

HOW WHAT YOU SEE CAN INFLUENCE WHAT YOU HEAR

Paula C. Stacey* and Christian J. Sumner

NTU Psychology, Nottingham Trent University, Nottingham, United Kingdom

YOUNG REVIEWERS:



NOVA
AGE: 10



RONI
AGE: 13

We automatically use more than one of our senses to understand the world around us. When having a conversation, most of us use our ears to listen to what other people are saying. But you might be surprised to know that what you see is also very important in helping you to understand others. This article looks at how hearing and sight are combined and explains how both of these senses help us to understand speech. We demonstrate how important it is to see the faces of people who are talking, particularly when we are in noisy places. This topic has important implications for the wearing of face masks! It is really important that we think about how face masks may affect the ability of people to communicate with each other.

INTRODUCTION

Do you ever struggle to understand what your friends or teachers are saying when they are wearing face masks? Why is this? Nearly everyone benefits from seeing people's faces while they are talking. The various sounds people make when they talk require them to move

their mouths in unique ways, and even though you might not be aware of it, your brain understands those specific mouth shapes! The shapes make it easier for us to understand people when we can see them talking.

The brain has specialized areas that process information from the senses and allow us to understand the world. Some of these brain areas automatically combine information across multiple senses. When information from one sense (hearing, for example) is unreliable, then we automatically rely more on information from another sense (like vision). This is why you will particularly notice the benefits of seeing people speak when you are in a noisy place.

In this article, we will begin by introducing you to the difference between sensation and perception, and we will use a well-known visual illusion to demonstrate how perception can sometimes be inaccurate. This will help you to understand that what you perceive does not necessarily correspond to the information that is in the world. We will then explain how our brains combine information across different senses, particularly sight and hearing.

MAKING SENSE OF THE WORLD

Your ears pick up sound waves and you hear them as sound. Your eyes pick up light energy, which gives you sight. We call these processes **sensation**. What your brain then does with the information from your senses is called **perception**. Most of the time, we assume that what we sense and what we perceive are the same thing. But sometimes our brains come to the wrong conclusions about what is out there! Perception can be a tricky job: it involves combining what our senses are telling us with what we already know about the world, to make the best possible guess about what is happening. We can misunderstand, or misinterpret, sensations, and then perceive something that just is not true.

Visual illusions can be fascinating to look at, and they give us a really good example of how sometimes the brain can misinterpret sensations to give us the wrong perception. Have a look at Figure 1. Which car do you think is the biggest? Now take out a ruler and measure them. Were you right?

At a first glance, you might think the top car is larger than the bottom one, but they are actually the same size! We call this a visual illusion because the picture fools your brain into thinking the cars are different sizes. This works because clues in the picture (the shape of the road, in this example) make the top car appear to be further away than the bottom car. Normally, things that are far away look smaller than things close to us (next time you are on a car journey look at the size of the cars in the distance compared to those close by), so the car at the top

SENSATION

The process where our senses pick up information in the world.

PERCEPTION

The process where our brains make sense of information provided by the senses.

Figure 1

Visual illusion. These cars are the same size, yet the top car is usually perceived to be larger. This is because the cues in the picture tell us the top car is further away, so the brain automatically “scales up” the image. This example shows us that what our senses observe and what our brains perceive are not always the same.

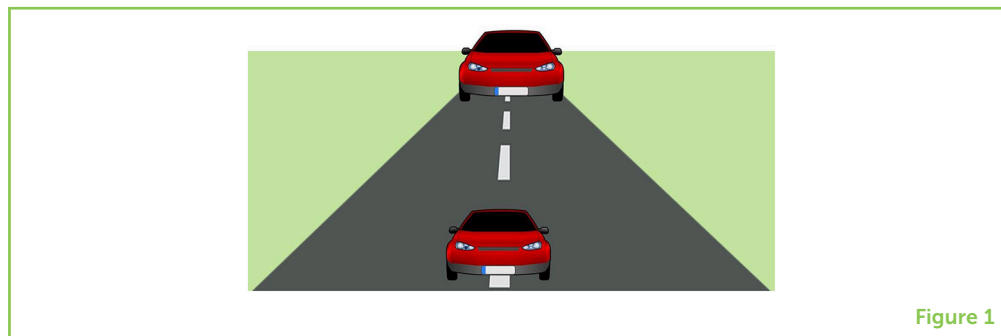


Figure 1

of the road *should* appear smaller because it is further away. The brain somehow knows this, so it “scales up” the perception of the car’s size automatically. In real life, this mental adjustment usually works well and helps our perception of the world to make sense.

