2004-06-28

Hearing the Difference: A Computer-Based Speech-Perception Diagnostic Tool for Non-Native Speakers of English

Justin Reed Shewell
Brigham Young University - Provo

HEARING THE DIFFERENCE: A COMPUTER-BASED

SPEECH-PERCEPTION DIAGNOSTIC TOOL FOR

NON-NATIVE SPEAKERS OF ENGLISH

by

Justin Reed Shewell

A project submitted to the faculty of

Brigham Young University

in partial fulfillment of the requirements for the degree of

Master of Arts

in

Teaching English to Speakers of Other Languages

Department of Linguistics and English Language

Brigham Young University

August 2004

BRIGHAM YOUNG UNIVERSITY

GRADUATE COMMITTEE APPROVAL

of a project submitted by

Justin Reed Shewell

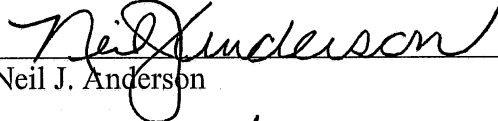This project has been read by each member of the following graduate committee and by majority vote has been found to be satisfactory.

28 June 2004

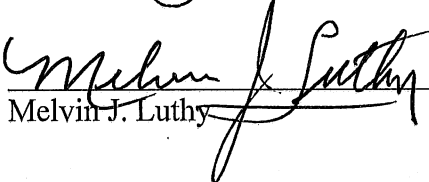Date _____   _____
                                  Lynn E. Henrichsen, Chair

25 June 2004

Date _____   _____
                                  Neil J. Anderson

28 June 2004

Date _____   _____
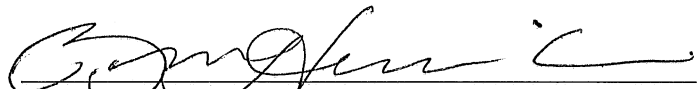   Melvin J. Luthy
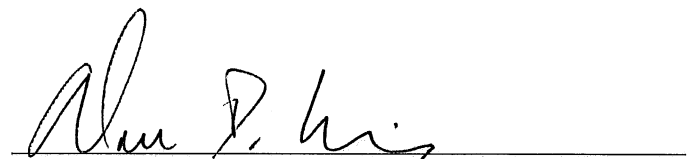
BRIGHAM YOUNG UNIVERSITY

As chair of the candidate's graduate committee, I have read the project of Justin Reed Shewell in its final form and have found that (1) its format, citations, and bibliographical style are consistent and acceptable and fulfill university and department style requirements; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the graduate committee and is ready for submission to the university library.

28 June 2004
Date

Lynn E. Henrichsen
Chair, Graduate Committee

Accepted for the Department

Alan D. Manning
Graduate Coordinator

Accepted for the College

Van C. Gessel
Dean, College of Humanities

ABSTRACT

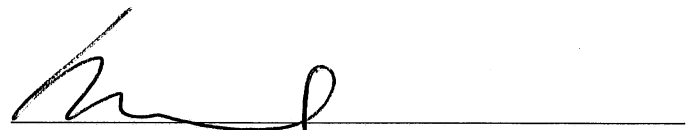HEARING THE DIFFERENCE: A COMPUTER-BASED

SPEECH-PERCEPTION DIAGNOSTIC TOOL FOR

NON-NATIVE SPEAKERS OF ENGLISH

Justin Reed Shewell

Department of Linguistics and English Language

Master of Arts

This project was completed to fill a need in the field of pronunciation teaching and learning by providing a computer-based, speech-perception diagnostic tool that helps determine learners' problem areas in the perception of English speech. Current diagnostic tools are few and very limited in their scope and application in the language classroom. The *Perception of Spoken English Test* diagnoses learners' specific speech perception problems, alerting teachers to areas that require special attention in a particular course or lesson. This project involved the development, production, piloting, evaluation, and revision of a computer-based instrument in an intensive English program. The data collected from the pilot experience led to several adjustments and improvements in the instrument, resulting in the version presented herein.

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

"English is now the dominant or official language in over 60 countries, and is represented in every continent….Most of the scientific, technological and academic information in the world is expressed in English and over 80% of all the information stored in electronic retrieval systems is in English." (Crystal, 1997, p. 106). As English becomes a lingua franca in science, business, and other fields, effective communication in English becomes increasingly important to millions of people.

Intelligible pronunciation and listening comprehension skills are a major factor in effective communication (Celce-Murcia, Brinton & Goodwin, 1996; Fayer & Krasinski, 1987; Morley, 1991).The teaching of pronunciation in language classes has had varied importance at different stages in history (Kelly, 1969). While accurate pronunciation was once considered an essential skill for any language learner, during the 1950s and 1960s pronunciation instruction began to be regarded as insignificant in the overall acquisition of a foreign language because many believed that native-like pronunciation ability was unobtainable (Scovel, 1969). However, as teachers and learners realized the importance of clear, understandable pronunciation in effective communication, pronunciation training once again found a place in English language programs (Celce-Murcia, Brinton, & Goodwin, 1996; Richards and Rogers, 2001).

However, many language teachers still overlook pronunciation in their language instruction goals for a number of reasons. One reason is that many teachers and professionals assume that this skill is one that cannot be taught, but must be learned through practice (Morley, 1994). However, research cited in Chapter 2 of this report

shows that the speech perception and production skills of adult and children learners alike can be improved through training (Borden, Gerber & Milsark, 1983; Cenoz & Lecumberri, 1999; Flege, 1995; Jamieson & Morosan, 1986; Strange & Dittmann, 1984; Underbakke, 1993). Other studies citied in this report show that speech perception plays an important role in the acquisition of speech production skills (Chan, 2001; Flege & Eefting, 1987; Ingram & Park, 1997; Schneiderman, Bourdages, & Champagne, 1988). Furthermore, teachers and researchers in the field point out that if students are to improve their speech production, they must take responsibility for their own mistakes (Acton, 1984; Morley, 1991).

Thus, if perception plays an important role in speech production, and speech perception can be improved through training, then it is helpful for language teachers to be aware of specific problem areas in their learners' speech perception abilities. By knowing these specific problems, teachers may help learners more effectively communicate by helping them overcome these difficulties in their speech perception. Learners need to be aware of problem areas in their own speech perception so they may better recognize mistakes in their production of English sounds and suprasegmental patterns, and thus take responsibility for correcting these mistakes. Awareness of speech perception problems may also help learners improve their listening comprehension skills.

Some teachers rely on contrastive analysis theory to indicate problems their learners may face in acquiring L2 phonology. Contrastive analysis is the process of analyzing a learner's native language and identifying specific features and characteristics that may interfere with the acquisition of similar characteristics of the target language. Other teachers rely on their own experience with a particular student to alert them of

problem areas. Both of these methods do not provide adequate evidence of specific, individual learner problems. Many teachers and learners turn to commercially available materials to aid them in diagnosing problems in speech perception. Chapter 3 investigates materials currently available for diagnosing speech perception problems. Most of these materials are inadequate and do not provide a clear picture of learners' specific problems. Those materials that do strive to provide a clear picture of learners' problems in speech perception are difficult to administer, requiring much teacher involvement and preparation, as well as a significant amount of contact time with learners.

