attempt at the objective measurement of lip-reading ability' (pg. 404). It was administered to 8300 pupils, but with 'little success' according to Markides (1980)
Goldman 1929	Germany	Words & sentences		Live			(Continued Göpfert's work)
Mason 1932	USA	24 words testing for word accuracy & initial consonant accuracy & 12 words testing for vowel accuracy	Talkers of different nationalities	Filmed			
Heider & Heider 1940	USA	3 parallel series of: 30 nouns; 30 independent sentences; 2 stories.	One	Filmed			Described by Reid (1946) as 'the most extensive report to date on the use of motion pictures' (pg. 404)

Mason 1943 A Cinemato- -graphic Technique for Testing Visual Speech Comprehension	USA	Test I (forms A & B): 5 simple nouns (e.g. baby) aimed at pre-school children; Test II (forms A & B): 10 slightly harder nouns (e.g. chair) Test III (forms A & B): An extension of 1 & 2 – adding 15 nouns, 5 included in the test but not introduced.	One (female)	Filmed (16mm motion-picture); black & white	Multiple choice: children were required to draw a large cross on the picture of the word spoken.	Deaf & Hearing impaired (HI) children	The test was too easy, and the author planned to continue its development. Unfortunately her work was not continued after her death in 1949 (Berger 1972)
Pauls 1947	USA	Continuous discourse: 'carefully selected and edited commercial shorts, cuttings from feature pictures, and also certain Navy training films' (pg. 269)	Several - varies between film clips	Filmed (could be presented with or without sound)	[not specified]	Deafened and HI adults at U.S. Naval Hospital	
Utley 1946 "How well can you read lips?"	USA	2 forms of 31 sentences; 2 forms of 36 words; 6 stories (5 questions about each).	One for words & sentences, four for stories	Silent film (1 hr 15)	Written; Max score: 190	Deaf / HI children & adults	This test was used to estimate speechreading ability at the Hearing and Speech Centre at Gallaudet College in its 1st 2 yrs ('58–'59). There has been some discussion of validity and difficulty (e.g. DiCarlo & Kataja, 1951)
Reid 1946	USA	3 forms, each with 5 parts: 17 vowels & diphthongs, 11 consonants, 10 unrelated sentences, a series of related sentences telling a story (title of story & character names given), & a short story with 4 questions	One for each form – each spoken by a different adult talker (2 female, 1 male)	Filmed (8mm, colour, lower ¾ of face & upper part of shoulders only),	Phonemes: multiple choice (underline word containing phoneme); Sentences: written repetition; Story: answer questions	Deaf children	Author records problems with asking children to give written responses: it doesn't allow children to demonstrate their full ability
Morkovin 1947	USA	10 everyday experiences followed by questions	--	Filmed	--	--	Markides (1980) reports this as a test, but Morkovin (1947) does not – he simply advocates teaching speechreading using real life material rather than drilling, based mostly on anecdotes.
Cavender 1949	USA	4 sets of 10 practice and 45 test sentences	one	Live, without voice	Multiple choice – underline the word (from a choice of 5) that occurred in the sentence	Hard-of-hearing children	
Kelly 1953	USA	3 sections: (1) 15 3-letter items (e.g. AIE, YBU, IGM, etc); (2) 10 'words out of context', 5 with 2 words to choose from, 5 with 3; (3) 10 sentences, 3-5 words long, 7 declarative, 3 interrogative.		Live			