COMBINING OUR SENSES

We know that our brains often use both vision and hearing to understand the world. This can make it easier for our brains to correctly determine what is going on than if we only used one sense. But the same way a visual illusion can play tricks on the brain *via* the eyes, an audio-visual illusion can confuse perception if the information coming from hearing and vision do not match. This was illustrated by an experiment performed by Harry McGurk and John McDonald in 1976 [1]. These scientists recorded a person saying the sound “ba.” They then made a silent video of the same person saying the sound “ga.” They played the sound recording (“ba”) alongside the video (“ga”) and asked volunteers to say what sound they heard. You might be surprised to learn that most volunteers said they heard the sound “da”! This is because their brains mashed up what they saw and heard and made a guess at what the sound really was. The amazing thing about this experiment is that it really does change the way the words sound, which is called the **McGurk effect** (Figure 2). The experiment showed that seeing a talking face can have a powerful effect on the sounds people perceive¹. The effect is also found with other combinations of sounds. The perception of speech is really the result of the brain making the best guess it can from what we see, what we hear, and what we already know about language.

HOW SIGHT HELPS US HEAR IN NOISY PLACES

In the real world, lip movements usually match the sounds coming from people’s mouths, so seeing someone talk usually helps us to perceive the right word. This visual information tends to be more helpful for understanding speech when the conversation takes place in a noisy setting. Try sitting in a quiet room with a friend. You will probably find it quite easy to understand what your friend is saying.

MCGURK EFFECT

An audio-visual illusion that occurs when hearing one sound (e.g., “ba”) while also seeing the mouth movements that accompany another sound (e.g., “ga”) results in perception of a new sound (e.g., “da”).

¹ Have a look at this video for a demonstration of the effect:
<https://www.youtube.com/watch?v=jtsfidRq2tw&t=20s>.

Figure 2

The McGurk effect. In this illusion, hearing one sound (e.g., “ba”) while also seeing the mouth movements that normally accompany another sound (e.g., “ga”), often results in listeners perceiving a new sound (e.g., “da”). This tells us that the brain automatically combines information across the senses, to help us understand the world.

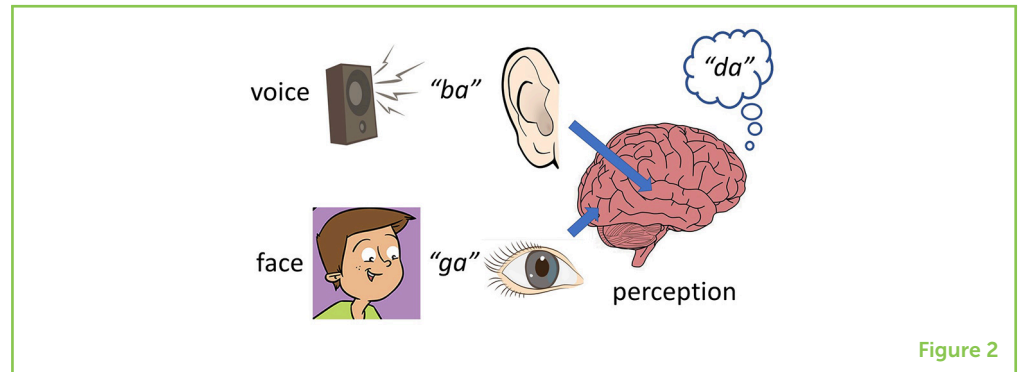


Figure 2

Now turn the television on with the volume up loud. How easy is it to understand your friend now? It is much harder to hear and understand what someone is saying if the background is very noisy. If you look at your friend’s face and lips, you are more likely to correctly guess what is being said.