This project sought to fill the need of teachers, both native and non-native speakers of English alike, and learners in accurately diagnosing speech perception problems. The term "project" refers to the design, production, piloting, evaluation, and revision of the *Perception of Spoken English* (POSE) *Test*. The POSE (pronounced /powz/) test is unique in many ways, as outlined in Chapter 4. First, the POSE test seeks to provide a clearer picture of speech perception problems by incorporating a large number of items for each aspect of speech perception included in the test: vowels, consonants, word stress, intonation, and sentence stress. Second, the POSE test is computer-based, allowing learners to complete the different sections outside of the classroom, thus eliminating the need for a teacher to be present during the administration of the test. Learners can also complete different sections of the test at their own leisure. The POSE test is delivered via the Internet, allowing more people and language programs access at nominal cost and requiring no more special equipment than a computer and an Internet connection.

The methods used during the pilot stage, and the participants involved, are outlined in Chapter 5. The POSE test was piloted at the English Language Center at Brigham Young University. The POSE test was split into two forms to accommodate more students and require less time during the pilot stage. Students from several different language background participated in the pilot of the POSE test, and their responses were recorded by the computer.

The data collected during the pilot stage of the project were analyzed and evaluated to determine the reliability and validity of the POSE test. The results of these evaluations are included in Chapter 6. Unfortunately, the data collected were insufficient to provide significant reliability data. However, several trends in the data suggested that the POSE test was reliable to a degree. The validity of the POSE test was difficult to determine, due in part to the uniqueness of this product and the lack of comparable, established measures. Two popular sources of contrastive analysis data were examined and compared to the data collected during the pilot stage of the project. While significant numbers and data were not available, several trends in the data suggested that the POSE test was, to some extent, a valid instrument for diagnosing speech perception problems. Future plans for the POSE test include further research in both the reliability and validity of this instrument.

After the completion of the pilot stage, and the analysis of the collected data, several changes and adjustments were made to the POSE test. These changes are explained in Chapter 7. The changes were made based on data and feedback from participants and objective observers. Additionally, some aspects of speech perception were outside the scope of the POSE test, and so were not incorporated. Some other

aspects of speech perception were not included due to limited time and resources. These limitations and delimitations are also discussed in Chapter 7. Finally, there are several additional changes and adjustments that will hopefully be incorporated in future versions of the POSE test, as well as plans for further research to help determine the reliability and validity of this instrument. Future research plans also include using the POSE test as an instrument in studying other aspects of speech perception and production. These plans for future research and revisions to the POSE test are outlined in Chapter 7 of this report.

The POSE test fills an important need in the area of diagnosing speech perception problems. The POSE test may be used by teachers to aid in the design of specific instruction aimed at overcoming speech perception and production problems. Learners may use the results of the POSE test to help raise awareness of their own speech perception problems, thus enabling them to improve their listening comprehension and self-monitoring abilities. The POSE test may also be used as a possible research tool to help provide more empirical evidence in the areas of speech perception, production, and contrastive analysis. Thus, the POSE test is an important tool for many different language situations, and may help teachers and learners in their quest for effective communication in English.

CHAPTER 2

A Review of the Literature

The literature review below looks at the history of pronunciation teaching and research and the importance of pronunciation instruction in the language teaching arena. The role of perception as it relates to pronunciation teaching and research is also examined. Finally, this review examines the teachability of perception and production of L2 sounds.

It should be noted that the focus of this review is the production and perception of speech in English as a second or foreign language. Some of the research reports presented below use other languages to examine pronunciation, but the principles in each can be applied to the instruction of English as well. Therefore, in the review that follows, it is assumed that all comments and references to pronunciation instruction, pronunciation, perception, and language programs refer to the teaching of English.

The History of Pronunciation Teaching and Research

Pronunciation instruction has long been a component of language teaching. Its place in the curriculum and prominence in research have varied however. In his book *25 Centuries of Language Teaching*, Louis G. Kelly (1969) offers evidence that Sanskrit linguists were aware of pronunciation as an aspect of language as far back as 1000 B.C. Kelly also mentions that many language texts and materials in the classic and medieval periods provided some type of phonetic analysis of the language being discussed. The focus on phonetics in these texts was mainly on imitation of sounds produced by native

speakers of the target language, which continued into the late 1800s and early 1900s with the Direct Method in Europe (Celce-Murcia, Brinton & Goodwin, 1996; Kelly, 1969).

In Asia, the Korean monarch, King Sejong the Great, commissioned scholars of his day to produce a unique alphabet whose letters reflect the positions of the different components of the speech mechanism (i.e. the tongue, teeth, etc.) in relation to the phonology of Korean. This alphabet was completed as early as 1446 (The Sigma Institute, n.d.), however little information is available as to how or when pronunciation was taught in this case.

In 1886, the International Phonetic Association was formed in Europe, which emphasized the teaching of the spoken form of the target language (Celce-Murcia, Brinton & Goodwin, 1996). Howatt (1984) reasons that this new focus on the spoken form of a language eventually led to the development of the Audiolingual and Oral Methods, as well as Situational Language Teaching, widely accepted in the 1940s, 1950s, and early 1960s, when pronunciation instruction was considered an essential part of almost every language program. Indeed, accurate pronunciation was a high-priority in both U.S. and British-based language programs (Morley, 1991) and Charles Fries (1945) felt pronunciation important enough to devote an entire chapter of his book *Teaching & Learning English as a Foreign Language* to teaching the sounds of English, noticeably before teaching any other component. These methods were based mainly on structural linguistics and held to the theorem of contrastive analysis (Robinett & Schacter, 1983; Wardhaugh, 1970), which stated that careful analysis of the learners' native languages and the target language would indicate potential problems in the language learning process, and that by knowing these problems, a teacher would be able to better facilitate

language learning (Richards & Rogers, 2001). Richards and Rogers cite a text produced by the American Council of Learned Societies (1952) whose title indicates the main philosophy of the day: *Structural Notes and Corpus: A Basis for the Preparation of Materials to Teach English as a Foreign Language*.

However, with the introduction of transformational-generative grammar (Chomsky, 1957, 1965) and cognitive psychology (Neisser, 1967), many people began to question this structural approach to language teaching. Researchers and language teachers began to look at language as more than simple habit-formation (Celce-Murcia, Brinton & Goodwin, 1996). Language programs began to base their curriculums and instruction on two new philosophies. The first one was that grammar and vocabulary were the most important components of language acquisition, and therefore should make up the bulk of instruction (Celce-Murcia, Brinton & Goodwin, 1996). The second philosophy was that native-like pronunciation was an unobtainable goal (Scovel, 1969), and if a learner could not learn to pronounce the language like a native speaker why bother wasting valuable instruction on this aspect of language acquisition. These two new ideas led to the reduction, or complete elimination in some cases, of pronunciation instruction in language programs around the world (Celce-Murcia, Brinton & Goodwin, 1996; Morley, 1991). This view lasted until the early 1980s, when the communicative approach began to gain prevalence in the language teaching arena (Celce-Murcia, Brinton & Goodwin, 1996; Richards & Rogers, 2001), and researchers once again began to study pronunciation and its place in the language curriculum.

