ICA 2013 Montreal
Montreal, Canada
2 - 7 June 2013

Animal Bioacoustics
Session 2pAB: Listening in the Natural Environment

2pAB6. Influences of perceptual continuity on everyday listening
Barbara Shinn-Cunningham*, Golbarg Mehrai, Scott Bressler and Salwa Masud

*Corresponding author's address: Center for Computational Neuroscience and Neural Technology, Boston University, 677 Beacon St., Boston, MA 02215, shinn@bu.edu

In the natural environment, listeners face the challenge of parsing the sound mixture reaching their ears into individual sources, and maintaining attention on a source of interest through time long enough to extract meaning. A number of studies have shown that continuity of certain acoustic features (including pitch, location, timbre, etc.) allows the brain to group sound from one acoustic sound source together through time to form an auditory object or stream. This presentation reviews results demonstrating that auditory feature continuity has important consequences on how listeners maintain attention on a stream through time. For instance, continuity of a sound feature that a listener knows is irrelevant to the task at hand nonetheless impacts the ability to maintain auditory attention based on some other sound feature. Moreover, the influence of auditory feature continuity decreases as the time between events in a given sound stream increases. Taken together, these behavioral results support the idea that auditory attention operates on auditory objects, rather than on individual sound features, and that feature continuity has an obligatory influence on the formation of auditory streams, and therefore on how selective auditory attention allows us to communicate in everyday settings.

Published by the Acoustical Society of America through the American Institute of Physics
PERCEPTUAL CONTINUITY

Sound conveys information through the changes in its content through time. Many important sound sources, like the voice of a colleague or the notes of an oboe, can create sound that has a relatively long duration, over the order of seconds (or longer). Most individual speech syllables or musical notes (or comparable spectro-temporal units, all of which will henceforth be called “syllables” for simplicity) typically carve out a continuous contour in time-frequency, an attribute that allows the listener to easily group together the syllable, perceptually, based on this local, continuous spectro-temporal structure. Importantly, there can be silent gaps or pauses between syllables, discontinuities in the spectro-temporal structure from one syllable to the next. Yet, despite this, listeners still hear the ongoing stream of syllables as one unified, coherent stream. In a typical perceptual stream, syllables are not identical, yet the listener has no trouble hearing these different items as coming from one perceptual source. This perceptual continuity arises because although syllables differ from one another, they typically also are similar to each other in one or more perceptual dimensions, or perceptual features (e.g., see Ihlefeld and Shinn-Cunningham, 2008; Maddox and Shinn-Cunningham, 2012). These features include, but are not limited to:
- location,
- pitch,
- voice quality,
- attack/decay profile, and
- other aspects of timbre.

It is the similarity of perceptual features like those listed above that allows a listener to link together syllables through time, forming an ongoing perceptual stream out of a sound mixture.

PERCEPTUAL CONTINUITY AIDS SELECTIVE AUDITORY ATTENTION

Object formation (including the perceptual continuity of objects through time) is important for listening in real-world settings. In most settings, competing sounds vie for attention; listeners are able to focus selective auditory attention on a source that is perceptually distinct from other competing sources. Selective attention requires both that a particular source be perceptually segregated from other sources, and that the desired source has some known feature that distinguishes it from other sources, allowing the listener to select it as the focus of attention (for discussion, see Shinn-Cunningham, 2008). Perceptual continuity of ongoing sources is important for enabling a listener to maintain focus on an ongoing sound source in order to extract its meaning, both by enhancing sound source segregation and by providing some feature or features that a listener can use to focus attention (e.g., listen to the source on the right, or the high-pitched source, or the male talker).

A previous series of experiments from our lab demonstrated these concepts. We showed that when listeners are asked to report a stream of target digits (indicated by a spatial cue) presented with simultaneous, masker digits, continuity of the digit location and the voice of the talker improved performance (Best et al., 2008). We then found that these effects are not due to having advance knowledge of the target location(s), but comes about from the location of the target not changing through time (Best et al., 2010). Moreover, we found that even when a particular feature is not relevant for performing a task, its continuity influences performance (Ihlefeld and Shinn-Cunningham, 2008; Maddox and Shinn-Cunningham, 2012).