MEASURING HOW VISION AFFECTS UNDERSTANDING OF SPEECH

We carried out some experiments on how background noise affects understanding of a sentence [2]. First, we videoed and recorded a volunteer saying several different sentences (such as “A large size in shoes is hard to sell.”). Then we played these recordings to other volunteers and asked them to repeat what they heard, and we counted how many words they got right. We repeated this experiment with lots of volunteers, adding various levels of background noise, or reducing the volume of the speech signal. The relationship between the speech signal and the background noise is called the **signal-to-noise ratio**. Sometimes we only played the sound of the person speaking, and other times we also showed the video of the person, so volunteers could both see and hear the speaking person.

We plotted the information we gathered on a graph (Figure 3). When the background noise was much louder than the person speaking, our volunteers could not identify any words correctly. But when the sentences were louder than the background noise, our volunteers heard most of the words correctly. Between these two situations, volunteers could identify some, but not all of the words correctly. Our results clearly show that people can understand speech with more background noise present if they can see the faces of talkers. This effect explains your experience of seeming to hear better when you can see people speaking, especially when you are in a noisy place. Previous studies have found similar results [3].

Why is this important? Imagine you are working in a noisy factory with loud, dangerous machinery. Now imagine your co-worker is trying to explain exactly how to use a machine. With all that noise in the

SIGNAL-TO-NOISE RATIO

A measure of the strength of the “signal” (in our case, speech) against the background noise.

Figure 3

When the background noise was much louder than the person speaking (low signal-to-noise ratio), volunteers could not identify any words correctly (A). But when sentences were louder than background noise, volunteers heard most of the words correctly (C). In between, volunteers could identify some, but not all, words correctly (B). When volunteers could both see and hear the person talking (red line), they got more words correct than when they could only hear the talker (blue line). So, it is easier to understand speech when we can both see and hear the speaker.

LIP READING

Understanding speech from mouth movements alone.

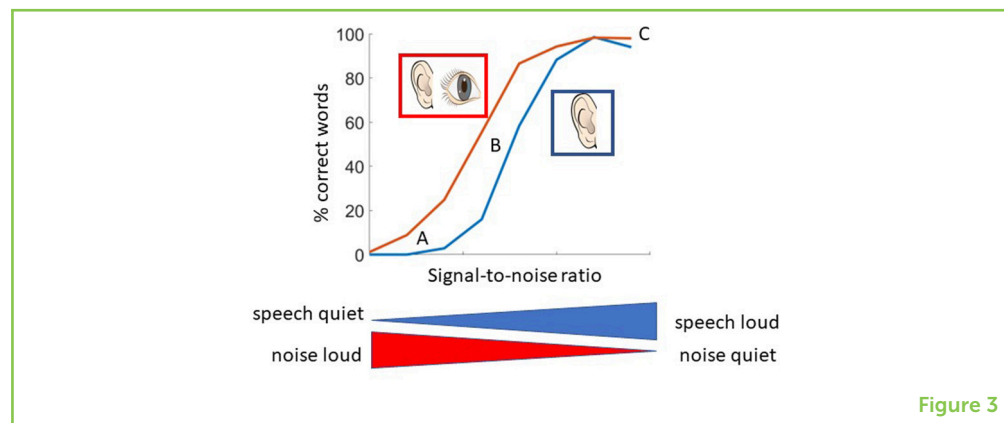


Figure 3

background, do you think you would be able to hear the instructions clearly? What if your co-worker was wearing a safety mask while talking? In this sort of environment, making a mistake could be very dangerous and might even result in injury.

BENEFITS OF VISUAL INFORMATION FOR PEOPLE WITH HEARING LOSS

Have you ever tried to understand what someone is saying by *just* looking at their face, without hearing their words? You may find you are better at this task if you have poorer hearing. People who are deaf or hearing impaired are generally better at **lip reading** than people with normal hearing [4]. This skill probably developed over time as these people lost their hearing and had to pay closer attention to their other senses.