This time, however, the role pronunciation played in overall language instruction was different. In the early 1940s, 1950s, and 1960s pronunciation was seen to be only a

component of language acquisition in a structural sense, and all instruction was based on phonemic contrasts. Fries (1945) devotes a good portion of his chapter on English sounds to the phonetic alphabet in an effort to help learners identify and produce the individual phonemes of the English language, as well as a few pages covering intonation and stress patterns. In the late 1980s pronunciation instruction moved beyond the structural aspects of pronunciation and began to focus on the role of pronunciation in overall communication, holding to the basic philosophy that "*intelligible pronunciation was an essential component in communicative competence*" (Morley, 1991, italics in original).

This new stance continues to thrive today among researchers. It is less common among teachers, however, as Morley (1994) points out. She says that teachers avoid spending time on pronunciation instruction due to several common myths about pronunciation, three of which are mentioned below:

1) Pronunciation is not important

2) Students will acquire the skills on their own

3) Pronunciation cannot be taught

The next segment of this literature review will look at each of these myths in turn and provide literature evidence that they are unfounded.

<center>The Importance of Pronunciation Instruction</center>

Pronunciation is important in language acquisition. One area where it can have the most influence is in the affective domain. If proper pronunciation is a problem, the fear the learner has of being misunderstood, or encountering awkward situations due to poor pronunciation may then influence how aggressive the learner is in seeking out

opportunities to use the target language. Flege (1995), citing studies from Lane (1963), Gumperz (1982), Fayer and Krasinski (1987), and Holden and Hogan (1993), observes that poor pronunciation may make it difficult for learners to be understood, especially when listening conditions are not ideal. Poor pronunciation may also encourage native speakers to judge learners' mental and affective state, or cause negative personal evaluations to be passed (Flege, 1995). Morley (1994) asserts that learners with poor pronunciation have long-term difficulties in becoming effective oral communicators, and that some learners in this group never reach this level of acquisition.

Focusing on the learners themselves, Morley (1991) proposes several learner groups to whom poor pronunciation might pose real problems, and thus may require specific attention in the area of pronunciation instruction. These groups are:

1) Adult and teenage refugees in vocational and language training programs.

2) Immigrant residents in English-speaking countries.

3) Nonnative speakers of English in technology, business, and industry.

4) College and university faculty members and researchers.

5) Graduate and undergraduate students in programs in English-speaking countries, and those wishing to enter such programs.

Celce-Murcia, Brinton, and Goodwin (1996), in their book on teaching pronunciation, add two more groups to this list:

1) Non-native speakers of English who teach English

2) Those who work with English speakers in foreign countries (such as tour guides and translators, for example).

Most learners fall into one of these seven groups. Thus, pronunciation should be considered an important part of most language teaching situations.

<div align="center">Factors that Influence Pronunciation Acquisition</div>

Like any other aspect of language acquisition, the level of pronunciation any one learner attains may be influenced by many different factors. Some of the more prominent influences include the age of the learner and the native language background from which the learner comes (Celce-Murcia, Brinton & Goodwin, 1996; Pennington, 1998). The next section of this review examines these two factors.

<div align="center">*Age of the Learner*</div>

It has long been noted, and is commonly believed today that children have the innate ability to learn languages, and specifically acquire pronunciation without the presence of a "foreign accent." Many have debated that the ability to acquire the phonetic system of a non-native language disappears, or is at least significantly reduced, after the age of puberty. Lenneberg (1967) proposed the Critical Period Hypothesis, which argues that there is a period where the functions of the brain with regard to language and language learning are assigned to certain sections within the brain. This is known as "lateralization" (p. 150). According to Lenneberg, after this period, called the "critical period" (p. 175), reassigning these functions becomes increasingly difficult. Krashen (1973) goes one step further by arguing that along with the lateralization of the brain comes a reduction of "cerebral plasticity" (p. 67), or the ability to add new information and functions to the ones already in place, making it impossible for any learner to achieve

native-like pronunciation in a second language. Seliger, Krashen, and Ladefoged (1975) reported that "puberty may be an important turning point in language learning ability" (p. 21), stating that the level of pronunciation obtainable by adult learners is limited at best.

Several inconsistencies in the study done by Seliger, Krashen, and Ladefoged (1975) must be accounted for, however. The study was conducted using linguistics students who surveyed three adult immigrants each as part of a class assignment. The students were assigned to ask each immigrant how accented his/her speech was perceived to be by native speakers of his/her target language (either English or Hebrew). No control was made for how opinionated the immigrants might be toward their own speech and no attempt was made to determine the accuracy of the reports of the immigrants in accordance with actual determinations made by native speakers. Furthermore, generalizations were made based on the survey responses and their correlation to the results of previous studies.

As could be expected, not all researchers completely support the idea of a "critical period." Flege (1995) cites data obtained when native speakers of English rated native Italian (NI) speakers of English according to how accented their speech was. The age at which the NI participants began learning English, or age of learning (AOL), ranged from 3 years to 21 years of age. All of the NI participants learned English while living in Canada. The correlation between the AOL and perceived accent by native speakers was strong ($r = 0.71$), indicating that the later the NI participants began learning English, the more accented their speech was perceived to be by native English speakers. Proponents of the critical period hypothesis claim that ability to change pronunciation habits greatly diminishes around puberty. Thus, the data should show stronger correlations for

participants who began learning English before puberty, and much weaker correlations for those participants who began learning English after puberty. However, the correlation data from this study show no such strengths or weakness, but instead show similar correlations for all participants. Flege concludes his examination of these data by saying, "If a critical period exists, it apparently does not result in a sharp discontinuity in L2 pronunciation ability at around puberty" (p. 234).

Kuhl and Iverson (1995) discuss the existence of a "perceptual magnet effect" (p. 121; see also Kuhl, 1991) wherein exposure to a particular language alters the phonetic perception of adults and infants alike. The data the researchers draw their conclusions from involved synthesized vowels in isolation, and participants were asked to respond when the stimuli presented were different from the ones preceding them in the sequence of sounds. The stimuli themselves varied only slightly from each other. In all practicality, this type of research situation does not even closely resemble a speaking environment in which most non-native speakers of English might find themselves. However, the data do suggest that the perceptual mindset of adults was influenced by exposure to the stimuli, indicating that adults possess some ability to change the way they perceive sound. Pennington (1998) supports this idea and suggests that while acquiring new skills in perception and production is difficult for adult learners, it is not altogether impossible, and that "the acquisition of phonology beyond childhood is a gradual and extended process" (p. 338). This idea is also supported by other researchers in the field (e.g. Jusczyk, 1993; Pennington, 1993).

