

# PSY 3393 Experimental Projects Spring 2008

Dr. Peter Assmann

## Dates

- Draft Methods section
  - *due date: Tue Apr 8*
- Hypothetical data write up
  - *due date: Tue Apr 15*
- Homework 6 – graphing 2-way interactions
  - *due date: Tue Apr 15*
- Revised Introduction section
  - *Optional but highly recommended*

## Hypothetical Data Writeup

- Before you collect the data for your second project, use the predictions you made in the Introduction section to develop (i.e., make up) a set of hypothetical data that fit these predictions.

**Note: the data you report in your actual project must be collected, not made up!!**

- Write up these hypothetical results in APA format.
- Use “dummy” (hypothetical) values for  $F$ ,  $df$ , and  $p$ .
- **Date due: April 15**

**Homework 6  
Due: April 15**

## Graphing interactions

- For each of the examples, construct a plausible set of data representing the outcome you think is most likely.
- Use the Excel graphing tools to make line plots of the interaction.
- Include figure caption (follow APA format)
- Include error bars on the graphs (optional).

## Two-way interactions

1. Jury decisions are influenced by the attractiveness of the defendant (male or female).
2. Visual imagery improves memory (immediate vs. delayed).
3. Providing courses in family planning in middle school reduce the incidence of teenage pregnancy (grade level).
4. A new program is developed to increase reading awareness in kindergarten children (normal and dyslexic).
5. Non-native speakers have more difficulty than native speakers understanding speech in noisy conditions (quiet and noise).

## Revised Introduction

- **Common problems that need to be addressed:**
  - Not enough background information
  - Too brief
  - Unsupported statements
  - Too few (or no) references
  - Informal writing style

[http://www.wadsworth.com/psychology\\_d/templates/student\\_resources/workshops/res\\_method/mistake/mistake\\_02.html](http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/res_method/mistake/mistake_02.html)

## Revised Introduction

- Common problems that need to be addressed:
  - Research problem not outlined in detail
    - Readers should be able to determine what gap in the literature your study aims to fill, and why the topic is a good choice for investigation
  - Rationale for the study too sketchy or missing
    - Theoretical reasons why the study is important
    - Practical reasons why the study is important

[http://www.wadsworth.com/psychology\\_d/templates/student\\_resources/workshops/res\\_method/mistake/mistake\\_02.html](http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/res_method/mistake/mistake_02.html)

## Revised Introduction

- Common problems that need to be addressed:
  - Description of experiment incomplete or missing
  - No discussion of research hypothesis
  - No predictions or expected outcome

[http://www.wadsworth.com/psychology\\_d/templates/student\\_resources/workshops/res\\_method/mistake/mistake\\_02.html](http://www.wadsworth.com/psychology_d/templates/student_resources/workshops/res_method/mistake/mistake_02.html)

## Concept → manuscript

- Hypothetical topic area: *voice recognition*
  - How do we recognize a voice as familiar?
  - Each familiar voice has a unique acoustic pattern that we can learn to distinguish from other voices.

## Literature search

- Topic area: *voice recognition*
  - Topic appears too broad: database search yields far too many articles, and many appear to be on unrelated topics. What to do next?

## Literature search

- What is known about the topic?
- Find some general background sources:
  - Nolan F. (1985). *The phonetic basis of speaker recognition*. (Cambridge Univ. Press).
  - Doddington G. (1985). *Speaker recognition: Identifying people by their voices*. Proceedings of the IEEE 73(11), 1651-1664.

## Concept → manuscript

- Find the right search terms:
  - Voice/speaker/talker recognition
  - By humans/computers
- Omit related topics (not directly relevant):
  - Automatic speaker identification
  - Automatic speaker verification

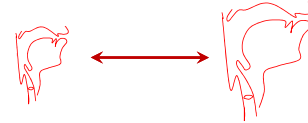
## Concept → manuscript

- People are very good at recognizing voices.
- Voices differ acoustically in several ways (e.g., the pitch and resonant frequencies differ).
- A familiar voice has a unique acoustic pattern that we can learn to distinguish from other voices.



## Research problem

- *Age-related variability*
  - But voices change over time. As children get older, their voices change: voice pitch drops and the resonances are lowered. How do we adjust to such changes?



