

HOMEWORK III: PERCEPTUAL CONFUSION EXPERIMENT

Ling 411/511 Phonetics: Guion

1. Aim

The goal of this assignment is to provide you with hands-on experience in designing and running a speech perception experiment in order that you gain first hand knowledge of perceptual phenomena as they relate to the acoustics of speech as well as a sense of experimental design, method, and reporting.

2. Overview

A perceptual confusion experiment is an experiment in which you test the confusability of a set of segments with one another. The assumption is that phonemic inventories are sound *systems*. That is, each sound can only be fully understood in relation to other sounds in the system. The logic of a perceptual confusion experiment follows directly from this assumption as it is meant to get at how well individual sounds are perceived with respect to other sounds in the system (think perceptual distinctiveness).

Although it is possible to test the relative perceptibility of all sounds within an entire phonological system in all possible syllable positions, it is usually preferable to ask a specific question about a subset of sounds. A specific question focuses the hypothesis, makes the experiment tractable, and makes for a better argument. For example, Redford & Diehl (1999) asked a question about the relative perceptibility of initial vs. final voiceless obstruents in CVC monosyllables. Guion (1998) looked at the relative perceptual confusion of velar stops and palatoalveolar affricates before front and back vowels. Also, many studies have looked at the cross-linguistic mapping of sounds from one language to another (Schmidt (1996) for Korean speakers listening to English consonants, Guion (2000) for Japanese speakers listening to English consonants, and Strange et al. (1998) for Japanese speakers listening to English vowels).

It is desirable to only look at a couple of variables at a time. For example, Redford & Diehl (1999) were concerned with the effect of syllabic position on perception and kept the consonants constant across syllabic position. As the voiceless obstruents in the stimuli varied by place and manner of articulation, they were also able to look at effects of place and manner. However the design was not overwhelmed with a huge number of potentially influential variables (e.g., voicing, nasality, openness, formant transitions). Guion (1998) was concerned about the effect of the following vowel frontness on misperceptions of velar stops as palatoalveolar affricates and, thus placed both the velar and the palatoalveolars before the same vowel types. As both voiced and voiceless velars and palatoalveolars were used, the effect of voicing was investigated but other extraneous confounding factors, such as syllabic position, were not introduced. Likewise, in cross-linguistic perceptual investigations, the consonants or vowels of interest are placed in similar phonetic conditions. For example, Schmidt used productions of English consonants in CV syllables where the vowels were /i/ /a/ and /u/ and the consonant perception was the focus of the study. In that way, the perception of English consonants under controlled conditions could be studied.

On the other hand, it is possible to overly restrict your question, i.e., your stimuli set.

When you do this you defeat yourself before you even begin because you are unable to collect sufficient data to make a compelling argument for any hypothesis. So, the crucial and difficult task in designing an experiment is the same as the crucial and difficult task in living: striking a balance. In the case of an experiment, you need to find a balance between what you want to know, what time and other resources you have available to you, and how compelling you want your argument to be. In this assignment, you will be asked to come up with a balance between your question, your design, and your argument by yourself.

3. Method

A. STIMULI

The stimuli you create will be entirely dependent on the kind of question you want to ask. Are you curious about the relative perceptibility of the segments that make up the natural class of sonorants? Are you interested in the effect of syllable position on consonant perception? Are you interested in the effect of vowel height on obstruent perceptibility? Are you interested in how English speakers would perceive Korean stops?

Whatever it is, make sure that you create the stimuli to answer the question, and make sure that you control for all other variables that could affect your outcome that you are not interested in studying. That is, if you are interested in the relative perceptibility of sonorants, then vary sonorants, but do not vary vowels, unless you plan on examining the effect of vowels on sonorant perception and how different vowel classes interact with different sonorant classes.

Make sure each type of stimuli is sufficiently represented so that you can get some real data. For instance, if you are studying nasals, make sure that each nasal type shows up several times (ideally 10 or more tokens) in your final set, so that you can collect several responses to this type. The more responses you have per type the easier it will be to compare between types. For example, if you are interested in the confusability of nasals in syllable final position by preceding vowel type [i, a, u], you may have 5 productions each of [tam], [tim] and [tum] for [m], 5 productions each of [tan], [tin] and [tun] for [n], and 5 productions each of [taŋ], [tiŋ] and [tuŋ] for [ŋ]. You should play each of the words once to each listener. Thus, each listener would hear about 45 words (3 nasal types X 3 vowel contexts X 5 tokens).

