

## Term Project

Over the course of the semester, we will have you work in groups to conduct the core of a project in phonetic research. We will have you work in groups of (approximately) three to make this tractable for all of us, allowing you the opportunity to pool your resources, and allowing us to focus on a smaller number of projects to maximize the likelihood that they will turn out well. Although the research itself will be a group effort, each student is expected to be involved with each of the phases of the research, and to submit his or her own write-up of the experiment. The reasons for this approach are two fold: 1) having individual write-ups encourages you all to understand what the research is showing - there have been cases in which two of three group members really had no idea what the project was for, and 2) this will give you your own independent record of the project for future reference. One final note about these projects is that they generally yield interesting and novel results and hence can easily form the core (or at least seed) of publishable research.

### Keep a log!

While doing the research, keep a careful log of what you are doing and why. Each individual ought to do this at every step of the research, including sketching out the topics of interest. This will be invaluable for you later on when you write up the research report. You will be surprised at how much you have forgotten in the space of a semester.

### Timeline:

The project will proceed according to the following steps and with the following time-line. (Bold-faced items here indicate documents you will turn in.)

- 1) Weeks 1 – 3: Form groups around related research questions
  - Jan. 26: submit **Draft Sketch** of research questions and proposed group
- 2) Weeks 4 – 5: Decide on research question and sharpen it
  - Feb. 9: get approval on **Sketch** for topics and groups
- 3) Weeks 6 – 7: Formulate **Research Prospectus** laying out procedure for project
  - Feb. 23: submit draft prospectus
- 4) Weeks 8 – 9: Sharpen **Research Prospectus** by incorporating feedback & trial runs
  - March 8: submit final **Research Prospectus**
- 5) Weeks 10 – 11: Run experiments
  - March 29: have recordings of subjects in hand
- 6) Weeks 11 – 12: Perform measurements and analyses
  - April 5: submit **Measurement and Analysis Scheme**
- 7) Weeks 13 – 14: Complete analyses and work through conclusions

**1) Form groups around research questions** relating to a topic in phonetics. You will have to figure out who else is in the class, and find people with whom you might have common interests. Common interests could be a common language of interest, or it could be topical (e.g. interests in second language learning). Begin discussing possible topics, and be open minded. Interesting phonetic topics come in many varieties. Here are a few suggestions:

**A.** Try to figure out some aspect of the acoustic (or articulatory) manifestation of a funny contrast. Note that 'funny' does not preclude familiar. *Example:* voiced and voiceless stops in English. Varieties of this topic would take into account differences in the manifestation of a contrast in different syllable positions or in different prosodic locations, or examine the effect of focus on a contrast.

**B.** The topics in A can be applied to the manifestation of a single sound as well, if there's some idea as to how variation in a single sound might say something interesting. *Example:* tense mid vowels in English and the manifestation of diphthongization. Are these things diphthongs or monophthongs? You can ask questions as to how such phonetic events vary with respect to factors such as region or social context, or in different consonantal environments.

**C.** Investigate a phonological rule to see if speakers really behave that way, and see how the rule might be sensitive to various prosodic or pragmatic factors. *Example:* word-final devoicing in German. One course project back in the late seventies on this topic turned into a research program which is still being pursued today up at SUNY Buffalo.

**D.** There are a host of questions related to second language learning. *Example:* how Korean speakers deal with word-final stop contrasts. Or how Japanese speakers deal with rising pitch accents.

Consult previous work on the subject, if available, since this often will seed some new ideas much quicker than having to find them from scratch.

Once you have discussed this and find a group of three or so that have some common ideas, formulate a **Draft Sketch**. Email the Sketch to both instructors and each proposed member of the group sometime *before* January 27.

**Draft Sketch** = one-paragraph descriptions of at least three projects which are at least potentially feasible. At the top, include a list of the proposed group members.

**2) Decide on research question and sharpen it.** After you submit the possible research questions, we will respond to the group with our observations about the feasibility of each of the questions, and suggestions for making them sharper, more interesting, or more feasible. Soon after this, the group should work out a final decision as to which topic will be pursued, remove the topics that are not going to be pursued, sharpen up the description of the topic that will be pursued, answering an questions posed by the instructors about the topic, and again email the final **Sketch** to the group members and instructors *before* February 10.

**3) Formulate Research Prospectus laying out procedure for project.** As soon as the topic has been decided and agreed upon by all of the parties, the group needs to work out the details of what is actually to be done, including why. This material should be collected in a draft of a **Research Prospectus** to be submitted by email to both instructors and cc'ed to all of the group members. All members of the group must approve the draft prospectus, though you may divide up the work of creating the prospectus however seems best among you. Submit the prospectus

articles follow the format used here.