Child's vocal tract

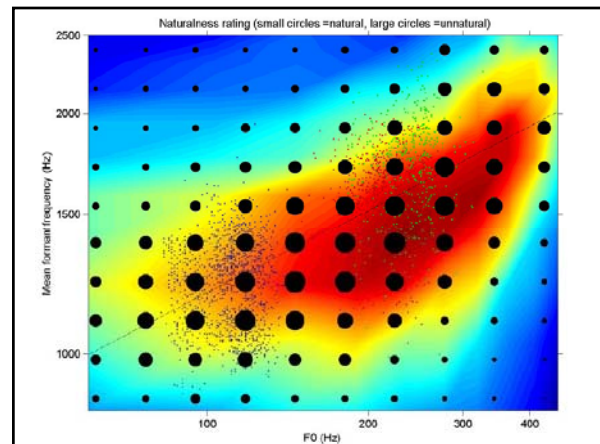
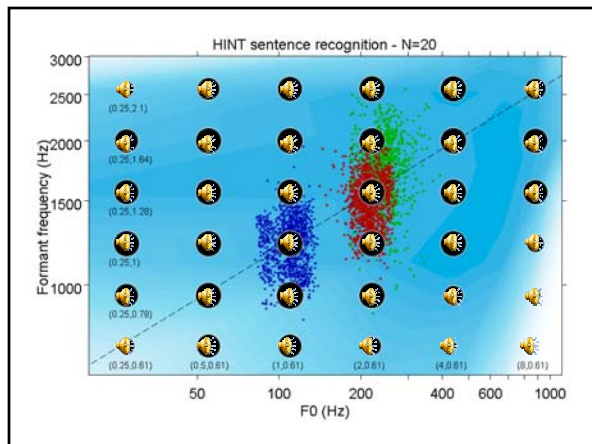
Adult vocal tract

## Research hypothesis

- Family members and friends hear changes in a person's voice on a daily basis.
- Acoustically, these short-term changes are fairly small. It might be difficult or impossible to recognize a person's voice if these maturational changes take place instantaneously rather than gradually.

## Research question

- What if we could use a computer to artificially change the voice of a child into an adult or vice versa? Would we still recognize a child's voice if the pitch and resonance frequencies were shifted to the adult range? Would we recognize our parents' voices as children?
- Recent developments in speech technology makes it possible to simulate such voice changes!



## Method

- Option 1: Use famous voices
  - Problem: familiarity may vary across listeners
- Option 2: Use a small set of voices and require listeners to learn to recognize them
  - Training stage
  - Test stage
  - Generalization

## Background study

- Sheffert SM, Pisoni DB, Fellowes JM & Remez RE. (2002). *Learning to recognize talkers from natural, sinewave, and reversed speech samples*. J. Exp. Psychol. Hum. Percept. Perform. 28(6): 1447-1469.



natural



sine-wave speech



reversed

## Sheffert et al. (2002)

- **Abstract:** In five experiments, the authors investigated how listeners learn to recognize unfamiliar talkers and how experience with specific utterances generalizes to novel instances. Listeners were trained over several days to identify 10 talkers from natural, sinewave, or reversed speech sentences.

## Sheffert et al. (2002)

- The sinewave signals preserved phonetic and some suprasegmental properties while eliminating natural vocal quality. In contrast, the reversed speech signals preserved vocal quality while distorting temporally based phonetic properties.

## Sheffert et al. (2002)

- The training results indicate that listeners learned to identify talkers even from acoustic signals lacking natural vocal quality. Generalization performance varied across the different signals and depended on the salience of phonetic information. The results suggest similarities in the phonetic attributes underlying talker recognition and phonetic perception.

## Procedure

- *Training phase.* Listeners were trained over several days to name the 10 talkers of the sinewave utterances. They were tested in groups of three or fewer in a quiet listening room. During each training session, each subject heard a random ordering of five repetitions of three sentences from each talker (150 items total). There was no blocking by talker or sentence. The same three sentences were used for each talker in each training session.

## Procedure

- *Familiarization phase.* Before beginning each of the generalization tests, all subjects completed a brief familiarization task to reinstate the correspondence between the sinewave tokens and the talker's names.

