Acoustics of Speech and Hearing

Lecture 2-7: Speech Perception Tests

Overview

1. **Tests of speech perception.** We can contrast tests that measure overall speech reception performance with tests that measure the perception of particular phonetic elements. Examples of the former are word or sentence intelligibility tests, where what is counted is the total number of words correctly recognised. Examples of the latter are experiments which analyse the phonetic confusions made by listeners, or which explicitly focus on a listener's ability to perceive a phonetic contrast. We will focus on this second type.

2. **Use of perception tests.** The experimental assessment of how listeners perceive speech sounds is important for a number of reasons. First, we can use perceptual tests to confirm hypotheses about acoustic cues that arise in studies of speech production. For example, although we have seen that Voice Onset Time varies systematically with phonological voicing, we do not know whether listeners make use of this cue. Second we would like to assess the relative importance of different cues, for example the relative importance of the burst cue to place and the F2 transition cue to place. Such information is useful in the design of communication aids such as hearing aids. Lastly we can use knowledge of how normal listeners perceive speech to assess an individual in an audiometric procedure. For example we can check that known stages of development have been reached by a child.

3. **Tests of perceptual confusions.** One way to test phonetic perception is to ask listeners to label (or *categorise*) simple stimuli. We can then analyse the listeners' confusions across categories to look for patterns which reveal aspects of their perceptual abilities. For example we could present VCV syllables with a small range of consonant types, then ask listeners to identify the consonants. Analysis of the consonantal confusions tell us about the robustness of the coding of the phonetic properties in the signal, and the ability of listeners to hear contrasts. To force errors to be made, such experiments are commonly presented in difficult listening conditions, such as in the presence of interfering noise.

4. **Tests of perceptual categorisation.** A characteristic of a developed sensitivity to a phonological category is that acoustic variants within a category seem more similar than variants which are at the border between two categories. It is as if our perceptual system is only interested in the category label for a sound rather than the degree to which it is typical of that category. Unless a /b/ sound is confusable with a /p/, then it's just a /b/. We can use this property of the perception of phonological categories to investigate what properties of a sound make it typical of its category. This type of test often uses pairs of words that differ only by a single phonological feature – called **minimal pairs**. Examples are "coat/goat", "date/gate", and "sip/ship". We generate stimuli that vary between one word and the other, either by editing natural speech signals, or by reconstructing signals artificially using a speech synthesizer. If we want to test responses to single cues, it is important to eliminate or neutralise differences in the stimuli corresponding to the other cues not under test. Typically, we create stimuli in which only a single cue is present, and all others removed or neutralised. Two stimuli representing the two categories are made and then a **stimulus continuum** is constructed in which one category is changed in uniform steps to the other category. The result is a set of stimuli that probe the range of perceptual responses across the range of values taken by one acoustic cue, see Figure 2-7.1. Given a two-choice task, we can calculate the average labelling behaviour for one of the choices as a function of position along the stimulus continuum. The shape of this **labelling graph** can be qualitatively labelled using terms such as **categorical** – showing

consistency of labelling across the continuum, **progressive** – showing a steady shift from one label to another across the continuum, and **random** – showing no clear pattern across the continuum, see figure 2-7.2. The **phoneme boundary** can be measured from the labelling graph; it is the point on the continuum where either label is equally likely. The **confidence** of labelling behaviour is often assessed by looking at the steepness of the labelling graph at the phoneme boundary; more confident labelling will give a steeper gradient.

Readings

At least one from:
- Hewlett & Beck, An introduction to the science of phonetics, Chapter 16: Speech Perception, pp227-236. *Simple description of speech perception testing.*
- Hazan et al, "Speech Pattern Audiometry for Clinical Use", European Journal of Disorders of Communication <u>30</u> (1995), pp116-123.
- Hazan & Fourcin, "Microprocessor-controlled Speech Pattern Audiometry: Preliminary Results", Audiology <u>24</u> (1985), pp325-35.

Historical interest:
- Miller & Nicely, "An analysis of perceptual confusions among some English consonants", J. Acoust.Soc.Am, <u>27</u> (2), 1955, 338-352. *Available on-line.*
- Lisker & Abramson, "The voicing dimension: some experiments in comparative phonetics", Proc. Sixth International Congress of Phonetic Science, Prague, 1967. *Available on-line.*

Learning Activities

You can help yourself understand and remember this week's teaching by doing the following activities before next week:

1. Find out about some purposes to which speech perception tests have been put – what were researchers trying to find out about our perception of speech?
2. An early influential model of speech perception was called the "Motor Theory". Read about the motor theory on the web and write a description in your own words about how it was meant to work.
3. Present some arguments for and against the use of artificial or synthetic speech stimuli in perceptual experiments compared to natural speech stimuli.
4. Write down some ideas for how speech perception stimuli might be used in the clinic for assessment purposes.