*a) Introduction & background.* The first section needs to lay out crucial background knowledge about the topic of the project that is necessary to understand what the topic is, and why it might be interesting. The introduction should have the necessary information on each crucial aspect of the project, and it should end with a paragraph stating what the project is and what you intend to accomplish by it. Crucial aspects of the project include basic information about the language or variety (or varieties) that are being examined, and the relevant phonological structures. Other crucial aspects would be features of the project that you will work hard to control, because it is a central part of the question you are asking. E.g., if you are looking at second language learners, then what is their first language, and why does this matter? If you are looking at allophonic variation, then what is the received wisdom as to the distribution of the different variants and what triggers them? Etc.

*b) Research question & hypotheses.* State in a short section the exact questions you are attempting to answer, along with likely outcomes to your project and what the different outcomes would mean.

*c) Methods*

- *Subjects.* Who are you going to work with, and why? How many are you going to work with? Where are you going to get them?

- *Corpus.* What linguistic structures are you going to look at and why?

- *Procedure.* What are you going to have the subjects do (exactly)?

- *Analysis.* Once you have the records from the procedures, what are you going to do with them to find the answers to your original questions?

*d) Division of labor.* Who in the group is going to do what?

*e) References.* Give citation information for any previous work that you cite in the work.

**4) Sharpen Research Prospectus by incorporating feedback & trial runs.** After you turn in the prospectus, we will mark up the document and return it to you with suggestions and questions. You should resolve these as a group. Some of these will require some rethinking of the basic idea or changes to the design of the study. In cases where we all are unsure if the project is likely to work, you can try it out on someone to get feel for how it might go. At any rate, you want to have a final, agreed-upon version of the prospectus emailed to both instructors and cc'ed to the group members sometime *before* March 9. We will then, as soon as we can, return it to you with some final feedback, after which you should move on to the data collection phase.

**5) Run experiments.** Having lined up the speakers during the process of figuring out the topic, you should be in a position to actually run the experiments sometime *before* March 30. We will work with you to sort out any details which were not resolved in the prospectus stage.

**6) Perform measurements and analyses.** Many analyses will make use of spectral analysis software, often the software in the Phonetics Lab. The exact nature of the analysis will depend, of course, on the topic. Things that are relatively straight-forward to measure are durations of

fairly straightforward to do, and provide part of the answer to your original questions.

You will have already specified your estimate of the analyses to be performed in the prospectus; however, guessing ahead of time and actually working with real signals can be two very different things. To help us through these hazards, you should email (or hand a hardcopy of) the **Measurement and Analysis Scheme** to us as soon as you have the data and have had a chance to attempt to do analyses. This had better be *before* April 6.

**Measurement and Analysis Scheme** = images and explanatory text indicating what measurements are to be drawn from the original data, *with it worked out on a particular example from the data itself*.

The scheme should include four items.

- a) A short narrative of how you go about extracting measurements from the recordings of subject behavior, with enough detail for us to know that you have actually attempted it and succeeded in getting the appropriate data. Before finishing this, you should do a sample of measurements with the various speakers and conditions you are examining, to make sure that you have an idea that this will actually work for most of your data.
- b) Any questions which arise in the process of figuring out how to measure items, so that we can address them.
- c) A print out of a (very) small number of representations of utterances that have the relevant information indicated on them. The exact form of the representations will depend on the project.
- d) Finally, tell us what the exact data that arise from the measurement scheme will be.

Particular examples are given below.

1) Suppose that you have a project that measures vowel quality and duration. You should give us a document specifying how the two types of measurements are to be done. For example, for vowel quality, you have to determine where in the vowel you are measuring, so give an example of a marked up spectrogram or waveform indicating the time of various measures. Then, at that time, if you use a formant extraction algorithm, you should indicate the settings used for the extraction algorithm, e.g. by printing out the dialogue box where the settings are entered. For vowel duration, indicate what visual display you will use (probably combination of broadband spectrograms and waveforms), along with marked up examples. This example will produce F1, F2, and F3 estimates at times X & Y (indicated on the spectrogram), as well as the time of release going into the vowel, and closure coming out of the vowel.

2) Suppose you are doing more of a typological classification of types of rhotics in language X that appear in different linguistic contexts. Here, you must include a list of the types that you will be classifying utterances into, along with a narrative as to how you know what class they belong in. Include example spectrograms of each type, with any quantitative information that you might get on the different types. Resulting data here are the number of instances of particular types for each speaker in various contexts, along with measurements of duration of items in types c – f (or something like that).