Watson 1957 The New Manchester Picture Test		UK	8 lists of 10 CVC words, 5 in each designed to test vowel discrimination & 5 consonant discrimination		Live	Multiple choice picture pointing (black & white pictures)	Developed as a test of hearing for speech for children from 5 yrs upwards	Test described by Hickson, 1987. Elphick (1996) reported that this test is commonly used by teachers of the deaf for quick assessment of children's AO, AV & VO ability.
A Film Test of Lipreading (from The John Tracy Clinic)	Lowell & Taaffe 1957	USA	2 forms of 30 sentences; I point for each correct word, possible total of 188 for each form. The two forms were constructed by ranking Keaster's 60 sentences & splitting them into 2 forms; (they had previously been in 6 lists of 10 sentences, each recorded with a different talker in b&w & colour).	One (male)	Film	Written	Deaf / HI adults & children	This test is also called the 'Iowa-Keaster Test', and the 'Keaster Film Test of Lipreading'; it used Keaster's sentences, developed at the State University of Iowa. Spitzer et al. (1987) reported this test as widely used. Their results showed it to be easier than the CID sentences and the Gold Rush paragraph
	Donnelly & Marshall 1967	USA	Development of Lowell & Taaffe's test, same 2 forms of 30 sentences, possible total now 30	One (male)	Film (with or without sound)	Multiple choice (derived from written responses)	Deaf / HI adults (university students)	
Moser, Oyer, O'Neill & Gardner 1960		USA	1-syllable words taken from Voelker's (1942) list of the 1000 most frequently spoken words	Four	Film			
Harris, Haines, Kelsey & Clack 1961 Harris' Revised CID Everyday Sentences Lists		USA	Revision of the 10 CID everyday sentence lists (Silverman & Hirsh 1955) – key words retained, sentence length controlled more stringently. Each list has 10 sentences, and 50 key words.		Varies	Open response (spoken / written)	Adults	Developed to assess a *listener's* understanding of speech
Montgomery 1966 Donaldson Lip-reading Test		UK (Scotland)	40 sentences increasing in difficulty; 10 pages of 6-9 black & white line drawings, each page used 4 times		Live	Multiple choice (Picture pointing with photographs)	Children of all school ages	The test was standardised on all profoundly prelingually deaf children born in Scotland in 1941, nearly all of those born in 1948, and some in other year groups.
Katt 1967		USA	Part of this test (the only part described by Smith & Kitchen, 1972) was 2 lists of 16 unrelated sentences, mean no. words: 4, mean no. syllables: 4.53; 26 declarative, 5 interrogative, 1 exclamatory		8mm film	Written		(unpublished masters thesis at Michigan State University)
Boothroyd 1968			Lists of 10 CVC words built from the	One adult	Usually live	Repetition	Adults	Elphick (1996) reports that this test

AB Isophonemic word test		same 10 vowels and 20 consonants					is commonly used by teachers of the deaf for quick assessment of children's AO, AV & VO ability
Myklebust & Neyhus 1970) Diagnostic Test of Speechreading	USA	Word, phrase and sentence stimuli. Lexical items recur in different sections of the test			Closed-set picture identification	Deaf children	
Nielsen 1970	Denmark	9 sentences (4-9 words) in an everyday scene: 2 adults drinking coffee. Simulated everyday situations used so that the analysis of situations through gestures was possible	Two adults	Colour film, without sound, length: approx. 4 minutes	Verbal repetition; Sentences scored as correct (1 point) or incorrect (0 point)	Hearing impaired adults	Test gives an estimate of ability. 0-1 point: weak (20%); 2-6 points: average (60%); 7-9 points: good speechreader (20%).
Ludvigsen (1974) (Pilot study: Ewertsen 1973) The Danish HELEN Test	Denmark	8 lists of 25 relatively simple questions in 5 broad categories (before/after, colours, opposites arithmetic, miscellaneous), (has been translated & adapted for other languages - see e.g. Plant et al. 1982)			Verbal – 1 or 2 word answer to questions	Adults	
Jones & Whitehead 1975 The NTID Phoneme Identification Test	USA	15 CV syllables (consonant + /a/), each appearing 8 times giving a total of 120 items		Video, with sound	Circle consonant from closed choice on an opscan answer sheet	Hearing impaired adults	
Binnie, Jackson & Montgomery 1976	USA	20 consonants in CV environment with /a/, each repeated 5 times in a random order (total 100 items)	One (female)	Film without sound	Written. Fixed choice of consonant.	Hearing impaired adults	
Kalikow Stevens & Elliott 1977 Speech Intelligibility in Noise (SPIN) Test	USA	8 lists of 50 sentences, varying from 5 to 8 words in length: 25 high predictability & 25 low predictability sentences mixed randomly in each list.		Usually presented with background noise (e.g. speech babble)	Verbal repetition of last word in sentence; Only the last word (noun) of each sentence is scored	Adults	reported to be difficult (Martin et al., 1983; Gagné et al., 1987), but useful for evaluating the benefit of hearing aids or for cochlear implant evaluation
De Filippo & Scott 1978 Continuous Discourse Tracking (CDT)	Varies according to script used			Live			reported to have been incorporated into many training programs for cochlear implant recipients (Gagné et al., 1987), but has limitations for assessing speech-reception performance
Bamford, Kowal & Bench 1979 BKB Sentence Lists	UK + Australia	21 lists of 16 sentences varying from 4 to 7 words in length, each list has a designated 50 key words	Usually one	Varies	Open response (spoken / written); key word scoring	> 8 yrs	A widely used English test of open-set sentence recognition; developed for speech audiometry