During the coronavirus outbreak, wearing face masks was often a requirement. Masks make lip reading very difficult for people who need to do so. In the UK, campaigners helped persuade the government to provide see-through face masks to help deaf people to communicate. In our research, we try to understand how lip reading helps communication, and especially how it helps people with hearing problems.

CONCLUSION

Visual information is really important to help us understand the sounds we hear. What we see is especially important for understanding speech when the sound we are listening to is affected by either background noise or poor hearing. There is still a lot we do not understand about how the brain combines what we see and what we hear. Next time you are in a noisy environment, think about how much you rely on seeing the faces of the people who are speaking, and about what could be

done to make visual speech information more accessible for people who have trouble hearing.

REFERENCES

1. McGurk, H., and MacDonald, J. 1976. Hearing lips and seeing voices. *Nature* 264:746–8.
2. Stacey, P. C., Kitterick, P. T., Morris, S. D., and Sumner, C. J. 2016. The contribution of visual information to the perception of speech in noise with and without informative temporal fine structure. *Hear. Res.* 336:17–28. doi: 10.1016/j.heares.2016.04.002
3. Sumbly, W. H., and Pollack, I. 1954. Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26:212–5.
4. Schreitmüller, S., Frenken, M., Bentz, L., Ortman, M., Walger, M., and Meister, H. 2018. Validating a method to assess lipreading, audiovisual gain, and integration during speech reception with cochlear-implanted and normal-hearing subjects using a talking head. *Ear Hear.* 39:503–16. doi: 10.1097/AUD.0000000000000502

SUBMITTED: 06 May 2021; **ACCEPTED:** 04 August 2022;

PUBLISHED ONLINE: 22 August 2022.

EDITOR: Guilherme Lucas, University of São Paulo, Brazil

SCIENCE MENTORS: Susana Martinez-Conde and Shlomit Yuval-Greenberg

CITATION: Stacey PC and Sumner CJ (2022) How What You See Can Influence What You Hear. *Front. Young Minds* 10:705956. doi: 10.3389/frym.2022.705956

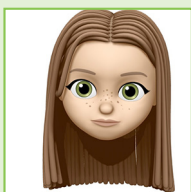
CONFLICT OF INTEREST: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

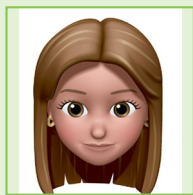
COPYRIGHT © 2022 Stacey and Sumner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

YOUNG REVIEWERS

NOVA, AGE: 10

My name is Nova and I am in 5th grade. My favorite subjects are writing, science, social studies and reading. When I grow up, I would like to be an architect, because I like art and I also like building. I think it is important for kids to be curious so they can learn. Albert Einstein said that he did not have a special brain, but he wondered how the universe worked, so he went out and learned, so he could figure it out.



**RONI, AGE: 13**

Roni was born in Jerusalem but lived most of her life in Tel Aviv. She is a 7th grade student in Pelech High School, Tel Aviv. She lived in the US for three years as a baby and toddler until moving back to Jerusalem and then to Tel Aviv. She studied elementary school in the Meshutaf school, Tel Aviv, except 5th grade which she spent in Berkeley CA. Roni is interested in science but also loves gymnastics and singing.

AUTHORS**PAULA C. STACEY**

Paula C. Stacey is a principal lecturer in psychology at the Nottingham Trent University. Paula completed her Ph.D. on the topic of auditory training for adults who use hearing-assistance devices called cochlear implants. She is currently interested in how people use visual information to help them to understand speech in noise. *paula.stacey@ntu.ac.uk

**CHRISTIAN J. SUMNER**

Christian J. Sumner is an associate professor in auditory neuroscience at the Nottingham Trent University. He has a background in computer science, and so views the brain as a big computer—and wants to understand the program for how we hear! He has wide interests in the field of hearing, from how the brain separates out simultaneous sounds (like instruments in an orchestra), to how single neurons in the brain process sound, to how hearing changes as we age.