



PDF of this article // 

print this article // 

The Science of Face-to-Face Communication: Q&A with Virginie van Wassenhove

By Cliff Atkinson

*If you make your face clearly visible when you speak, your audience will understand you faster, according to a recent **study** co-authored by **Virginie van Wassenhove**, PhD, while pursuing her doctorate in neuroscience and cognitive science at the University of Maryland. Surprisingly, when your audience sees your face clearly, they may actually be able to anticipate what you're going to say before you say it, speeding up communication.*

Cliff Atkinson: *Virginie, the terms "visual speech" and "auditory speech" are central to your research -- how do you define them?*

Virginie van Wassenhove: The term *speech* is often mistakenly used to refer to *auditory speech*. In fact, speech can be heard (auditory speech), seen (visual speech) or produced (speech production). *Auditory speech* thus corresponds to the acoustic signals received by the listener's ears. In face-to-face conversation, *visual speech* corresponds to the speaker's articulatory movements during production, and some of these articulatory movements are visible to the listener's eyes. You will also have heard the term *lip-reading*, which is a type of visual speech perception in which the lip-reader focuses on the speaker's lip movements to extract speech information. However, there is no particular focus on the lip movements when we speak about visual speech.

For instance, in face-to-face conversation, one naturally looks at the speaker's face. In auditory speech, the acoustic signals are often sufficiently intelligible to fully grasp the content of the utterance (for instance, over the phone). In visual speech, conformational changes of the vocal tract are hidden and unavailable (for instance when you pronounce 'daddy') while others can provide visible cues (for instance if you pronounce 'baby').

CA: *What is the difference between a visual stimulus such as the face of a presenter, and a visual stimulus such as a PowerPoint slide projected onto a screen?*

VvW: A major difference lies in the dynamics of the inputs. In visual speech perception (i.e. seeing the presenter's face), the information provided by the face is congruent in space and in time with the auditory inputs. Hence, we can now talk about *auditory-visual speech*, where one hears and sees the presenter's face. Additionally, the movements of the articulators provided by the presenter's face and the produced speech sounds that one sees and hears pertain to the same type of perceptual representations in the brain.

In a PowerPoint slide, none of the natural relations between auditory and visual speech are present. Rather, one may extract some information from what they hear independently from what they read. Reading and visual speech are two different types of perceptual strategies. Hence, while the informational content provided by the presenter may relate to the content of the PowerPoint slides, the underlying mechanism by which these two inputs are related in the brain may fundamentally differ.

CA: *Some presenters use PowerPoint to present only simple images, with no text on the screen. Do these visual slides qualify as "visual speech"?*

VvW: In general, static images will not qualify as visual speech (the static image of a face will be considered 'speech' in only a few instances). The addition of pertinent visual information (i.e. visual information that is 'semantically congruent' with what one hears) will most probably facilitate the assimilation of information. As was mentioned earlier, two sensory modalities are more efficient than one, but only so if their semantic values are congruent.

CA: *Based on your recent experiment, what impact does visual speech have on auditory speech?*

VvW: In the **study** David Poeppel, Ken Grant and I conducted, we found that visual speech information (i.e. seeing the speaker's face) facilitates the processing of auditory speech. One of the major results in the study suggests that the neural facilitation is specific to how much speech information one can extract from the interlocutor's face. The more salient the visual speech information is, the faster the auditory speech will be processed.

CA: *Let's say I make the same presentation twice - in the first one the audience cannot see my face, and in the second one they can. How much faster can the audience understand the second version where they can see my face?*

VvW: This is an interesting empirical question, although a difficult one to test and predict at this time. In our study, we limited our experimental conditions to syllables. During running speech (i.e. during the normal flow of a conversation), a realm of contextual information will also come into play. This information can potentially facilitate the speech system very early on (as we showed in our study, within a couple hundreds of milliseconds after hearing the sounds) but also later on at yet different levels in the hierarchy of speech processing.

CA: *What can presenters take away from your research findings? For example, how could a presenter speed up the processing of auditory speech?*

VvW: From a neurophysiological approach, much research needs to be pursued in order to specify what may enable more efficient processing of auditory speech. However, it is a definite advantage to have the presenter's face readily visible in front of an audience. For instance, we do know that multisensory (speech and non speech) information enables a faster reaction time (e.g. **Stein and Meredith**, 1993). For example, hearing and seeing a lion will enable you to react faster than if you were to hear it only.

From the speech point of view, seeing the presenter's face permits better intelligibility. This advantage may be more profound when a room is noisy or when the acoustics of the room are not optimal. Ken W. Grant has specifically addressed these questions for both normal-hearing and hearing-impaired populations **here**.

CA: *What do you think presenters might do to ensure they make maximum use of their faces to enhance visual speech? For example, should they make full use of facial expressions, make sure their face is clearly visible, or even use a IMAG (image magnification) camera to make sure their face is clearly visible to everyone in a large audience?*

VvW: I would answer yes to all of your suggestions. Face information does not only provide visual cues relevant to the comprehension of speech but also to prosody (i.e. the stresses and emphasis of speech such as a rising pitch at the end of a question). For instance, recent studies by **Beatrice de Gelder** and her colleagues at the University of Tilburg, Netherlands points out the importance of emotions conveyed by the face in auditory-visual speech processing.

CA: *Do your research findings apply only to live presentations, or do they apply to visual and auditory speech that is conveyed via technology channels such as television, film, and videoconferences?*

VvW: Our experiments made use of technology and the results were found based on digital videos. Hence, regardless of the medium by which visual information is made available, auditory speech can benefit from visual speech information. Our research is certainly relevant to any technologies that wish to incorporate a more naturalistic and efficient use of speech information. The timing of auditory-visual speech information needs to be respected as carefully as possible. For instance, if visual movements lag (instead of naturally preceding) the auditory signal by as little as 50 to 100 milliseconds, the benefit of having visual speech is already diminished.

*Cliff Atkinson is an independent management consultant who helps organizations solve problems related to PowerPoint. Visit his website at www.sociablemedia.com, read more articles and interviews **here**, or email him **here**.*