Markides 1980 The Manchester Speechreading (Lipreading) Test		UK	2 lists of 33 CVC words (99 phonemes); 2 lists of 25 sentences (5 each of 2 wd, 3 wd, 4 wd, 5 wd, 6 wd = 100 words). Words & sentences selected after testing larger pool with 120 hearing 6 year olds Later papers refer to 4 lists	One	Originally live, no voice	Written or Verbal repetition (tester writes response down); Words scored per phoneme, sentences per word	British HI children & adults	
Plant & MacRae 1981 NAL (National Acoustics Laboratory) Test		Australia	50 questions in 5 categories (e.g. questions about you). Topic given prior to each set of questions.			Responses to questions.		
Bannister & Britten 1982 Eyes and Spoken Language (EASL)		USA	5 tasks: (I) fill in the blank (noun); (II) sentence completion (modifier / verb complement); (III) sentence recognition, 1 key word given; (IV) sentence recognition, no cues; (V) question response	One	Video, no sound	Written	Hearing impaired adults	Test developed to assess how well a person uses linguistic constraints in responding to visual only spoken language
Tyler, Preece & Tye-Murray 1986	The Iowa Laser Videodisc Consonant Confusion Test	USA	13 consonants presented in an [aCa] context	One (male)	Laser videodisc			'created for study of visual speech perception'
	The Iowa Sentence Tests	USA	6 lists of 30 sentences varying from 4 to 7 words in length, with 88 key words			Repetition	Adults	
Bernstein & Eberhardt 1986 John Hopkins Lipreading Corpus		USA	Corpus I-II: Disc I: CVC words Corpus III-IV: Disc 2: sentences ('E-B sentences')	Two (1 male, 1 female)	Laser videodisc	Typed repetition		
Spitzer, Leder, Milner, Flevaris-Phillips & Giolas 1987 The Gold Rush Paragraph		USA	A paragraph followed by 6 yes/no questions	One (either male or female)	Video (viewed on either a b&w or colour monitor)	Written responses to questions.	Cochlear implant candidates	This test did not provide an adequate range of scores, and was considered unsuitable for testing connected discourse.
Cox, Alexander & Gilmore 1987 The Connected Speech Test			Sets of related sentences spoken conversationally, topic given before each set	One		Verbal repetition		
Gagné Seewald & Stouffer 1987 Visual Consonant Recognition Test		Canada	18 English consonants in /aCa/ context, each presented 5 times in random order (making 90 items)	One (female Canadian experienced in monitored live-voice speech production)	Video (colour)		Hearing / HI Canadian adults	

Tye-Murray, Tyler, Bong & Nares 1988	USA					HI children & adults	
Tye-Murray, Purdy, Woodworth & Tyler 1990	USA	50 primary sentences	Six	Laser videodisc	Verbal repetition		
Gagné, Tugby & Michaud 1991 Speechreading Test on the Utilization of Contextual Cues (STUCC)	Canada	208 test items consisting of 104 sentences, each presented in a related & unrelated context. Each item consists of an introductory sentence (related / unrelated to test sentence), then the test sentence.	One (female)	Video, without sound, slowed rate of speech, introductory sentences were spoken (silent) & captioned	Written; Key word scoring	Adults with an acquired hearing loss	Test items selected from 198 modified SPIN sentences. The authors reported that further test development was needed & warranted.
Boothroyd 1991 CUNY sentence lists		72 lists of 12 sentences related to 12 known topics, varying from 3 to 14 words in length			Repetition	Adults	
Tye-Murray, Witt & Castelloe 1996 The Sentence Gist Recognition Test	USA	50 sentences, length 5 to 8 words, in 6 topically related sets (e.g. a restaurant), cued by short film clips. When subjects get an item wrong they are given a choice of 5 repair strategies (repeat, key word, elaborate, simplify, rephrase). Two practice sets were presented before testing began.	Thirteen talkers (8 female, 5 male; aged from childhood to middle age)	Laser videodisc; AV with 6 talker babble;	Multiple choice – picture illustrating each item from choice of 6	HI adults	

3.1 **Speechreading Tests:** Reference List

Bannister, M. L. & Britten, C. F. 1982. Linguistically based speechreading assessment. *Journal of Communication Disorders* 15: 475-479.