Native Language Background

We acquire the sound system of our native language as children. Some researchers (Flege, 1987; Leather & James, 1991) suggest that this acquisition process involves the formation of phonemic categories or classes. Children generally are able to add to existing classes fairly easily because their "central representations for sound categories are still evolving" (Flege, 1987, p. 172). On the other hand, when an adult learner sets out to acquire a second language, the acquisition of the sound system is often influenced by the phonemic categories or classes previously established (Flege, 1987; Leather & James 1991). Werker and Polka (1993) state that infants have the ability to perceive almost all sounds in any language, but that this ability is diminished by the time the learner reaches adulthood as a result of having learned a particular language. This indicates the sound system of the target language is either processed according to existing phonemic classes, or new categories are formed based on the differences of the two sound systems (Best, 1995; Flege, 1995). Thus, as emphasized by contrastive analysis theory (Robinett & Schacter, 1983; Wardhaugh, 1970) and error analysis theory (Banathy & Madarasz, 1969; Schachter, 1986), the problems a learner has in acquiring accurate pronunciation of the L2 may be based in part on the native language background of that learner.

The Role of Perception in Pronunciation

Many researchers have noted that there is a relationship between what a learner can hear, and what sounds a learner can produce. Speech production is a process, and not a product, made up of four different phases—perception, programming, processing, and

execution (Ferguson & Macken, 1980; Hewlett, 1990). Much of the research in pronunciation has focused on the execution phase of speech production, but there is a fair amount written on the relationship between perception and speech production.

Flege and Eefting (1987) examined the relationship between perception and production by testing the perceptual performance of two groups of learners. The first group was made up of English monolinguals between the ages of 9 and 10. The second group consisted of native speakers of Spanish of the same age who began learning English between the ages of 5 and 6. Each group of subjects participated in a perceptual decision task that required them to decide if the stimulus presented was English /dɑ/ or /tɑ/. Results from this study indicated that the some of the Spanish speakers had poorer production performance when compared with the English group, while the others in the Spanish group did not differ from the English-speaking group in their production of these sounds. The entire group of Spanish-speaking children did not differ from the English-speaking children in their identification of the English sounds. The authors then concluded that as the group with poorer performance in production had more experience with English, they would produce these stops authentically, noting indications in their data that accurate perception of English /da/ and /ta/ lead to accurate production.

The relationship between perceptual training and production was also examined by Schneiderman, Bourdages, and Champagne (1988). These researchers utilized L2 learners of French at beginning and low-intermediate levels of ability. The participants were from several different language backgrounds, including English, Chinese, Tamil, Hindi, Turkish, Spanish and German, and were between the ages of 18 and 60. The participants were divided into two groups. One group received specific perceptual

training, while the other group spent their time doing "listening comprehension exercises" (p. 7). Each group was given a perception and production test before and after this training period. For the perception test, participants were asked to determine whether the stimuli presented were the same or different. The first group of stimuli was made up of pairs of French words which were either identical or which differed by only one segment. The second group of stimuli consisted of sentences of the same or different intonation and rhythmic patterns. For the production test, subjects were tested on their ability to imitate French words and sentences. The researchers found that the group of participants who received specific perceptual training scored significantly higher on all post-training measures (except for the discrimination of rhythmic patterns) than the untrained group, indicating a relationship between ability to perceive and ability to produce sounds in the target language, at least on a segmental level. The researchers concluded that an improved ability to discriminate L2 sounds would result in more native-like production.

Supported by these early findings, two more important studies were conducted to examine the relationship between perception and production. Ingram and Park (1997), employing a carefully controlled research design, examined the perception of non-native vowels in Japanese and Korean learners of English. They had a total of five groups of participants. Two of the groups were lower-level learners, having only been in Australia for less than 12 months. Both groups were in their twenties, and both groups used English everyday in their studies at Australian universities. Two more groups were made up of more experienced learners of English, having been in Australia for at least five years, and all but two of the members of these groups taught Japanese or Korean at Australian

universities. The participants in these groups were in their thirties. The remaining group was a control group made up of native Australian English speakers.

The groups were given a forced-choice perception test, using tape-recorded stimuli. The participants were asked to circle one of five English words according to what they heard. The words were presented in the form of /h_d/ where the blank was filled in with one of the following vowels: /i, ɪ, e, æ, ɑ/. The Korean participants confused the Australian /e/ - /æ/ contrast in the perception task much more so than did the Japanese participants. In the production task, while not able to produce the contrast with native-like accuracy, the Japanese participants were able to produce acoustically different sounds, while the Korean participants produced sounds that acoustically overlapped.

In a separate, but related experiment (reported in the same article), the Japanese participants classified the /e/ and /æ/ into separate native language phonetic categories, but the Korean participants classified them in the same category. This may be due in part to the differences between the vowel systems of Japanese and Korean. Japanese vowels can be either long or short in duration, the difference in duration being phonemic in Japanese. The Korean vowel system, on the other hand, which once held duration as a phonemic contrast, has lost this distinction in all but older speakers. Ingram and Park (1997) reported that the Japanese subjects, while not being able to discriminate the two sounds based on their vowel quality alone, were able to discriminate them based on the duration of the two vowel sounds. The Korean subjects (except the group of older speakers, to whom vowel duration was still a phonemic classifier) were not able to make this distinction, and thus classified the two vowel sounds into the same category. While these findings provide empirical evidence that native language background effects the

perception and production of L2 sounds, they also provide evidence that perception and production are related.

In another study (Chan, 2001), native Cantonese L2 learners of English were divided into two groups of 30 participants each. The first group consisted of learners who consistently mispronounced all English words with word-initial consonants /v, θ, ð, z, r/. The group was tested three times (non-consecutively). The second group was comprised of Cantonese speakers who consistently pronounced the same words correctly in the same production test. Chan controlled for age, length of English learning experience, English educational level, and hearing and oro[1]-motor function. When Chan compared the perception scores of the two groups, she found that those in the first group (those who consistently mispronounced the tested words) had significantly poorer perception scores than those in the other study group, suggesting that perception performance positively relates to production performance.

In looking at the evidence presented above, it must be noted that in most of the perception/production studies participants were asked only to determine if the stimuli presented were either the same or different from each other, or if the stimulus presented was a certain English sound. As most communication does not involve phonemes in isolation, it is difficult to generalize the findings of these studies to the acquisition of native like fluency in oral communication. However, the findings of these studies do provide some evidence as to the positive relationship between perception and production. If this is true, then the question is, what can be done to help improve the perception (and in turn, the production abilities) of L2 learners of English.

---

[1] CancerWEB's On-line Medical Dictionary defines *oro* as a prefix relating to the mouth. (Retrieved May 24, 2004, from http://cancerweb.ncl.ac.uk/cgi-bin/omd?query=Oro&action=Search+OMD).