3) Projects may involve perceptual identification. If this is the case, then indicate in your narrative what the stimuli are, and what the exact task is. For example, if you have a continuum of different durations in a quantity comparison study, state how many continua you have, and how many steps there are in the continua. Then, if the task is identification, indicate what the

**7) Compile and make sense of your analyses.** Your data should be presented in some interpretable form, including relevant graphs or tables. In my experience, graphical devices which show relationships between measurements provide the most information. Since we do not have a uniform background in statistical reasoning in the class, we will not be requiring the use of inferential statistics, though you will (generally) be using measures of central tendency (means) and dispersion (standard deviations) to get a general idea about the amount of spread and divergence between groups in your data. Important tools for doing this are various methods of graphical visualization of the data (figures). These figures will act as an important part of our interpretation of the overall outcome of the study. To help us interact over these things, you should email to us and to each group member a set of **Summary Figures** sometime *before* April 20.

**Summary Figures** = a document with

- a) a list of questions you are asking with the research.
- b) graphical representations of the data that contribute substantially toward an answer to the questions.

Exact forms of what you turn in will depend on the topic, so we should converse about this as we go along. Examples are given below.

- 1) Suppose that you are asking about the degree of approximation of the production of a vowel quantity distinction by English learners of that language, and you have a native speaker to compare with. The representation here should allow for comparison across the speakers, either by having one figure with averages for the speakers superimposed, or by having closely related figures. The most likely graph here would be a scatter-plot with F1 (Hz) as a function of duration (ms), on the hypothesis that English speakers will try to do quantity by using stress which disproportionately increases F1 for a given durational lengthening.
- 2) Suppose you are doing a typological classification of rhotics. The representation here (at least for the typological classification part) would be a bar chart indicating the proportional frequency of the different types for different linguistic contexts or different speakers.
- 3) Suppose you are doing the identification task with a continuum. The most appropriate representation here would be identification functions for the continua; identification functions plot the ordinal location of the stimulus on the continuum scale on the horizontal axis against the proportion of times that stimulus was identified as X. Since studies like this compare different continua or different speakers, multiple speakers or continua can be plotted in the same figure.

**8) Write the Final Paper.** A thoughtfully-written description of the project is an under-rated portion of the research. Research is not completed until you are able to tell someone about it. Also, you will find that writing up a project convinces you that you found out something different than you thought you found out. This step should require a lot of thought at this juncture, but not the generation of a lot of text! The prospectus should form the shell of the paper, and if you are generating the results write up as you do the analysis, much of the ms will have been done before this stage.

**Final Paper** = a document including an *introduction*, *statement of questions and hypotheses*, *methods* including subjects, corpora, and analyses, *results* presenting the figures which summarize the data along with parallel text to explain what's in the figures, *discussion* relating these figures briefly to the original questions being asked, along with brief *commentary on issues that the results raise for future research*, a *conclusion* stating the basic

### **Additional tips for setting up and running a phonetics experiment.**

**Illustrative suggestions of topics:** Following are examples of topics. Um, the exact topics listed below are actually ones which have been done quite a bit, but extending them to other languages is likely to be novel.

1. Effect of phonemic voicing on vowel duration and obstruent duration in German (or English or...). This is generally an interesting topic, and is especially interesting in German where there is phonemic merger of the voiced and voiceless stops. Another angle on durational studies is to vary speech rate, stress, or focus to see how it affects the contrasts you are investigating.

- pick 3 or 4 minimal pairs; get 4 repetitions from three speakers for a total of 60 tokens. Measure the stop, fricative and vowel durations.

2. Effect of high-vowel devoicing on timing in Japanese.

- pick 2 minimal pairs like /kakusi/ and /kakesi/, get 5 repetitions from three speakers. Then have them focus on the relevant word. Measure the durations of the various parts of the words and when and where devoicing happens and whether it is affected by speaker attention.

3. What is the relevant dimension of contrast corresponding to [ATR] in Twi?

- pick 2 or three minimal pairs for 3 of the five vowels, get 5 repetitions from two or three speakers. Measure the formant values and durations in the various vowels.

4. Effect of speech rate or prosodic position on diphthongization of American vowels.

- pick a subset of vowels, perhaps tense/lax pairs, which may or may not diphthongize. Put them in various prosodic positions and compare formant values from the first half and last half of the vowel, as well as rate of change.

### More complicated designs (careful to make them simple):

5. Compare pharyngealized and unpharyngealized vowels in Arabic in two conditions; one condition in which pharyngealization is clear, and one in which you're not sure. What makes this complicated is the number of linguistic conditions which could be compared. Back in the 1980's there was a cottage industry in formulating rules for feature spreading, such as is involved in this case.

6. Compare affricates and fricatives in language X. The relevant measure here would be how quickly the frication turns on. What makes this more complicated is you need to come up with a measure which is not typical in phonetic studies.

### Other directions for thought:

7. Do a description of some aspect of an undocumented language; for example, plot out the vowel space.

8. Find out if the fricatives in Korean are really the same as those in English. Use a small number of tokens and measure fricative quality, and/or make palatograms of one or two speakers.

9. Documenting foreign accent. Any topic – probably has not been done.