Boothroyd, A. 1968. Developments in speech audiometry. *Sound* 2: 3-10.

Boothroyd, A. 1991. CASPER: a user-friendly system for Computer Assisted Speech Perception Testing and Training. New York: City University of New York.

Elphick, R. 1996. Issues in comparing the speechreading abilities of hearing-impaired and hearing 15 to 16 year-old pupils. *British Journal of Educational Psychology* 66: 357-365.

Ewertsen, H. W. 1973. Auditive, visual and audio-visual perception of speech. Copenhagen: The Helen Group, State Hearing Centre.

Kalikow, D. H., Stevens, J. N. & Elliott, L. L. 1977. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America* 71: 1211-1224.

Myklebust, H. R. & Neyhus, A. I. 1970. Diagnostic Test of Speechreading. New York: Grune and Stratton.

Plant, G., Phillips, D. & Tsembis, J. 1982. An auditory-visual speech text for the elderly hearing impaired. *Australian Journal of Audiology* 4: 62-68.

Plant, G. L. & MacRae, J. H. 1981. The NAL lipreading test development, standardisation and validation. *Australian Journal of Audiology* 3: 49-57.

Silverman, S. R. & Hirsh, I. 1955. Problems related to the use of speech in clinical audiometry. *Annals of Otology, Rhynology and Lanryngology* 64: 1234-1244.

Tyler, R. S., Preece, J. P. & Lowder, M. 1983. Iowa cochlear implant tests. University of Iowa, Iowa.

Tyler, R. S., Preece, J. P. & Tye-Murray, N. 1986. The Iowa phoneme and sentence tests. University of Iowa, Iowa.

4 Further sources of scientific research on lipreading

4.1 Research groups and teams

http://mambo.ucsc.edu/psl/lipr.html (last updated around 2000)

http://www.haskins.yale.edu/AVISA/AVISA.html (last updated 2000)

AVSP annual conference: this annual conference of researchers working on speechreading includes researchers in 'automatic' speech recognition (ie speechreading by computational engines), brain bases for speechreading – and linguistic and psychological explorations of audiovisual speech and speechreading. The list of attendees reflects the current research teams involved http**://www.avsp2009.co.uk/ (last updated, October 09. Links to earlier AVSP meetings (archives))**

4.2 Books

J.Jeffers & M.Barley (1971 , updated 1980) *Speechreading* Charles C Thomas, Springfield Ill USA

D.Stork and M.Henneke (1996) *Speechreading by Man and Machines* NATO symposium, Springer, Germany (available as Google book)

B.Dodd and R. Campbell (1986) *Hearing by Eye* Erlbaum, Hove,UK

R.Campbell, B.Dodd and D Burnham (1997) *Hearing by Eye II,* Psychology Press, Hove, UK

D.Massaro (1997) *Speech perception by ear and by eye* MIT Press

4.3 Thesis

Tara-Jane Mohammed (nee Ellis) 2007 *An Investigation of Speechreading in profoundly congenitally deaf British adults* PhD Thesis, University of London . accessible via http://catalogue.ulrls.lon.ac.uk/search~S1/q?author=Mohammed%2C+Tara&title=speechreading

5 Biography

Ruth Campbell is an experimental psychologist and neuropsychologist who retired from her position as Chair of Communication Disorder and Deputy Director of DCAL (Deafness, Cognition and Language Centre), University College London, in September 2008. She remains a member of DCAL and an emeritus professor at UCL. In addition to a PhD (London, 1979) she holds an honorary research position at Kings College London (Institute of Psychiatry) and an honorary doctorate from Linkoping University, Sweden (2007).

One of her primary research aims, over thirty years, has been to understand the psychological and neural bases of speechreading, and, in addition to over 150 or so peer-reviewed journal papers on this topic, she has edited or co-edited several volumes dedicated to putting speechreading on the map as a topic for psychological and interdisciplinary research. She was elected patron of ATLA in 2004. She gave an invited lecture course on speechreading for the Australian Research Council's annual interdisciplinary postgraduate 'Summerfest' (Sydney, Australia, 2008). An ABC featured radio interview with Professor Campbell explored speechreading at length. [51]

As an expert on speechreading (though not herself an expert speechreader) she has worked to develop a consistent and scientifically based approach to the forensic uses of lipreading. She has served as advisor and as an expert witness on landmark English cases including R. V Luttrell *et al*., a case now cited in authoritative legal reports on forensic evidence. [52] She collaborated with Laraine Callow (Director, *Deafworks*) to develop a unique innovative training course for Deaf forensic speechreaders in 2008. [53] This course was supported by a business grant from the Economic and Social Sciences Research Council of Great Britain (ESRC) to DCAL, and provided a basis for the collaborative project between expert deaf lipreaders and speech scientists reflected in the present report.