Morley (1991) and Acton (1984) stress the importance of developing self-monitoring skills in L2 learners. A significant part of self-monitoring includes recognizing mistakes in one's own pronunciation. It is self-evident that accurate speech perception plays a significant role in this important skill. Training in speech perception helps to heighten the awareness of L2 learners to their own mistakes and thus may help learner in accepting responsibility for these mistakes and eventually overcoming these problems in effective communication. Thus the teachability of perception becomes an important issue in this review.

The Teachability of Perception and Production

Research on the teachability of pronunciation is rare; however a few studies do exist. Previous to 1980, some people believed that "pronunciation instruction had no effect on the acquisition of phonology" (Pennington, 1998, p. 325). The study most widely cited to support this claim was conducted by Suter (1976). Suter surveyed 61 non-native speakers of English in an attempt to discover what features of their background most influenced their pronunciation. First, Suter questioned them on their language learning background, including the amount of time they had spent specifically studying pronunciation. Then the pronunciation of the 61 participants was assessed by native-speaker judges.

Suter (1976) then examined the data to find correlations between the background information the participants had given and the pronunciation scores they had received. Suter found that the number of weeks the participants had spent in formal pronunciation training had no influence on how accurate their pronunciation was perceived to be. Suter

listed twelve of the variables examined that had high correlations with the pronunciation accuracy of the participants. Purcell and Suter (1981) followed up with another similar study that examined these twelve variables more closely and found that only four of the twelve were meaningful predictors of pronunciation accuracy. Once again, the amount of formal pronunciation training was not among those listed.

Several problems may be noted in this study, however. First, Suter (1976) used overall pronunciation accuracy (this includes segmental and suprasegmental accuracy) as a measure of pronunciation ability. Pennington (1998) observes that a learner may have accurate pronunciation at the segmental level, but still be poor in oral communication and overall fluency. While Suter instructed his native-speaker judges to rate on both the accuracy of sounds and of rhythm, stress, and intonation, it is almost impossible to say which influenced the ratings more. Some participants may have been rated high because they had good pronunciation in terms of intonation and rhythm, but those skills in suprasegmentals may have masked their poor ability in accurately pronouncing segmentals. On the other hand, some participants may have been rated low because they did not accurately produce English rhythms and intonation, but may have been able to accurately produce English segmentals.

Pennington (1998) also suggests that it is difficult for native speakers to separate accuracy from the myriad of oral language features that influence evaluation of oral performance, and that ratings of each feature (even after lengthy rater training sessions) tend to correlate very strongly with each other, suggesting that raters have a difficult time determining the difference between such features as accuracy, fluency, correct articulation of vowels and consonants, intonation, and stress and rhythm.

Another problem with the study concerns the type of information obtained from the participants. Suter (1976) questioned his participants as to the amount (number of weeks) of training they had received in pronunciation, but never reported the types of methods and strategies that training employed. In at least some of the cases, the type of training received would have affected the acquisition of pronunciation (Derwing, Munro, & Wiebe, 1998).

Lastly, Suter (1976) examined the relationship of the variables mentioned in the study with the pronunciation accuracy of the participants by looking at the correlation coefficients of each of the variables and the pronunciation scores. Correlation in and of itself does not provide evidence as to the cause of the relationship. The fact that the amount of overall English training and the pronunciation scores of the participants do not correlate significantly does not provide evidence that pronunciation training is ineffective. Indeed, several other studies have shown that training does have an effect on the acquisition of L2 phonology (e.g. Cenoz & Lecumberri, 1999; Champagne-Muzar, Schneiderman & Bourdages, 1993; de Bot, 1983).

To examine the effect of perception and production training on Korean speakers of English, Borden, Gerber, and Milsark (1983) used a total of ten participants between the ages of 19 and 48. All of the participants had been in the United States for no longer than three years, and each of the participants varied in their use of English outside of their ESL courses at Temple University (Philadelphia, PA). Participants were tested in English speech production, identification, AX discrimination, and self-perception of the /r/ - /l/ contrast. While the results showed no long-term effect of training on pronunciation acquisition, they did show that those learners who scored lower on the production

measures indicated the most carryover from the training in post-training measures. Those who had higher production scores showed no significant improvement, however, suggesting that there is a point at which higher-level learners resist change in speech patterns. Borden et al. also concluded that there was a strong relationship between the ability to perceive the /r/ - /l/ contrast accurately, and being able to accurately produce those same sounds without confusion.

Strange and Dittmann (1984) looked at the influence of training on the acquisition of L2 phonology in native Japanese speakers of English. As with Borden et al. (1983), they focused on the /r/ - /l/ contrast, as this contrast seems to be difficult for all Japanese learners of English to perceive and produce as it is not present in this language (Goto, 1971; Miyawaki et al., 1975). Participants for this study were eight Japanese females 25 to 33 years old, who had lived in the United States from 5 to 30 months. English experience and ability varied for each subject, but all subjects had difficulty in differentiating /r/ from /l/. In comparing pre- and post-training scores, all eight participants showed improvement in both perception and production regardless of their level of perception prior to training. The training in this study involved asking the participants to distinguish between sets of stimuli, determining if they were the same or different. Immediate feedback was given at the end of each response.

Another similar study (Jamieson & Morosan, 1986) examined the effects of training on French speakers who had difficulty perceiving and producing the /ð/ - /θ/ contrast in English. For this study, ten male and ten female subjects, ages 18 to 32, were chosen because they had scored below the 50th percentile on the English placement test offered by the institution. The participants were given a pretest, and then randomly

assigned to either a control group or a training group so that each group had subjects with an equal range of pretest scores. Participants in the training group were then given specific training on the contrasting sounds. As in the previously mentioned study (Strange & Dittmann, 1984), this training involved having the participants distinguish between sets of stimuli, determining if they were the same or different. While Strange and Dittmann used individual sounds, however, Jamieson and Morosan used minimal pair syllables (CV). Results from this study indicated that training had a positive effect on participants' ability to identify and discriminate between the two sounds.

Underbakke (1993) also looked at the pronunciation of /r/ and /l/ by native Japanese speakers of English. This study used 39 students (17 males and 22 females) at the English Language Institute at the University of South Florida. Most of the students had spent less than three months in the United States. Students were given a pretest and then training on identification of the initial sounds of the stimuli. The training included a familiarization and practice section, two 60-item blocks of trials which included feedback, one 60-item block of trials with no feedback, and then a 60-item identification test. After training, participants were given a post-test on their ability to identify the specific /r/ - /l/ contrast. Underbakke reports that the treatment group showed significant improvement on all measures. The control group, which had received training on the /b/ - /v/ contrast, also showed improvement; however, it was not as significant as the treatment group. Underbakke accredits this to the fact that the control group may have learned something from taking the tests, and the fact that the control group also focused on initial sound contrasts in their training sessions, although it was a different contrast than that focused on by the treatment group.