If you are unsure about any of these, make sure you ask questions in the lab or in tutorial.

Reflections

You can improve your learning by reflecting on your understanding. Here are some suggestions for questions related to this week's teaching.

1. What is an acoustic cue?
2. Why do multiple cues to a contrast help the listener?
3. Do all languages use the same acoustic cues? How might bilingual speakers use acoustic cues differently to monolingual speakers?
4. What happens when cues are in conflict with one another (in some artificially generated signal)?
5. What are the advantages and disadvantages of using synthetic speech to make stimuli?
6. In the lab session you did a labelling test in which one sound was presented and you had two choices to make. Think of some other ways in which you might do a perceptual test.

7. What information can you get from a labelling test that you can't get with an intelligibility test?
8. Why does the slope of the labelling curve relate to confidence?

Figure 2-7.1 Design of a synthetic speech pattern test

Procedure

1. Choose a phonetic contrast (e.g. rising/falling intonation)

2. Find a suitable minimal pair (Ah?/Ah!)

3. Make a recording of some natural utterances

4. Measure acoustic parameters (Fx, F1, F2, F3, durations, etc)

5. Decide which acoustic cue to vary and which cues to neutralise

6. Synthesize the two extremes with only this cue varying

7. Create a continuum between the two extremes by changing the cue in small steps.

8. Present the stimuli to the listeners multiple times in random order

9. Plot average labelling for each stimulus step

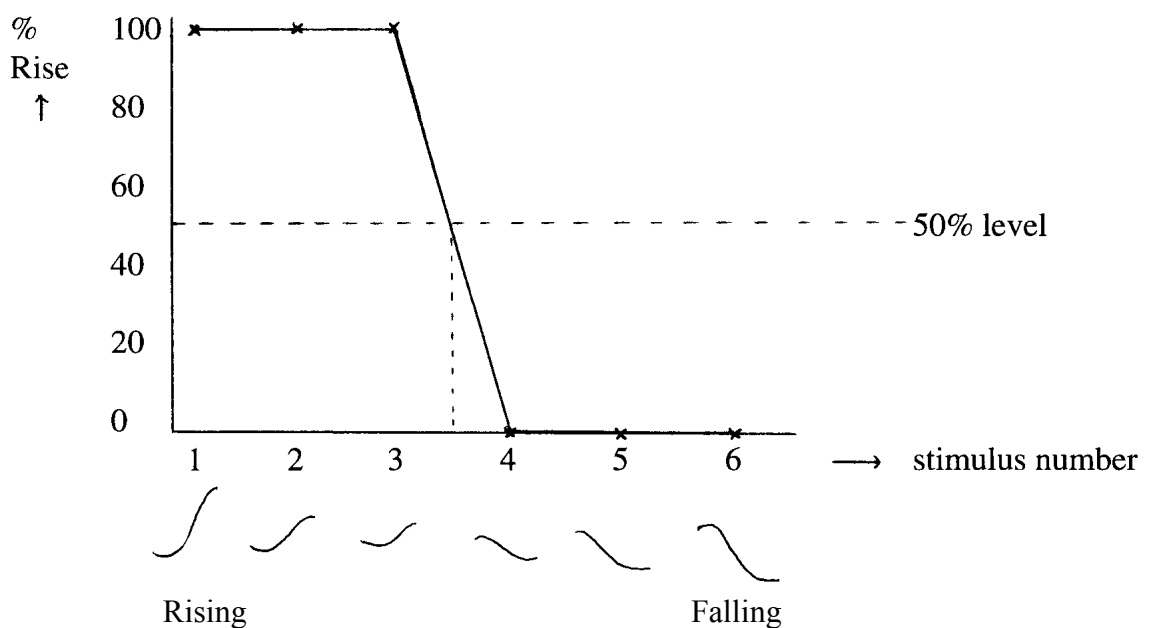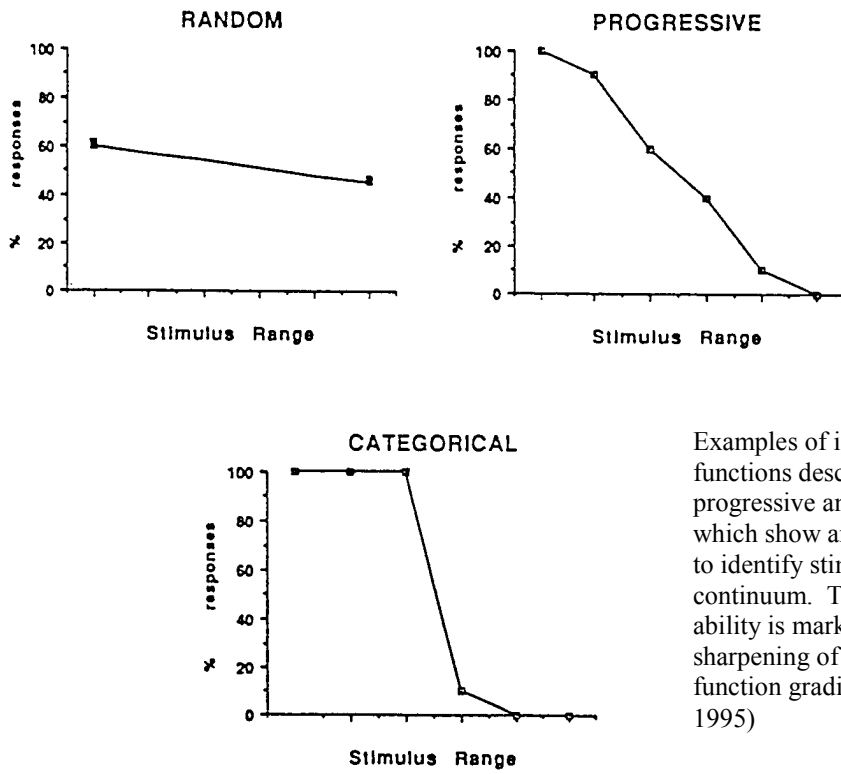10. Measure phoneme boundary and steepness at boundary