B. RECORDING

Randomize the order of your stimuli and record them in a frame sentence as *.wav files. When you present the stimuli to the listeners, you may either edit the target words out of the frame sentence, or present them in the frame sentence. Label each of the stimuli phrases with a random order number. That way, when you grab all the files to play them to your subjects, they will play forward in a random order, and the subject will not be able to guess the appropriate response from the file name.

If you choose to present your stimuli set as a continuous audio stream (i.e., one right after the other), then you need to make sure that each token is buffered with enough silent time so as to give the subject enough time for a response.

Make sure that you keep a list of what token occurs in what random order. If you

have no list, you will not be able to evaluate the subjects' responses and you will have no data.

C. SUBJECTS

You need to test a minimum of 2 subjects, so that you can make sure that your results are not due simply to individual weirdness.

D. TASK

Subjects will be instructed in the task ahead of time, and informed that they are participating purely on a voluntary basis, and that they can quit the task at any time for any reason.

Subjects will be asked to wear earplugs (make sure you specify in your methods section the advertised attenuation qualities of the earplugs that you are getting your subjects to use), and to listen to your stimuli set at a preselected volume (that you will also note in your methods section). Subjects will be instructed to write down or select the nonsense word that they hear on the response sheet that you provide. Depending on the sounds to be analyzed, you may have to address issues of English orthography with your subjects. For example, if you are testing both [ð] and [θ], you should use key words like 'th' in *the* and 'th' in *thank*. English vowels will also require key words.

E. ANALYSES

After you have run your subjects, you will need to score their data, in order to conduct your analyses. Note the response for each item. Then tabulate the number of correct responses for each type. You should also tabulate the number of incorrect responses by incorrect response type. Continuing with our final nasal example, 3 subjects, each hearing 5 tokens of [tam] responded with [m] (correct) 10 times, with [n] (incorrect) 1 times and with [ŋ] 4 (incorrect) times.

4. Results

It is important that you present your findings in a coherent and organized fashion. In particular, you should be organizing your findings in a way that seems to get at the question that you were asking in your design. For instance, if you tested the relative perceptibility of initial vs. final consonants, then you should definitely be comparing the perceptibility of initial and final consonants in your results section. In addition, remember to explore other factors that may be working for or against your question. For example, show how place and manner of articulation influence your overall finding that final consonants are less well perceived than initial consonants.

All of these comparisons and rhetorical points should be made in tables and charts. I suggest using confusion matrices like those found in Johnson Ch 4 (Table 4.1, 4.3) (on e-reserve) and in many of the papers in the reference section below. You can report the number of responses or proportion (or percent) of responses. Either way, state the total number of responses. The text in the result section (and there should be text in the results section) should be built around the tables and charts where you are displaying your main results. Minor results should also be presented, but these should only be commented upon

in the prose of this section.

5. Discussion

This section is devoted to your argument based on your findings. Think about what you have learned in your experiment, and coordinate that with what you have learned from class and from the readings. Make the whole discussion gel around a single idea or related set of ideas.

6. Conclusion

This is a brief note about what you have concluded from your study and what you think the next logical step should be in your search for the ultimate understanding of X.

7. References

- Guion, S.G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, 55, 18-52. (available on e-reserves)
- Guion, S.G., Flege, J.E., Akahane-Yamada, R. & Pruitt, J. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America* 107, 2711-2724. (available by taking link here: [PDF](#))
- Redford, M.A., Diehl, R.L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America*, 106, 1555-1565. (available by taking link here: [PDF](#))
- Schmidt, A.M. (1996). Cross-language identification of consonants. Part 1. Korean perception of English. *Journal of the Acoustical Society of America*, 99, 3201-3211. (available on e-reserves)
- Strange, W., Akahane-Yamada, R., Kubo, R., Trent, S.A., Nishi, K., & Jenkins, J. (1998). Perceptual assimilation of American English vowels by Japanese listeners. *Journal of Phonetics*, 26, 311-344. (available on e-reserves)