

From "twos" to "tunes" in 15 easy steps: Cue trading in speech perception.

The "listener as a source of sound change" model (Ohala, 1981) of sound change, and ultimately linguistic sound patterns, assumes that listeners "parse" the acoustic signal into linguistic units and that some sound changes occur because of misparsing. One key aspect of this model was explored in the exercise "cutting up damp skunk". There we found that the acoustic cues that lead to the perception of particular features (place of articulation, voicing) or sounds (presence or absence of a nasal segment) are distributed over non-contiguous time, and vary according to their local context and are thus not invariant acoustic images or patterns of articulator movement.

In this exercise we explore the phenomenon of "trading relations" among acoustic cues in speech perception and particularly in the perception of a nasal segment. We trade two cues for post-vocalic /n/ against each other; distinguishing the English words "twos" and "tunes". One cue for the presence of /n/ is the duration of the nasal [n] segment itself. If the nasal segment is very short then the word probably sounds more like "twos" than like "tunes". The other cue is the coarticulated nasality on the vowel. If the vowel is nasalized then the word is more likely to be "tunes" than "twos". Cue trading is one indication that listeners are parsing the acoustic signal, imposing coherence on speech by taking in information from a variety of signal properties.

A. Collect some perception data.

We'll start the exercise by conducting a short experiment with you as the listener. Here are the steps to follow:

- 1) download and unzip the file "tunes_twos.zip". When unzipped this should create a new directory/folder on your computer called "toons_twos". This directory contains 32 sound files and a file called "identify.txt". We'll use these files to run a speech perception experiment.
- 2) Use headphones (ear buds are OK) and find a quiet place to do the experiment.
- 3) Start Praat, and open the experiment using the "read from file" command, just as you would open a sound file. Run the experiment by choosing the object "ExperimentMFC identify" and clicking the "run" button.
- 4) When you have finished, click on "extract results". Highlight the object "ResultsMFC identify" and click on "Collect to Table". Highlight "Table allResults" and from the "write" menu select "write to Table file" and email the resulting text file to kjohnson@berkeley.edu so I can compile results for the whole class.
- 5) Select "Table allResults" and click on "Edit" to see your results. Enter the number of "twos" responses to each stimulus in the following table (for example, the first cell in the table is for the number of "twos" responses to the stimulus token "twos_1"):

| | | | | | | | | | | | | | | | | |
|---------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|
| %"twos" | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| twos | | | | | | | | | | | | | | | | |
| toons | | | | | | | | | | | | | | | | |

B. Make some acoustic measurements of the stimuli.

Measure the durations (in ms) of the vowels and of the nasal segments in each of the 32 stimuli and enter your measurements into the tables below. For instance, the value that you put in the first cell of the "vowel" table is the duration of the vowel in token "twos_1.wav".

| | | | | | | | | | | | | | | | | |
|-------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|
| vowel | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| twos | | | | | | | | | | | | | | | | |
| toons | | | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | | | | |
|-------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|----|
| nasal | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| twos | | | | | | | | | | | | | | | | |
| toons | | | | | | | | | | | | | | | | |

C. Answer some questions.

Question 1. The boundary along a continuum of edited speech tokens like these (more on that later) is the point where labeling shifts from one category to the other. For example when I did this experiment I labeled token "twos_10" as "tunes" both times I heard it, token "twos_11" as "twos" once and "tunes" once, and both times I heard token "twos_12" I called it "twos". So my boundary (the point where my labeling function crossed 50% "twos") was at #11 on the "twos" continuum. What were your boundaries for the "twos" and "toons" continua? Did you have the same boundary for them? If the boundaries are different what explanation for the difference would you give?

Question 2. The files "toons_1.wav" and "twos_16.wav" are the tokens that are closest to the natural productions of "tunes" and "twos" (toons_1.wav actually has the final [z] of "twos"). How do the vowel and [n] durations vary with each other in these natural tokens? In the discussion above I said that we would be using the duration of the [n] segment as one "cue" for /n/, but here you find that I manipulated the vowel duration as well. What are the cues that are being manipulated in these speech tokens?

Question 3. Listen to the four endpoints (twos_1.wav, twos_16.wav, toons_1.wav and toons_16.wav). How does the vowel in the "twos" tokens differ from the vowel in the "toons" tokens? Which of these four tokens sounds the least like a natural speech pronunciation? i.e. which sounds most obviously edited? Why does this one sound least natural? How might your native language influence your choice?

Question 4. If we consider the action of the tongue and velum we can imagine a sequence of events in a

production of the word "tunes" (toons_1.wav).

| | | | | |
|----------------------|--------|------|--------|----|
| | [t | u | n | z] |
| tongue tip: | up | down | up | |
| velopharyngeal port: | closed | open | closed | |

What sequence of events is implied by the acoustics of tokens toons_16.wav, twos_1.wav, and twos_16.wav? Produce a diagram like the one we provide for toons_1.wav for each of these others.

Discuss the articulatory variation and resulting misperceptions that might take place for tokens twos_1.wav and toons_16.wav - and the sound changes that might take place as a result.

D. Try making your own experiment.

If you are feeling ambitious or geeky you might try making your own nasal segment cue trading experiment. The file "basefiles.zip" contains sound files for the words "tease" and "teens", and for "toes" and "tones". Questions that you can address in this optional section: (1) You might want to try manipulating nasal duration while holding vowel duration constant to see if the inverse vowel and nasal duration that were used in the "twos" - "tunes" continuum was really necessary. (2) Does it matter whether you use the [s] of "tease" or "teens" (or of "toes" or "tones")? (3) Do vowel nasality and nasal segment duration trade off with each other for the vowels [i] and [o]?