Example Labelling Function

Figure 2-7.2 Examples of Labelling Curves

1. Classes of labelling curve

Examples of identification functions described as random, progressive and categorical, which show an increased ability to identify stimuli across a continuum. This increasing ability is marked by a sharpening of the identification function gradient. (Hazan et al, 1995)

2. Development in ability to process a contrast (/ɜ/-/ɑ/) Hazan & Fourcin, 1985.
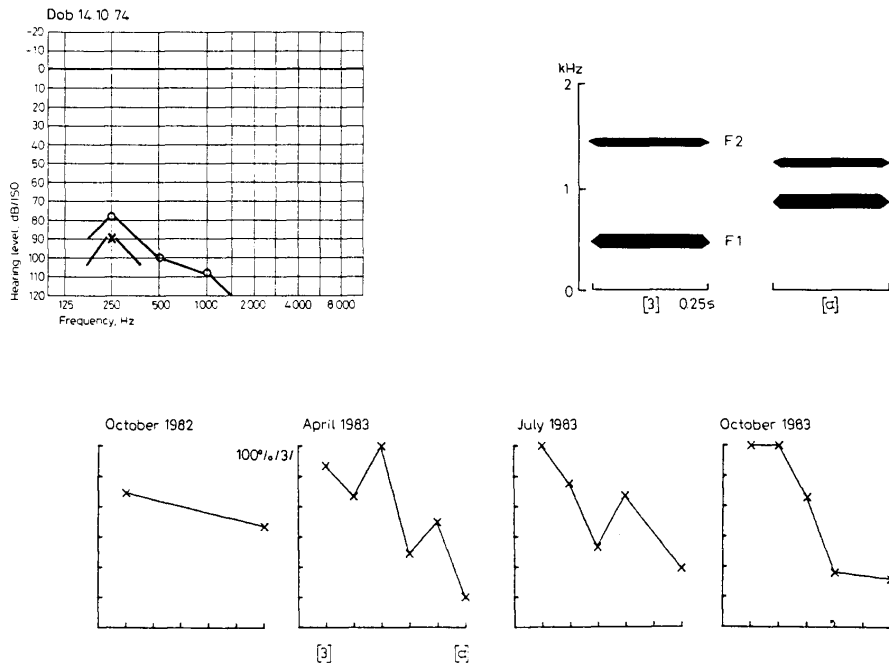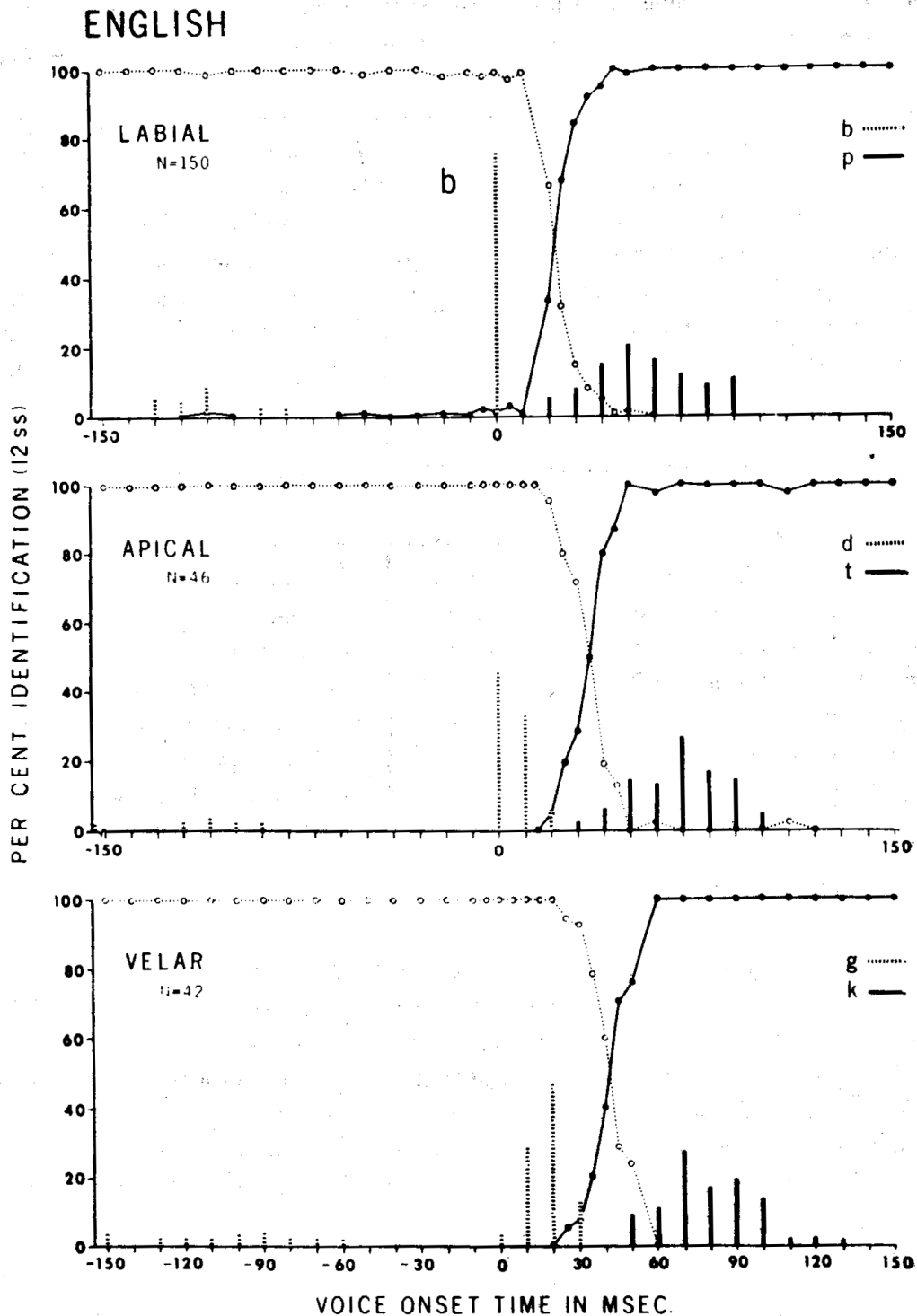
Fig 2-7.3 Categorical labelling of Voice Onset Time Continua

[Lisker & Abramson, 1967]

Synthetic CV continua are constructed for bilabial, alveolar and velar places of articulation in which Voice Onset Time is varied systematically from -150 to +150ms. Listeners are asked to label these with the respective voiced or voiceless phoneme. Note the sharpness of the cross-over and the shift in boundary with place.

Lab 2-7: Speech Perception Tests

Introduction

Tests of speech perception can be used to tell us about the acoustic properties of speech sounds which are important to listener's identification of phonological categories. Such tests can also be used to compare the differences between two listeners when presented with the same speech sounds. In this experiment you will take part in a test of phonetic labelling of some speech sounds and then analyse your own performance. The significance of the stimuli will be explained to you *after* you have done the tests.

