Guidance for organizations planning to use lipreading

for information gathering

Ruth Campbell Ph.D

Division of Psychology and Language Sciences

University College London

Spring 2010

These guidance notes are based on input from *Speechreading science forum* (London, November 25th 2009) and *Expert Speechreaders' forum* (London, December 8th, 2009), organised by Laraine Callow (Deafworks) and Ruth Campbell (UCL).

The evidence-base for the opinions expressed here can be found in an accompanying report, '*Speechreading for information gathering: a survey of scientific sources'* by Ruth Campbell (to be obtained from r.campbell@ucl.ac.uk).

Two further reports, *Experience of Expert Lipreaders – (1) Interim Report, (2) Final Report*, by Laraine Callow & Sally Reynolds (London, January and March 2010) summarise the experiences and recommendations of expert speechreaders in relation to using speechreading for information gathering. These may be obtained from Laraine Callow (lc@deafworks.org.uk). Material from those reports also informs these guidance notes.

These notes should not be distributed without permission of the author (r.campbell@ucl.ac.uk)

Background   There are large amounts of video data showing people talking, without any sound track. Could people who lipread in their daily life (mainly deaf and deafened people) help make sense of such 'silent speech'? These notes were developed in relation to collecting suitable material, and working with expert speechreaders.

Speechreading: what is it? Speechreading (lipreading is an older term, but the terms are used interchangeably here) is the skill of interpreting speech content from a silent or very noisy video-track. A good (silent) video record carries much information related to speech – actions of head (eyes, brows as well as mouth/lips), body (secretive or open posture etc).. and context. Speechreading can never be as reliable as a clear audio track, but that does not mean that it is useless. **However, because it is inherently unreliable, all and every means of *checking* any single interpretation should be taken.**

- Example: in a restaurant – one talker looks as if he might be saying '"tea please". That can be checked for accuracy if he then gets a cup of tea.
- Example: talker looks as if he's saying "!'d better go now.." and then leaves the room/shakes hand with other party etc.

These examples show that it is not enough to focus on a small segment of potentially important conversation: the speechreader needs as much information as possible from the video track in order to support her interpretation. This could include 'unimportant' speech, such as in the examples above. This will help firm up interpretation of more difficult and contentious parts of the video.

Videotapes and video clips: what's likely to be useful?

 If you are planning on obtaining a video record of speech, here are some points to watch out for.

- **Image capture.** Obviously, talkers' mouth actions should be visible for much of the conversation: full or ¾ views preferred, but profile is possible. Avoid shadow on the mouth region (top-lit and back-lit talkers will be relatively hard to speechread). Torso and whole body actions are informative. **SUDDEN CHANGES IN ZOOM SHOULD BE AVOIDED**: these make it difficult to follow a conversation. The speechreader needs to note reactions of people who are *not* talking (nods of the head, direction of eye-gaze) and other events beyond what the speaker is saying (as in examples above).
- **Length of video**. More is better. If there are records, including audio records, of a talker obtained from another source, they could be useful in (for instance) establishing the talker's accent and speech style**.**
- **Talkers' language and speech style**. Most casual conversation is conducted in 'non-standard English' (if in English at all). 'Street speech' can be particularly difficult to decode unless the speechreader is familiar with it. Information about *accent* (are the talkers known to be from a particular region?) and *language* (English or another language?) may be needed in order for the speechreader to attempt an interpretation. **FAMILIARITY (WITH A TALKER'S SPEECH) BREEDS UNDERSTANDING**

An expert lipreader CAN

- Advise whether a video record is 'lip readable'. Many video records are not useful for speechreading purposes. However, before ditching them, it can be worth asking an expert speechreader to check them out.
- *Sometimes* indicate whether a particular (target) phrase has been said – or not, and whether the speech is in a familiar language (i.e. in English or another language known to the lip reader).
- Prepare a transcript of what *may* have been said.
- Prepare a transcript showing a clear timeline in relation to speech.
- Indicate (in a transcript) where she is 'sure', 'less sure' and 'unsure' about speech content.
- Back up the transcript by demonstrating her interpretation of relevant parts of the video record to clients.
- Help in live surveillance: that is in viewing and interpreting on-line (live) talk, and advising on camera placement and zoom to capture talk.
- Be prepared to have her interpretation tested by other lipreaders, and to undertake testing of her own lipreading skill.

An expert lipreader CANNOT

- Give a *definitive* account of what was said; only a *possible interpretation.*
- Always identify the language a talker uses.
- Always interpret body language accurately.
- Avoid making use of (mis-)leading cues and clues. Speechreading makes maximum use of context. This is helpful when it is accurate, but blocks correct interpretation if it is not. **CLIENTS SHOUL D AVOID GIVING POSSIBLY UNRELIABLE INFORMATION TO THE SPEECHREADER**
- Be assumed to have knowledge of possible languages, accents and speech styles, unless s/he has had some training/familiarisation with them.
- Prepare a transcript immediately. It takes many passes through a videotape to prepare a reliable transcript: for three-to-four minutes of tape, about 10 hours of work may be required.

How good is your expert lipreader?

An expert lipreader is someone who uses lipreading a lot – often a deaf person, but it could be someone who works or lives with deaf people – and who has professional competencies for the task required. They should be aware of the different agency requirements, have good written language skills and access to appropriate technologies for viewing and transcript preparation as needed. The expert is likely to be aged between 20 and 60 years and should be able to show that s/he has undertaken professional tasks that call on his/her lipreading skills. These need not be restricted to information gathering: for instance, experts may use their skills to interpret for tracheotomy patients who cannot use their voice. The expert should be as familiar as possible with the language, and language *style* (accent, 'street-speech' etc.) used by the talkers.

Lipreaders differ in confidence: the most accurate lipreader is not necessarily the one who is surest. An excellent speechreader might give more 'unsure' responses than one who is less good, but is overconfident. Here are some suggestions to help achieve the *most likely* fit of seen speech to transcript.

- **If the expertise of the lipreader is in doubt, she should be willing to be tested. To test, make a short video clip of speech as similar in style to that required for interpretation. The audio track provides the 'correct' interpretation and the speechreader's transcript of the silent video should match that reasonably closely. If it does not, the speechreader should be able to indicate where and why her interpretation differs.**
- **Use more than one lipreader. They could work together or independently. Where they agree, the interpretation is more likely to be accurate.**
- **Be prepared to accept that *all* interpretations of lipspeech are provisional, and try to elicit as many possible interpretations as could fit.**