EXPERIMENT 1: NON-SPATIAL PERCEPTUAL CONTINUITY

These initial studies focused solely on the perceptual feature of location, or pitted location against other attributes (pitch, voice). Since perceptual continuity of other features also affects performance, we decided to focus on sound mixtures where there are no spatial differences in the competing sources to see whether non-spatial feature continuity alone affects selective auditory attention performance. To test these questions, in Experiment 1 we created mixtures of target digits embedded in time-reversed digits. Listeners were presented with diotic mixtures of simultaneous target and masker digits, and asked to report the forward, target digits. In some blocks, the target talker was the same throughout a given trial. In other blocks, the identity of the target talker changed (randomly) from digit to digit. We found that voice continuity improved selective auditory attention performance (see Figure 1). These results show that perceptual continuity of non-spatial features acts in much the same way as previously demonstrated effects of spatial continuity.
FIGURE 1. Results of Experiment 1. The solid line represents performance as a function of digit position for fixed voice trials, where within trial, the target talker remained constant. The dotted line represents performance for random-voice trials, where the identity of the target talker always changed from one digit to the next. Overall, for each digit position performance during fixed voice trials is better than performance during random voice trials. For a 0ms inter-digit delay (left plot) performance was worse than for a 500ms delay (right plot).

EXPERIMENT 2: VOLITIONAL VS. AUTOMATIC EFFECTS OF CONTINUITY

Because of the way the trials were blocked in Experiment 1, it was not clear whether perceptual continuity improved performance because listeners knew, within a block, that once they heard a target digit, they could simply direct attention to the target talker, volitionally. Therefore, in Experiment 2, we manipulated the talker transitions within each trial randomly, without informing the subjects. Each couplet of digits within the target string (1-2, 2-3, 3-4, etc.) could either have the same target talker, or different talkers. The repeated talker transitions were randomly occurring, unpredictable, and rare, so even if listeners figured out that repeated-talk transitions occurred, it would hurt performance to try to listen for the same talker in a subsequent digit. We found that talker continuity improved performance in this case (see Figure 2). This result shows that the effects of perceptual continuity on auditory attention are involuntary and automatic.

FIGURE 2. Results of Experiment 2. The solid line represents performance as a function of digit position for fixed repeated voice transitions, where the preceding target digit and the current digit had the same voice. The dotted line represents performance for novel voice transitions. Performance was always better when the target talker was the same in the preceding digit, even though these transitions were not predictable and not relevant to the overall performance of the task.
EXPERIMENT 3: EFFECT OF TEMPORAL SEPARATION

In a third experiment, we explored how the timing between digits affects perceptual continuity. In this experiment, listeners were instructed to pay attention either to location or to talker gender. The irrelevant feature either was continuous or discontinuous, with equal likelihood. We find, as in previous studies (e.g., Maddox and Shinn-Cunningham, 2012), that continuity of a task-irrelevant feature aids performance. However, we also find that as the gap between syllables increases, the effects of feature continuity decrease. This result demonstrates that feature continuity depends on the spectro-temporal separation of syllables in a stream. The longer the separation between syllables, the less strong are the influences of continuity.

![Figure 3: Average performance for continuous (fixed) and discontinuous (switch) irrelevant feature conditions. X-axis depicts the feature to which listeners directed attention. Each panel indicates the inter-word delay (IWD) between digits within a trial. Notice that the influence of continuity weakens as IWD increases (left to right).](image)

SUMMARY AND FUTURE STEPS

These experiments demonstrate that once a listener is focused on a particular sound source, whatever subsequent sound is perceptually continuous (part of the same, ongoing perceptual stream) is preferentially likely to remain in the focus of attention. At least some portion of this effect comes strictly from the continuity of perceptual features in the stimulus, not from top-down, volitional control of attention. We are now exploring how perceptual continuity affects the auditory attentional network in the brain, using magneto-encephalography (MEG) to see whether discontinuity of perceptual features causes different patterns of neural activity, revealing a cortical correlate of these perceptual results.

ACKNOWLEDGMENTS

This work was supported by funding from the National Institute for Deafness and Communication Disorders, the Office of Naval Research, and a National Science and Engineering Faculty Fellowship.

REFERENCES