Ruth Campbell was senior academic supervisor for Tara-Jane Ellis Mohammed's PhD thesis (2007), from which section 3 (Lipreading Tests) has been extracted.

6 References

[1] The opinions expressed are those of the author. The material in this report should not be distributed without the permission of the author (r.campbell@ucl.ac.uk). Companion reports from this collaborative project, integrating the experiences of expert lipreaders and scientists working in audiovisual speech processing, are *Experience of Expert Lipreaders – (1)Interim Report, (2) Final Report*" (Laraine Callow & Sally Reynolds), January and March 2010. These may be obtained from Laraine Callow (lc@deafworks.co.uk).

[2]Berger KW (1972)*Speechreading: Principles and Methods,* National Educational Press, USA. See also Arnold P (1997) The structure and optimization of speechreading. *J.Deaf studies and deaf education 2 (4), 199-217*

[3] ditto

[4] Sumby WH, Pollock I.(1954) Visual contribution to speech intelligibility in noise *J.Acoustic Soc. America* , 26 (2) 212-15

[5] Lyxell B. 1994. Skilled speechreading: a single-case study. *Scandinavian J. Psychology* 35: 212-19; Rönnberg J *et al.(*1999) A speechreading expert: The case of MM *J. Speech, Language and Hearing Research* 42 (1), 5-20. Other single cases have also been reported.

[6] Bernstein LE *et al.* 2000. Speech perception without hearing. *Perception & Psychophysics* 62: 233-252; Auer ET, Bernstein, LE. (2007) Enhanced visual speech perception in individuals with early-onset hearing impairment. *J. Speech, Language, and Hearing Research*. Vol 50(5), 1157-1165; Ellis, T *et al.* (2001): "TAS: A new test of adult speechreading - deaf people really can be better speechreaders", In D.W.Massaro, J.Light, K.Geraci (Eds)*Auditory-Visual Speech Processing (AVSP 2001)*, Aalborg, Denmark, September 7-9, 2001, ISCA Archive, http://www.isca speech.org/archive/avsp01

[7] http://en.wikipedia.org/wiki/Amoskeag Manufacturing Company.

[8] Lyxell B, Holmberg I (2000) Visual speechreading and cognitive performance in hearing impaired and normal hearing children *Brit. J. Educational Psychology* 70, 505-18. Davies R. *et al* (2009) Investigating the psycholinguistic correlates of speechreading in preschool age children. *Int. J. Language and Communication Disorders* 44 (2) 164-74

[9] Mohammed T *et al* (2006) Speechreading and its association with reading *Clinical Linguistics and Phonetics* 20 (7-8),621-630; Harris M, Moreno, C. (2006). Speech reading and learning to read: a comparison of deaf children with good and poor reading ability. *J. Deaf Studies and Deaf Education. 11,* 189-201.

[10] Soto-Faraco S *et al* (2007) Discriminating Languages by speechreading *Attention, Perception & Psychophysics* 69, 218-31

[11] Mohammed T *et al* (2005) Speechreading skill and visual movement sensitivity are related in deaf speechreaders. *Perception* 34, 205-216

[12] Kitson, H. D. 1915. Psychological tests for lip-reading ability. *Volta Review* 17: 471-476; Lyxell, B. & Rönnberg, J. 1993. The effects of background noise and working memory capacity on speechreading performance. *Scandinavian Audiology* 22: 67-70

[13] Lander K, Davies R (2008) Does face familiarity influence speechreading? *Quarterly J. Experimental Psychology* 61 (7), 961-7; Irwin A *et al* (2007) Regional Accent Familiarity and Speechreading Performance. *Proceedings of the international conference AVSP2007* (eds J.Vroomen, M.Swarts and E.Kramer), ISCA available at: http://spitswww.uvt.nl/Fsw/Psychologie/AVSP2007/

[14] Lyxell B. Rönnberg, J. 1987. Guessing and speech-reading. *British Journal of Audiology* 21: 13-20; Mohammed T (2006) *TAS: a new test of adult speechreading* PhD thesis, University College London. Chapter 10: Risk-taking personality style, as assessed by questionnaire, was associated with better speechreading

[15] Tye-Murray N *et al* (2007) The effects of age and gender on lipreading abilities. *J American Academy of Audiology* 18(10):883-92.

