Project Abstract

Temporal Dynamics of Phonological Expectations in Language Comprehension and Development

0433567
Delphine Dahan and Daniel Swingley
University of Pennsylvania

Overall objective and preliminary outcomes

Imagine a situation where your task is to evaluate some information in order to guide a decision, but the information is ambiguous, noisy, and keeps changing. How could you ever perform this task? Amazingly, this situation captures the challenge we face when we perceive and interpret what someone is saying. Speech is a continuous, highly complex, and rapidly changing signal. A spoken word is a temporal sequence, which becomes available only gradually. Contrary to written language, where the stimulus usually remains in view, speech is transient. Because humans' sensory memory is limited, speech must be evaluated and interpreted incrementally. Speech is also a noisy and ambiguous signal. For example, the same acoustic detail in the signal can mark the presence of different speech sounds (e.g., *g* or *d*), depending on the nature of the sound that precedes or follows it (e.g., *l* or *r*). As a result, the recognition of spoken words requires the on-line management of a continually shifting set of expectations, to work out an interpretation in the face of substantial ambiguity.

The objective of this project is to understand the nature of the expectations that adult and infant listeners bring to the task of interpreting speech, and how these expectations change to reflect listeners' continuous adaptation to the information they have accumulated. In order to assess listeners' interpretations, we monitor their eye movements to potential referents of the spoken stimulus. Adults or young children hear spoken instructions referring to one of a set of objects displayed on a computer screen (e.g., "*Look at the cat!*"). Listeners' eye movements to an object upon hearing a spoken name are taken to reflect the degree to which the object is thought to be the named target. Thus, eye movements serve as an index of listeners' interpretation of the target's name. Previous research from our labs and others has established that eye movements provide a useful measure of speech interpretation over time. For instance, if adult listeners are asked to "*click on the beaker*" with the computer mouse, by the end of the word *beaker*, they are more likely to have launched an eye movement to the picture of a beaker (the target) or a beetle (which sounds like *beaker* at first) than to the picture of, for example, a baby carriage. This demonstrates that (1) people process speech incrementally and generate hypotheses about the identity of the spoken word they hear before they have received complete information, and (2) eye movements to visual potential referents can reflect this incremental process.

We examine how evidence from the acoustic signal, combined with prior expectations, is evaluated and integrated over time by both adults and very young children. Evidence that children as young as one year of age continuously monitor the speech signal and can revise an interpretation based on new information is scarce. Our first study aimed to establish continuous uptake and evaluation of phonetic information in the signal by one-year olds. Children between

14 and 20 months of age and their caregivers were invited to the lab. During testing, the child was sitting on his/her caregiver's lap, facing a large monitor. On each trial, two familiar objects were displayed, and one was named using a correct pronunciation (e.g., *Where's the dog?*) or mispronunciation (e.g., *Where's the tog/dawb?*). A concealed video camera recorded the child's face, and infants' eye movements were coded as being directed to one or the other of the two displayed pictures. Each child heard correct pronunciations and onset or offset mispronunciations of known names of objects (as assessed by parental reports). The mispronunciations were natural recordings, created by changing one phonetic feature to the first or last consonant of the original name. Overall, children looked at the referent less when its name was mispronounced than when it was correctly pronounced. Importantly, the timing of this difference reflects where in the word the mispronunciation occurred. This result demonstrates that even very young children continuously evaluate the speech signal with respect to a mental representation of the referent's name, and that eye movements can capture this dynamic process. This opens the way to testing the perception of artificially modified pronunciations that are even more subtle, for which different theories make different predictions.

A second line of research focuses on adults' ability to adapt to the characteristics of a talker, and the processing consequences of this adaptation. In some American English dialects, the vowel *a* before *g* (as in *bag*) is raised to a vowel approaching *e* (making *bag* sound almost like *beg*). Importantly, this "raising" phenomenon does not occur before a *k* context (so *back* does not sound like *beck*). This phenomenon in effect reduces the phonetic overlap that exists in standard American English between (e.g.) *bag* and *back*, because with raising of *bag*, *bag* and *back* no longer have the same vowel. We took advantage of this dialectal phenomenon to study listeners' adaptation to speaker differences. In this study, participants saw four written words on a computer screen (e.g., *bag, back, dog, dock*) and heard a spoken word. Their task was to indicate which word they heard. Participants' eye movements to the written words were recorded. First, participants in the "raising" group heard *bag*-like words containing the raised vowel *e*; participants in the "control" group heard *bag*-like words containing standard *a*. Then, acoustically identical *back*-like words were presented to both groups. Participants in the "raising" group made fewer fixations to the competitor *bag* than control-group participants did. This result indicates that (1) prior exposure to raised realizations of *bag* has changed the pronunciation of *bag* that listeners expect from this talker, and (2) this knowledge is involved when evaluating the initial portion of the spoken word *back*: The evaluation of the vowel *a* in *back* predicts *bag* to a lesser degree for the raising-group participants than for the control-group participants, even though both groups are hearing the same "standard" pronunciation of *back*

Adult listeners evaluate speech with respect to expectations. This is often modeled as Bayesian inference. This study's main contribution is showing that the computation of "likelihood" (the probability of the signal, given a hypothesis) is context- and talker-dependent. Incoming data from the speech signal are used for two purposes simultaneously—as input to be interpreted, and as input to processes that modify the perceiver so as to optimize future interpretation.

Broader impacts

The project promotes teaching and training by involving undergraduate students in the laboratory. Under the direction of the PI and co-PI, several undergraduate students have participated in all phases of the research effort, from assisting in the design and creation of experimental stimulus materials, to testing participants and analyzing results. Some of these students have participated as part of Research Experience courses that the PIs offer annually.

To date, all of the undergraduates who have participated in the research this way have been women, and a high proportion have been from underrepresented groups. Given that nearly all of the students who have worked in our labs have then gone on to seek higher education or employment in research careers, the present project may be viewed as a springboard for young researchers with promising futures in science.

Because the research participants in the infant studies can be anyone in the community who is raising a native speaker of English, the studies have an impact outside the University and, most likely, outside the range of much exposure to science. Minority group participation in these studies has been encouraging. All parents who participate are then placed on our mailing list for the infant lab newsletter, which presents our research and other scientific research studies of early development in understandable terms. In this way, parents can place their own participation in the broader context of incremental scientific progress.