Scientific Objectives

- to investigate how intelligibility is affected by background noise
- to investigate which phonetic features are most robustly perceived in noise
- to investigate the sensitivity of listeners to a voicing cue in an identification task.

Learning Objectives

- to show you how speech perception tests are run with listeners.
- to give you an opportunity to analyse perceptual confusions
- to give you an opportunity to generate a 'labelling curve' of your own performance.
- to provide you with your own data so that you can practise interpretation of perception test results.
- to provide you with group data to highlight group trends and individual differences in the perception of phonetic contrasts.
- to revise the acoustic cues to plosive voicing

Method

Audio recordings of the sound stimuli will be played over headphones. You will be provided with a stimulus response sheet to record your judgements. You will need to react quickly to 7each stimulus and to *guess* if you are not sure what to respond rather than missing a response.

1. **Consonant intelligibility**
   You will hear a number of VCV syllables and you need to identify the consonant only from the set /b,p,d,t,g,k,v,f,z,s,m,n/. All stimuli will be played in noise to make the task more difficult.

2. **Plosive Voicing**
   You will hear a number of CV syllables and you need to identify the consonant only from the set /b,p,d,t,g,k/.

Observations

Key-graph sheets will be distributed to help you analyse your responses. For each stimulus number you will see the list of presentation numbers corresponding to a presentation of that stimulus. Transfer your responses to the key-graph sheet, and then plot graphs as described below:

Intelligibility test: plot percentage correct against signal-to-noise ratio for yourself and the class average. Calculate a confusions table for your own results.

Labelling test: plot class average %Voiced responses as a function of VOT for all three conditions on one graph.
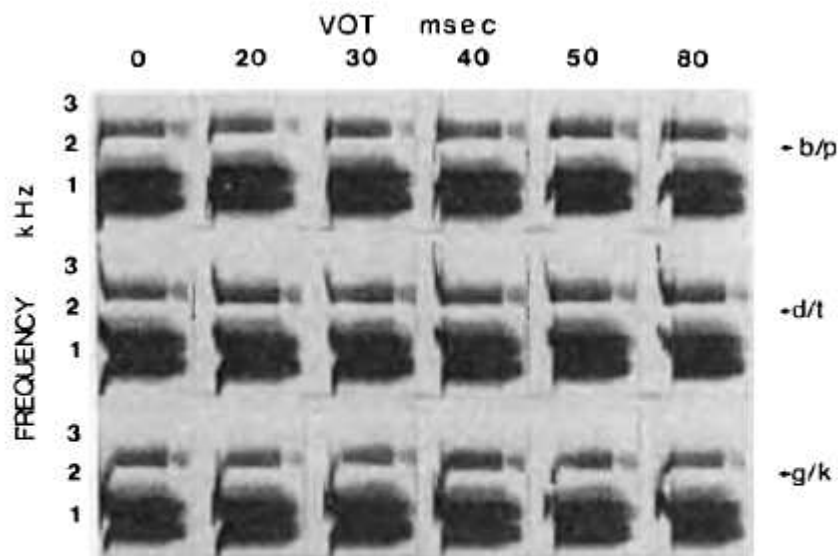
FIG. 1. Spectrograms of synthetic sounds used in these experiments. The voice-onset times (VOT's) are shown across the top of the figure (note that VOT's of 10, 60, and 70 ms are omitted). Each row represents a different place of articulation. These are the stimuli developed by Abramson and Lisker (1970).

Concluding Remarks

1. Describe the stimuli used in the phoneme intelligibility test.
2. Interpret the class graph of intelligibility against signal-to-noise ratio.
3. Looking at the consonant confusion errors you made, do you see any pattern in terms of the relative frequency of voice, place or manner errors? Why might one type of error be more common than another?
4. Describe the stimuli used in the labelling experiment.
5. Discuss the class results on the Plosive Voicing test. Can you link this week's perception experiment with last week's production experiment?

Examination Questions

These are questions from past exam papers. You may like to write outline answers to these, or to discuss them in tutorial.

1. Describe an experiment that could be used to investigate the link between the production and perception of voice onset time (VOT) differences among plosives. Suggest what results might be obtained for EITHER a group of English speakers and a group of French speakers, OR for English speakers at different places of articulation. [2006/7]
2. Describe how a perceptual test of phoneme contrast can be designed and run. Discuss what results might be found for the plosive voicing contrast as a function of voice-onset time (VOT). How do the perceptual results in this case relate to the way plosive voicing is articulated? [2010].