[16] Rodríguez Ortiz Ide L. (2008)Lipreading in the prelingually deaf: what makes a skilled speechreader? *Spanish J Psychology* 11(2):488-502.

[17] Houtgast T, Festen JM (2008). On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise. *International Journal of Audiology, 47:*287-295.

[18] Mills AE (1987) Development of phonology in the blind child. In B.Dodd,R.Campbell (Eds) *Hearing by Eye*, London, Erlbaum, 145-61. Among other things this study found that some blind toddlers' speech showed (temporary) confusions because of the unavailability of visible speech contrasts ('**m**an' vs. '**n**an').

[19] Summerfield AQ. (1992) Lipreading and audio-visual speech perception. *Philosophical Transactions Royal Society London B*, 335(1273):71-8.

20 http://en.wikipedia.org/wiki/Lip_reading

[21] Auer ET Jr, Bernstein LE. (1997) Speechreading and the structure of the lexicon: computationally modeling the effects of reduced phonetic distinctiveness on lexical uniqueness. *J Acoustic Society of America* 102(6):3704-10.

22 Jiang J, Auer ET Jr, Alwan A, Keating PA, Bernstein LE. (2007) Similarity structure in visual speech perception and optical phonetic signals .*Perception and Psychophysics*. 69(7):1070-83

23 Campbell R. (2006). Audiovisual Speech processing. in Brown,K. et al. (ed.) *The Encyclopedia of language and linguistics*. Amsterdam, London, NY: Elsevier, 562-569

[24] ditto

[25] Numbers, M. E. (1939) An experiment in lip reading. *Volta Review*. 41, 261-264; Summerfield AQ (1986) Some preliminaries to a comprehensive account of audiovisual speech perception. In B.Dodd & R.Campbell (Eds) *Hearing by Eye* Erlbaum, Hove UK, 3-51

[26] See, for examples of context-driven speechreading: Rönnberg J *et al* (1996) Lipreading with auditory low-frequency information. Contextual constraints. *Scandinavian Audiology* 25 (2), 127-32; Samuelsson S, Rönnberg J (1991) Script activation in lipreading *Scandinavian J. Psychology* 32 (2), 124-43

[27] McGurk effects, see http://en.wikipedia.org/wiki/McGurk_effect

[28] Bernstein LE *et al* (2001) Enhanced speechreading in deaf adults: can short-term training/practice close the gap for hearing adults? *J Speech and Language, Hearing Research* 44(1):5-18; Blumsack JT, Bower CR, Ross ME.(2007) Comparison of speechreading training regimens. *Perceptual and Motor Skills*. 105(3 Pt 1):988-96; Binnie CA (1977) Attitude changes following speechreading training *Scandinavian Audiology* 6,(1), 13-19; Gagné JP *et al* (1991) An evaluation of CAST: a computer-aided speech training program *J.Speech and Hearing Research* 34 (1) 213-21 . Richie C (2004) *The effects of auditory-visual vowel and consonant training on speechreading performance* PhD dissertation, Indiana University, http://proquest.umi.com.libproxy.ucl.ac.uk/pqdweb?did=828422001&Fmt=2&clientId=54625&RQT=309&VName=PQD

[29] http://sportsillustrated.cnn.com/inside_game/magazine/life_of_reilly/news/2002/01/09/life_of_reilly/

http://www.timesonline.co.uk/tol/sport/football/international/article685555.ece

[30] http://www.lipreadingtranslation.com/livetranslation.htm ; http://www.amphl.org/articles/carroll2003.pdf

[31] Soto-Faraco, S., et al. (2007) Discriminating languages by speech-reading. *Perception and Psychophysics.* 69(2):218-31.

[32] Appeal court decision on the admissibility of lipread evidence Luttrell & Ors, R v [2004] EWCA Crim 1344 (28 May 2004)

[33] See *Experience of Expert Lipreaders – (1)Interim Report, (2) Final Report*" (Laraine Callow & Sally Reynolds), January and March 2010.

