

# Lexical and audiovisual bases of perceptual adaptation in speech

Citation for published version (APA):

Ullas, S. (2020). *Lexical and audiovisual bases of perceptual adaptation in speech*. [Doctoral Thesis, Maastricht University]. Ipskamp Printing BV. <https://doi.org/10.26481/dis.20200617su>

**Document status and date:**

Published: 01/01/2020

**DOI:**

[10.26481/dis.20200617su](https://doi.org/10.26481/dis.20200617su)

**Document Version:**

Publisher's PDF, also known as Version of record

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# 6

## Knowledge valorization



## Valorization

Speech is essential for human interaction, but it is not accessible or experienced by everyone in the same manner. 6.1% of the world's population is estimated to have hearing loss (World Health Organization, 2020), and many rely on hearing devices or cochlear implants to be able to listen and communicate. However, these devices do not always operate optimally; they may amplify background noise or other sounds irrelevant to the listener, so users of these devices cannot solely rely on the now-amplified auditory signal in order to understand speech. Consequently, users of hearing devices, as well as others with hearing impairments who do not use such devices, may rely on information other than the acoustic signal itself to guide speech perception. Such populations may utilize lip-reading (also known as speech-reading) to support speech comprehension when the available acoustic signal is inadequate.

The studies presented in this dissertation have touched upon lip-reading, and specifically addressed various ways in which listeners can use contextual information to guide perceptual shifts of phonetic categories, particularly through knowledge of the lexicon and by attending to lip-reading cues. This line of inquiry has highlighted the importance of the non-acoustic contextual cues contained in speech, and how they can reshape what a listener hears and lead to shifts in internal representations of phoneme categories. The results of these studies hold implications for improving and refining educational strategies for lip-reading. Lip-reading can support speech comprehension, and while most listeners use lip-reading cues to some extent (and uniquely evidenced by the McGurk effect), for listeners with hearing impairments, lip-reading may supplement or even replace the auditory signal. Training in lip-reading involves conscious concentration on lip-movements being produced by the speaker in order to enhance recognition. Listeners thereby learn to build stronger links between singular and/or sequences of lip-movements with phonemes, syllables, and words. However, lip-movements alone may not convey enough information for the listener to interpret the speaker, as multiple phonemes map onto the same viseme (i.e. /pa/ and /ba/ are visually identical). Therefore, lexical knowledge also plays an important role in lip-

reading and can be an additional source of clarification. Training and educational strategies that incorporate both components may be more useful than either on their own, as each cue individually may be insufficient. Lexical knowledge and semantic context can be useful for the listener, so as to narrow the possible items of what the speaker is most likely to be saying, such as a word rather than a non-word (*bottle* versus *pottle*), the word most probable depending on the remainder of the sentence or phrase (baseball *bat* versus *pat*), or based on word frequency within a language (*pear* versus *bear*). Building strong links between visemes and sequences of lip-movements, along with their respective words may make lip-reading faster and more efficient. Lip-reading education already incorporates both lexical and audiovisual aspects, but potential advancements in lip-reading should place further emphasis on strengthening the mappings between phonemes, visemes, syllables and the lexicon. A multimodal approach to lip-reading and speech recognition featuring salient, non-acoustic contextual cues is more likely to benefit listeners struggling to comprehend speech, than strategies focused entirely on learning lip-movements and visemes themselves.

In conclusion, it is important to consider combining contextual cues when training listeners in lip-reading, as the combination of multiple contextual sources may be more useful to listeners who cannot rely on the auditory signal alone, and each source individually might not be a sufficient source of guidance. Investigating speech perception is not only essential for understanding a fundamental human experience, but is also necessary in order to make improvements upon technological devices designed for speech and communication purposes.

