

**6.542J, 24.966J, HST.712J LABORATORY ON THE PHYSIOLOGY,
ACOUSTICS, AND PERCEPTION OF SPEECH
Fall 2005**

Kenneth N. Stevens

Joseph S. Perkell

Stefanie Shattuck-Hufnagel

Lab 13

10/20/05

Speech intelligibility and confusion matrices

References

G.A. Miller and P.E. Nicely (1955) *Analysis of perceptual confusions among some English consonants*, **J. Acoust. Soc. Am.** **27**, 338-352.

D.N. Kalikow, K.N. Stevens and L.L. Elliott (1977) *Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability*, **J. Acoust. Soc. Am.** **61**, 1337-1351.

In this laboratory we will examine two methods of assessing the intelligibility of speech passed through a communication channel. One method examines the intelligibility of phonetic segments and features in noise, and the other examines the effect of context on the intelligibility of words in sentences.

1. You will hear a series of consonant-vowel (CV) nonsense syllables in which the vowel is always / α /, and the consonants are selected from the list /b d g p t k f s š v z ž m n w y/. The speech is mixed with various amounts of white noise. The syllables are presented in random order (with an extra pause after every 10 items), and there are several 90-item tests. After you hear each item, you must write down which of the 16 consonants you heard. You must *always* make a response of one of these consonants, even if you have to guess.

Each group will grade their results and organize them into confusion matrices. Is there a pattern to the errors? Note that the consonants can be organized into features as follows:

| | b | d | g | p | t | k | f | s | š | v | z | ž | m | n | w | y |
|------------|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| continuant | - | - | - | - | - | - | + | + | + | + | + | + | - | - | + | + |
| sonorant | - | - | - | - | - | - | - | - | - | | - | - | + | + | + | + |
| voiced | + | + | + | - | - | - | - | - | - | + | + | + | + | + | + | + |
| labial | + | - | - | + | - | - | + | - | - | + | - | - | + | - | + | - |
| coronal | - | + | - | - | + | - | - | + | + | - | + | + | - | + | - | + |
| palatal | - | - | - | - | - | - | - | - | + | - | - | + | - | - | - | + |
| velar | - | - | + | - | - | + | - | - | - | - | - | - | - | - | - | - |

What can you say about the intelligibility of the individual features?

2. You will hear two lists of sentences in noise, and your task is to write down the last word in each sentence. For some sentences, the context provides some information that helps to predict the final word, and for others it does not.

Score your results, separating the data for the two types of sentences. The signal-to-babble ratios for the two tests are 0 dB and +10 dB. Can you estimate how much the signal-to-babble ratio must be increased if context is lacking in the sentences, to obtain an equivalent intelligibility?

A brief report giving your results and making a few comments will be adequate.

3. As part of this lab, we will play the set of brief utterances that each group excised from their recording of an ad lib description of the recording room. For each item, you should write down what you hear, even if you are guessing. Most of the utterances are one or two syllables; some may be three syllables, or more than one word.