[34] Hill H, Bruce V (1996) Effects of lighting on the perception of facial surfaces *J.Experimental Psychology, Human Perception and Performance* 22 (4), 986-1004

[35] Erber N (1974) Effects of angle, distance and illumination on visual reception of speech by profoundly deaf children (1974)*J. Speech and Hearing Research* 17 (March),99-112

[36] Ijsseldijk F J (1992) Speechreading performance under different conditions of video image, repetition, and speech rate. *J. Speech and Hearing Research* 35: 466-471.

[37] McCormick B. 1979. A comparison between a two-dimensional and a three-dimensional lipreading test. *IRCS Medical Science* 7: 324

[38] *Experience of Expert Lipreaders – (1)Interim Report, (2) Final Report*" (Laraine Callow & Sally Reynolds), January and March 2010.

[39] McCotter MV, Jordan TR.(2003)The role of facial colour and luminance in visual and audiovisual speech perception. *Perception.* 32(8):921-36; Also Jordan TR, McCotter MV, Thomas SM.(2000)Visual and audiovisual speech perception with color and gray-scale facial images. *Perception and Psychophys.* 62(7):1394-404.

[40] Erber, N. P. 1971. Effects of distance on the visual reception of speech. *Journal of Speech and Hearing Research* 14: 848-857.

In this study Speechreading scores decreased by 0.8% per foot with increasing distance within the range 5-70 ft. Attributed to increased articulatory detail: participants could see within the mouth of the talker when they were closer to her (participants were aided by extra illumination at the mouth level of the talker).

[41] Jordan, T.R., Sergeant, P.C. (2000). Effects of distance on visual and audiovisual speech recognition. *Language and Speech*, 43*,* 107-124

[42] For general overview – see Munhall, K.G., & Vatikiotis-Bateson, E. (1998). The moving face during speech communication. In R. Campbell, B. Dodd & D. Burnham (Eds.), *Hearing by Eye II: The psychology of speechreading and audiovisual speech.* London: Taylor & Francis, Psychology Press, 123-139

[43] Lansing C R, McConkie GW (1994) A new method for speechreading research: Tracking observers' eye movements. *J. Academy of Rehabilitative Audiology*. 27,25-43; Lansing C R, McConkie G W, (2003) Word identification and eye fixation locations in visual and visual-plus-auditory presentations of spoken sentences *Perception & Psychophysics* 65, 536 -552; Vatikiotis-Bateson E *et al* (1998) Eye movements of perceivers during audiovisual speech perception *Perception & Psychophysics* 60 926 - 940

[44] ditto

[45] http://www.itu.int/ITU-T/worksem/qos/200606/presentations/s6p2-saks.pdf.

[46] Vitkovitch M. & Barber PJ (1996). Visible speech as a function of image quality: effects of display parameters on lip-reading ability. *Applied Cognitive Psychology, 10*, 121–140. Also see Munhall KG *et al* (2004) Spatial frequency requirements for audiovisual speech perception. *Perception & Psychophysics* 66: 574-583

[47] Thomas SM, Jordan TR (2004) Contributions of Oral and Extraoral Facial Movement to Visual and Audiovisual Speech Perception. *J. Experimental Psychology: Human Perception and Performance* 30, 873-888

[48] Rosenblum LD, Saldaña H (1997) Time-varying information for visual speech perception *I*n R.Campbell, B.Dodd & D.Burnham (Eds*) Hearing by Eye II,* Hove, Psychology Press, and see Jiang *et al* (2007) *op cit*

[49] Lesner SA (1988). The talker. *Volta Review* 90: 89-98.; Kricos PB, Lesner SA 1985. Effect of talker differences on the speechreading of hearing-impaired teenagers. *Volta Review* 87: 5-14; Bench, J. *et al* (1995) Choosing talkers for the BKB/a speechreading test: a procedure with observations on talker age and gender. *British J. Audiology* 29: 172-187.

[50] From Mohammed T (2007) *An Investigation of Speechreading in profoundly congenitally deaf British adults* PhD Thesis, University of London . This is a near-complete list and description of accessible and/or referenced tests of speechreading in English up to the date of writing. Further details are available from the thesis

[52] See *CrimeLine* Issue 117, Week ending 19 December 2004, also available as http://www.qebholliswhiteman.co.uk/downloads/article20File1.pdf

[53] http://www.dcal.ucl.ac.uk/course/Forensic_lipreading_course_2008.html