## Procedure

- *Familiarization phase.* The familiarization task was simply an abbreviated version of a training session in which subjects listened and responded to one instance of each sentence produced by each talker (30 items total). The items were presented in a random order, and accuracy feedback was given after each response. The familiarization task lasted approximately 8 min.

## Procedure

- *Generalization tests.* After reaching a 70% correct criterion in the sinewave training phase, each subject completed two generalization tests. One generalization test presented three unfamiliar sinewave sentences, whereas a second test presented three unfamiliar naturally produced sentences.

## Procedure

- *Generalization tests.* Half the subjects received the natural generalization test before the sinewave generalization test, whereas the other half received the tests in the opposite order. Each test presented five repetitions of each of the three sentences in a random order (150 items total). Once again, subjects were provided with a transcription of the sentences they would be hearing. Their responses were not corrected during either of the two generalization tests.

## Talker recognition

- What is the dependent variable?
- What are the independent variables?
  - should be apparent when reading the abstract

## Sheffert et al. (2002)

- **Abstract:** In five experiments, the authors investigated how listeners learn to recognize unfamiliar talkers and how experience with specific utterances generalizes to novel instances. Listeners were trained over several days to identify 10 talkers from natural, sinewave, or reversed speech sentences.

## Talker recognition

- Dependent variable: proportion (percentage) of voices correctly recognized (out of 10)
  - What about bias (e.g., all of the sinewave voices might sound like talker #1)?
- What if the task is too easy? Too hard?

## Talker recognition

- Independent variables:
  - Type of speech used in training stage (natural, sine-wave, reversed)
  - Type of speech used in test phase (natural, sine-wave, reversed)
  - Talker gender
  - Number of training days

## Data table: Length of training

Table 2  
Median Number of Training Days and Mean Proportion Correct for Talker Recognition in Experiments 1–5

Experiment	Training condition	Training days	Generalization test performance*		
			Sinewave	Natural	Reversed
1	Sinewave	13	.44	.46	
2	Natural	1	.27	.88	
3	Natural	2	.22		.53
4	Reversed	5	.16		.59
5	Reversed	5	.23	.56	

\* Estimated proportion correct from guessing alone is approximately .10.

## Generalization Performance

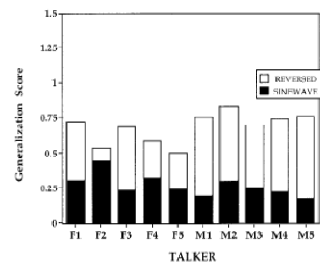


Figure 6. Mean generalization scores on the reversed speech generalization test and the sinewave replica speech generalization test in Experiment 1. Performance is displayed using a value-added stacked bar graph. Performance on the reversed speech test is represented by the height of the entire bar; performance on the sinewave replica test is represented by the dark section of each bar. Performance is displayed as a function of talker. F1 through F5 refer to the female talkers; M1 through M5 refer to the male talkers.

## Follow-up study

Sheffert SM, Olson E. (2004). Audiovisual speech facilitates voice learning. *Percept Psychophys.* 66(2): 352-362.

In this research, we investigated the effects of voice and face information on the perceptual learning of talkers and on long-term memory for spoken words. In the first phase, listeners were trained over several days to identify voices from words presented auditorily or audiovisually. The training data showed that visual information about speakers enhanced voice learning, revealing cross-modal connections in talker processing akin to those observed in speech processing. In the second phase, the listeners completed an auditory or audiovisual word recognition memory test in which equal numbers of words were spoken by familiar and unfamiliar talkers. The data showed that words presented by familiar talkers were more likely to be retrieved from episodic memory, regardless of modality. Together, these findings provide new information about the representational code underlying familiar talker recognition and the role of stimulus familiarity in episodic word recognition.

## Introduction section

When a talker produces an utterance, the listener simultaneously apprehends the linguistic form of the message as well as the nonlinguistic attributes of the talker's unique vocal anatomy and pronunciation habits. Anatomical and stylistic differences in articulation convey an array of personal or indexical qualities, such as personal identity, sex, approximate age, ethnicity, personality, intentions or emotional state, level of alcohol intoxication, and facial expression (see Bricker & Pruzansky, 1976; Chin & Pisoni, 1997; Cook & Wilding, 1997; Doddington, 1985; Kreiman, 1997; Scherer, 1986; Tartter, 1980; Walton & Orlikoff, 1994).