While the data from these studies support the premise that training positively affects performance in the perception and production of English sounds, these studies focused on single phonemic contrast in English: /l/ - /r/ (Strange & Dittmann, 1984; Underbakke, 1993) and /ð/ - /θ/ (Jamieson & Morosan, 1986). Therefore, it is difficult to generalize these findings and say that all training in perception will improve performance. Also, the tasks in these studies consisted of sets of isolated stimuli, something that rarely occurs in real-world language situations.

A different study (Cenoz & Lecumberri, 1999) examined the effect of training on the perception of eleven English vowels and eight diphthongs. Participants in this study were 109 university students (mostly female), with most being between the ages of 18 and 21 years old. All of the participants were in their first year of English studies at the university. An interesting thing to note with this study is that 67 percent of the participants had visited an English-speaking country at least once, and all had a desire to improve their English pronunciation. Participants were asked to complete several questionnaires regarding their background, motivation, and English proficiency. Then participants were given a pre-training and post-training aural discrimination test focusing on the discrimination of all the sounds in the RP vowel system, except schwa (/ə/). The training was included as part of a course the participants were enrolled in and focused on the theoretical description of English sounds, as well as 14 hours of aural discrimination training on vowels, diphthongs, and consonants. All stimuli represented British English pronunciation.

The researchers report that participants showed significant improvement after only a few hours of training, and that the overall improvement was significant as well.

They mention in their report that the sounds which were not easily identifiable by participants in the pre-training discrimination tests showed the most improvement, while sounds that were easily distinguishable showed no major improvement on the post-training measures. This supports Flege's (1995) Speech Learning Model.

As has been mentioned before, the studies above focused on specific phonemic contrasts found in English, and perception training and testing was done in isolation of other oral communication skills. Also, the training programs employed in the above studies usually involved some type of audio and visual feedback, for example, visual representations of intonation contours. As mentioned before, most oral communication (a prime goal of communicative language instruction) does not involve recognizing and producing sounds or prosodic features in isolation; however it is reasonable to expect that some of the relationships and predictabilities that occur in oral communication would filter into the tasks participants were expected to do in the studies mentioned above, and thus the results of these studies may be cautiously applied to communicative language instruction in the real world.

Conclusion

While researchers' conclusions vary as to the amount of training, and the type of training that should be used, specific training on production and perception of English sounds can have a positive effect on the acquisition of English phonology by L2 learners. Many of the difficulties in acquisition may arise from the learner's native language background, as well as the age at which the learner began learning English. However, it is

agreed that, with training, adult learners can acquire some degree of fluency in the phonology of their target language.

The role of perception in the acquisition of L2 phonology is important, as supported by the literature showing that training and awareness of problems in perception of English sounds has a positive effect on both the perception and the production of these sounds by L2 learners of English. Thus, knowing the problems a learner has in the perception of English sounds and suprasegmental patterns would be helpful for that learner in obtaining a more native-like level of pronunciation and listening comprehension. This is an important goal of this project.

CHAPTER 3

Purpose and Rationale for the Project

The literature review in Chapter 2 provides evidence that awareness of problems in speech perception can lead to a more native-like level of pronunciation and listening comprehension. Some researchers and language teachers (Acton, 1984; Morley, 1991) stress the importance of self-monitoring in pronunciation instruction. Self-monitoring is the process of noticing specific errors in one's own production of the target language. In many cases, this process leads to correction of these errors and improvement in language skills. The importance of speech perception in the self-monitoring of speech production, as well as in the comprehension of speech in the target language is self-evident. Hence, problems in speech perception contribute to problems in other areas of language acquisition, i.e. speech production and communicative competence. Therefore, knowing what problems exist, and overcoming these problems through training (see Cenoz & Lecumberri, 1999; Pennington, 1998; Underbakke, 1993), can help to improve the acquisition of L2 phonology.

The *Perception of Spoken English Test*, or POSE test (pronounced /powz/), helps learners and teachers be aware of specific problems in speech perception. Learners can use this knowledge of problem areas to increase their ability to self-monitor their own speech production. Teachers can use this knowledge of problem areas to guide their instruction on specific sounds and suprasegmental patterns. Thus, the POSE test is an important tool for any teacher or learner concerned about effective communication. The POSE test is not the only method for diagnosing speech perception problems, however.

Currently, several different methods are used to help discover the specific problems learners have in speech perception. Some teachers and learners rely on contrastive analysis to help them determine possible problems in the acquisition of L2 phonology. Contrastive analysis is the process of analyzing a learner's native language and comparing that language to the target language. Teachers focus on the specific features of the target language to be taught, but use the knowledge of the native-language features to identify and explain problems in the acquisition of target language features. Thus, all learners from a particular language group are considered alike. However, not all learners are alike, and even though learners may be from the same language group, they will exhibit different problems in language acquisition. This practice of comparing the features of two languages creates a stereotype the learner is then thrust into without regard to these individual problems. This may not be the most effective method for determining problem areas in speech perception.

Other teachers and learners rely on their own intuition and experiences in the target language to guide them in determining problems in speech perception. But this raises two major concerns. First, most teachers and learners usually have access to data that reflects production abilities; however, though perception does play an important role in the acquisition of speech production skills, the exact role it plays is complex and difficult for many researchers to define. One author says, "…there is a complex link between production and perception of L2 sounds. Although it seems that perception in general might precede production, direct inferences about pronunciation accuracy cannot…be made from perceptual abilities in a straightforward manner" (Llisterri, 1995, The Production and Perception of L2 Vowels section, para. 14). Thus, using these data as

an indicator of learners' perception abilities can be misleading and inaccurate if not done with these complexities in mind. Second, even though data reflecting perception abilities may be available, many teachers themselves are non-native speakers of English and may have speech perception problems of their own that interfere in accurately determining problems their learners are having. Thus, many teachers and learners turn to commercially available materials to aid them in diagnosing problems in speech perception.

Market Analysis

Currently, there are only a few materials available that help diagnose speech perception problems. The perception tasks used in the studies examined in the literature review (see Chapter 2) focused solely on the perception of segmentals. While these may shed some light on problem areas, they do not provide a clear picture of a learner's speech perception ability. Also, many of the tasks utilized in the aforementioned studies are very specific and difficult to adapt to the language classroom. As a result, these tasks are unsuitable in many instances for diagnosing speech perception problems.

Other materials do exist. Many pronunciation texts or instructional materials include some form of diagnostic or "pre-test" that helps teachers and learners determine what areas to focus on in their instruction or individual study. Some texts even provide materials that help students analyze their own speech and find problem areas in their production of English sounds and suprasegmentals. There are few texts, however, that provide any type of diagnostic material for speech perception.

One text that includes a perception diagnostic test is *Exercises in American English Pronunciation* (Sudlow & Bischof, 1994). This diagnostic, and the text in general, focus solely on segmentals. This test has a large number of items, and covers a variety of phonemic contrasts in both the vowel and consonant sections. However, by focusing only on segmentals, it does not provide an accurate picture of speech perception problems. Also, the test is designed to be read by the teacher during a class or other contact time with the students. This method can be daunting to many non-native English speaking teachers who may worry about their own pronunciation influencing the results of the test.