Sheffert et al. (2002). *J. Exp. Psychol. Hum. Percept. Perform.* 28(6): 1447-1469

## Introduction section

Personal characteristics play an important role in communicative interactions. This is especially true for listeners who are unable to use indexical attributes available in other modalities as a result of neurological impairments in face recognition (*prosopagnosia*: Benton & Van Allen, 1968; Bodamer, 1947; Damasio, Damasio, & Van Hoesen, 1982) or visual impairments (Bull, Rathborn, & Clifford, 1983; Yarmey, 1986). Over the course of a lifetime, listeners acquire very detailed and enduring knowledge about many different talkers. The ability to recognize a talker begins in utero (Hepper, Scott, & Shahidullah, 1993) and develops rapidly throughout infancy and childhood (DeCasper & Fifer, 1980; Jusczyk, Hohne, Jusczyk, & Redanz, 1993; Mandel, Jusczyk, & Pisoni, 1995), reaching adult levels of proficiency by age 10 (Mann, Diamond, & Carey, 1979).

*Sheffert et al. (2002). J. Exp. Psychol. Hum. Percept. Perform. 28(6): 1447-1469*

## Introduction section

An extensive literature on human talker recognition dates to the work of McGehee (1937), who examined the reliability of ear-witness testimony, and the studies by Peters (1955) and Pollack, Pickett, and Sumby (1954), who examined laboratory effects of linguistic content on talker recognition. This literature describes the effects of acoustic, procedural, and individual attributes that affect the recognition and discrimination of unfamiliar talkers (see Bricker & Pruzansky, 1976; Clifford, 1980; Hecker, 1971; Kreiman, 1997; and Read & Craik, 1995, for reviews).

*Sheffert et al. (2002). J. Exp. Psychol. Hum. Percept. Perform. 28(6): 1447-1469*

## Introduction section

In contrast, much less is known about how a listener recognizes a familiar talker beyond the benchmarks that reveal perceptual, cognitive, and neural differences in the classification of familiar and unfamiliar talkers (Papçun, Kreiman, & Davis, 1989; Schmidt-Nielsen & Stern, 1985; Schweinberger, Herholz, & Sommer, 1997; Van Lancker & Canter, 1982; Van Lancker & Kreiman, 1987; Van Lancker, Kreiman, & Cummings, 1989). Moreover, few studies have examined how a listener becomes familiar with a talker (Legge, Grosman, & Pieper, 1984; Nygaard & Pisoni, 1998;

*Sheffert et al. (2002). J. Exp. Psychol. Hum. Percept. Perform. 28(6): 1447-1469*

## Introduction section

Nygaard, Sommers, & Pisoni, 1994). These studies show that repeated or extended exposure to a talker's speech increases a listener's sensitivity to talker-specific attributes, improving the ability to differentiate familiar from unfamiliar talkers. Left unspecified, however, are the properties of the speech signal that are most relevant for learning and recognizing familiar talkers from novel utterances.

*Sheffert et al. (2002). J. Exp. Psychol. Hum. Percept. Perform. 28(6): 1447-1469*

## Introduction section

The research described in this article investigated the recognition of familiar talkers, examining the contribution of different talker-specific properties of a speech signal to perceptual learning. To set the task in this experimental design, we trained our listeners to identify different talkers using signals that were acoustically modified to preserve different properties that were arguably talker-specific. Listeners heard sentence-length natural, sinewave, or reversed speech samples. Their knowledge of the talker was then

*Sheffert et al. (2002). J. Exp. Psychol. Hum. Percept. Perform. 28(6): 1447-1469*

## Introduction section

assessed using generalization tests in which a novel set of natural, sinewave, or reversed speech samples were used and listeners were asked again to identify the talkers. Our intention was to permit a comparison of the attributes available in the learning conditions and in the generalization tests with those proposed in several classic and recent accounts of individual identification. This comparison allowed us to assess the extent to which talker identification exploits segmental phonetic attributes and to evaluate evidence favoring a dissociation between indexical and phonetic processing in speech perception.

*Sheffert et al. (2002). J. Exp. Psychol. Hum. Percept. Perform. 28(6): 1447-1469*