Another text that provides a speech perception diagnostic component is *Clear Speech* (Gilbert, 1993), which includes the *Clear Listening Test*. This diagnostic test focuses mainly on suprasegmentals. The first section is titled "Sounds," but the items contained in this section are mainly designed to detect the habitual addition or omission of consonants and/or vowels and do not highlight specific phonemic contrasts that may pose a problem for learners in communication settings. While some teachers may choose to focus on these general problems, other teachers and learners may desire knowledge about specific problems that this test cannot provide. Also, this test is designed to be administered during class time, and therefore is designed to require "about 20 minutes" (Gilbert, 1993, p. vii). While this may be practical, it does not provide a very clear picture of speech perception problems because the number of items that are available to test the different contrasts and suprasegmental patterns is very limited.

One text that seems to provide suitable diagnostic material in the perception of both segmentals and suprasegmentals is *Pronouncing American English* (Orion, 1997). This text is broken up into several units that examine specific vowel and consonant

contrasts as well as intonation, rhythm and stress patterns. Included with most units is a section entitled "Check Your Listening" that contains items students and teachers can use to help determine problem areas in listening comprehension and speech perception. This text has some disadvantages, however, that should be mentioned. First, the diagnostic items are broken up and placed in different sections of the text. This is inconvenient for teachers who would like to examine problems in several different contrasts at once (i.e. focus on all vowel contrasts, or on several consonant contrasts at the same time). Focusing on several contrasts at once in diagnosing problems is helpful in noticing patterns in the problems identified. Learners and teachers can then focus on these patterns when time does not permit focus on specific problem areas. Another disadvantage to the items in this text is that the contrasts are tested using minimal pair words. Phonemes do not usually occur in isolation, and thus placing the sounds being tested in combination with other sounds, as in words or syllables, is desirable. However, single words are also not usually heard in isolation during spoken language interactions. Thus, to better emulate the real-world speaking environment, sentences are more desirable than single words. Finally, while this text provides a large number of diagnostic items for each contrast, the items are in book form, and not computerized. Thus, administering this test would require several hours of teacher contact time with learners, a condition that is impractical in many language programs where resources are limited.

In terms of computer-based materials, there are many different programs available that are designed to help L2 learners of English improve their pronunciation; however, there are virtually no programs that provide diagnostic material designed to help discover

problems in speech perception. Thus, the POSE test fulfills an important need in the diagnosing of speech perception problems.

Special Features

In short, very few speech perception diagnostic instruments exist. The few in existence are often inadequate in giving teachers an accurate picture of learners' speech perception problems, or are difficult to administer in many language teaching situations. The POSE test is designed to help learners and teachers of English identify specific perceptual problem areas and allow teachers to focus specific training on those areas. Learners can use the POSE test to help increase their own awareness of their speech perception problems and may use the results to help guide their self-study of English perception and production. The POSE test differs from other available materials for diagnosing speech perception problems in that it features both segmentals and suprasegmentals, and is of sufficient length to provide a much clearer picture of individual problems.

The POSE test features five sections: 1) vowels, 2) consonants, 3) word stress, 4) intonation, and 5) sentence stress (sometimes called prominence). Each section consists of at least forty items. For each contrast or aspect of perception included in each section, there are at least two (and in most cases four or more) items that focus on that contrast or aspect. This allows for a more accurate diagnosis than a test with only a few items that focus on general areas instead of specific problems.

The POSE test is designed to be flexible in the way it is administered. Learners have the option of completing any number of sections in the POSE test during any one

administration or sitting. Thus, students with a lot of time may complete all the sections in one sitting. Students who have only one or two hours may complete one or two sections of the POSE test and then return at a later date to complete the remaining sections if so desired. This allows teachers and learners to adjust the POSE test to their needs regarding time and the specific sections of the POSE test they wish to use.

The POSE test is administered by a computer, so it is easily accessible to many students at any one time, and it can be administered outside the classroom, allowing teachers to use valuable class time helping students to overcome the perception problems from which they are suffering. Also, because the presence of the teacher is not required, individual learners may use the POSE test in their own self-study of English. This makes the POSE test as practical as many of the self-administered diagnostic tests in terms of teacher-time required without sacrificing accuracy and reliability. Administering the POSE test via computer also allows teachers who are non-native speakers of English themselves to accurately diagnose speech perception problems without having their own perception and/or production errors influence the diagnostic results.

The POSE test is currently distributed via the Internet, making it easily accessible to many programs and individuals that might not have the equipment required to utilize more sophisticated computer applications. Thus, many more users can access and utilize the POSE test without having to purchase expensive equipment. Another advantage of using the Internet as a distribution tool is that learners of English not enrolled in any specific language program can access the POSE test at home or from their office computer without having to buy expensive CD-ROMs or install other complicated programs.

Conclusion

The literature review in the preceding chapter (see Chapter 2) has shown that pronunciation instruction has found a new role in communicative language teaching. Perception is important to speech production, and perception (as well as production) can be improved through specific training. Learners have a need to identify problem areas in their perception abilities, thus allowing them to focus on these areas in their instruction and language practice. Currently diagnosing these problem areas has been left to the teacher or learner and the methods for this diagnosis have not been easy to implement in many situations.

The features and technology of the POSE test allow access to more people than may be possible with a single text or other instructional materials. The POSE test can be administered at home, in a computer lab at a language institution, or from an office computer during a lunch break. Thus, the POSE test has the potential to assist learners and teachers around the world in identifying speech perception problems, hopefully leading to increased communicative competence.

CHAPTER 4

A Description of the *Perception of Spoken English* (POSE) *Test*

The *Perception of Spoken English* (POSE) *Test* is a computer-based test that focuses on diagnosing problems in the perception of vowels, consonants, word stress, intonation, and sentence stress. Each section is described in detail below. Screenshots, or images of the way the test looks on a computer screen, are also provided. A list of the items contained in each section can be found in Appendices A through E.

The POSE test is designed to be flexible as to how and when students access the test. Students have the option of completing all sections at once, or completing each section individually and separate from the other sections. This feature allows students to be responsible for their study of perception. It also allows teachers to focus on different aspects of speech perception at different times. A teacher who chooses to focus only on the perception and production of vowels during the length of a course or class, for example, could have his/her students complete only the vowel section, leaving the other sections for individual study at a later date. This flexibility makes the POSE test adaptable to many different language teaching situations.

In each section of the POSE test, for each item, the learner listens to a recording of a native speaker. Then the learner is asked to choose from a set of responses the response that best matches the recording which he/she just heard. If the learner cannot distinguish between the possible responses, he/she can indicate this by selecting an option appropriately labeled "I don't know." The recordings in the POSE test reflect the pronunciation and contrasts found in North American English. The POSE test focuses only on North American English for two main reasons. First, there is an increasing

demand for accurate perception and production of North American English around the world. With the advent of the TOEFL, the TOEIC and other such instruments produced in the United States as a means of determining language ability and their weight in such matters as admission to U.S. colleges and institutions of higher learning, obtaining employment, etc., learners of English worldwide are anxious to improve their perception and listening comprehension skills in North American English versus other Englishes represented around the world. This can be seen in the number of programs springing up designed to help learners prepare for these language situations. Another factor in the use of North American English as the basis for the POSE test is the fact that the author and creator of the POSE test is a native speaker of North American English, and accessibility to native speakers of other Englishes was limited.

The Vowel and Consonant Sections

The vowel section consists of 38 items designed to diagnose problems in 11 different vowel contrasts (see Table 1). The consonant section consists of 84 items designed to diagnose problems in 25 different consonant contrasts (see Table 2 and Table 3). The specific vowel and consonant contrasts were chosen based on their *functional load* (Catford, 1987). This term refers to the "number of pairs of words in the lexicon that [each vowel or consonant contrast] serves to keep distinct" (p. 88).

The consonant section focuses solely on syllable-initial (see Table 2) and syllable-final (see Table 3) consonants for two reasons. First, no information or research was available on the functional load of syllable-medial contrasts. Second, the creation of several sets of minimal pair sentences that are authentic in terms of plausibility for each

medial contrast was nearly impossible. This may be supported by the fact that no information on the functional load of these contrasts was available.

In both the vowel and consonant sections, each item is presented in isolation, and the order in which the items are presented is randomized so the same test can be used any number of times with the same learner. Each item consists of a set of minimal pair sentences, as opposed to minimal pair words, syllables or isolated phonemes. Minimal pair sentences were used to create a diagnostic environment comparable to communication environments in the real world. Thus, a clearer and better picture of the learner's problems is obtainable and more effective communication can result.

In order to diminish the effects of reading ability and vocabulary knowledge on the results of these two sections, each item is presented with illustrations that highlight the difference between the two sentences in the minimal pair set. Similar methods have been used in other perception instruments as presented in the literature review (e.g. Borden, Gerber, & Milsark, 1983). For some of the items in both the vowel and consonant sections, the illustrations and text of the sentences were borrowed from *Pronunciation Matters* (Henrichsen, Green, Nishitani & Bagley, 1999). For the remaining items, original illustrations and sentences were used. These illustrations were rendered by Dr. Lynn E. Henrichsen, co-author of *Pronunciation Matters*. Efforts were made to maintain consistency of illustrative style throughout. Example items from the vowel and consonant sections can be seen in Figures 1 and 2.

Table 1. *Vowel Contrasts in the POSE Test*

Phonemic Contrast	Minimal Pair Example	Functional Load[a]	Number of Items
/i/ - /ɪ/	beet/bit	95%	4
/ɪ/ - /ɛ/	bit/bet	54%	4
/ɛ/ - /e/	bet/bait	53%	4
/ɛ/ - /æ/	bet/bat	51%	4
/æ/ - /ɑ/	cat/cot	76%	4
/ɑ/ - /ɑr/	cot/cart	31.5%	4
/ɑ/ - /ʌ/	cot/cut	65%	4
/ɑ/ - /ow/	cot/coat	-----[b]	4
/ʌ/ - /ɚ/	cut/curt	40%	2
/ʌ/ - /ʊ/	putt/put	9%	2
/ʊ/ - /uw/	pull/pool	7%	2

[a]Source: Catford, 1987, pp. 89-90. [b]This contrast was included because of its existence in *Pronunciation Matters* (Henrichsen, Green, Nishitani & Bagley, 1999) even though no functional load information was available.

Table 2. *Syllable-Initial Consonant Contrasts in the*

POSE Test

Phonemic Contrast	Minimal Pair Example	Functional Load[a]	Number of Items
/p/ - /b/	pill/bill	98%	4
/p/ - /f/	pan/fan	77%	4
/v/ - /b/	vote/boat	29%	4
/v/ - /w/	vet/wet	22%	4
/f/ - /v/	fan/van	23%	4
/f/ - /θ/	free/three	15%	4
/θ/ - /t/	thin/tin	18%	4
/θ/ - /s/	think/sink	21%	2
/ð/ - /d/	they/day	19%	2
/n/ - /l/	nap/lap	61%	4
/l/ - /r/	lice/rice	83%	4
/s/ - /ʃ/	sip/ship	53%	2
/ʃ/ - /tʃ/	shin/chin	26%	2
/tʃ/ - /dʒ/	choke/joke	19%	2
/dʒ/ - /y/	jail/Yale	20.5%	2
/k/ - /g/	coat/goat	50%	4

[a]Source: Catford, 1987, pp. 89-90.

Table 3. *Syllable-Final Consonant Contrasts in the*

POSE Test

Phonemic Contrast	Minimal Pair Example	Functional Load[a]	Number of Items
/p/ - /b/	cap/cab	14%	4
/p/ - /f/	cup/cuff	17%	4
/θ/ - /t/	bath/bat	27%	4
/θ/ - /s/	faith/face	17%	2
/t/ - /d/	cart/card	72%	4
/n/ - /l/	bone/bowl	75%	4
/s/ - /z/	ice/eyes	38%	4
/ʃ/ - /tʃ/	wash/watch	12%	2
/k/ - /g/	tack/tag	29%	4

[a]Source: Catford, 1987, pp. 89-90.

The Word Stress Section

This section of the POSE test contains forty words consisting of two to five syllables each. The items in this section were chosen to reflect the different parts of speech in English. In English, word stress can differ depending on the part of speech of any particular word. For example, the word *contest* can be a noun if the first syllable is stressed (i.e. [kán tɛst]), or a verb if the second syllable is stressed (i.e. [kən tést]).

Figure 1. *Example Item from the Vowel Section*

Figure 2. *Example Item from the Consonant Section*

Again, every effort was made to make the diagnostic environment as comparable to real-world communication environments as possible.

The items in this section differed from those of the other sections in that the words in each item were presented in isolation. Different parts of speech receive stress on different syllables. If a learner happened to know that nouns, for example, were stressed on the initial syllable in two-syllable words, and was presented with a sentence in which a noun was missing, the learner would be able to correctly guess the answer of that particular item without really being able to perceive which syllable was actually stressed in the recording. The items in this section were presented as single words and not in sentences so that this type of learner knowledge would not influence the results. An example item from this section can be seen in Figure 3.

The Intonation Section

The items in the intonation section focused on the intonation at the end of an utterance. However, this section consisted of two different types of items. The first twenty items are sentences that could either be questions or statements. Learners are presented with the sentence minus any ending punctuation on the screen, and hear the sentence in a recording. They are then asked to indicate whether the sentence they hear in the recording is a question (rising intonation) or a statement (falling intonation). Again, if they cannot distinguish the difference, they are allowed to select "I don't know." To help distinguish between the two choices, the question choice is presented with the image of a question mark (?), and the statement choice is presented with the image of a period (.). An example of this type of item can be seen in Figure 4.

Figure 3. *Example Item from the Word Stress Section*

Figure 4. *Example Item from the Intonation Section (Part 1)*

The last twenty items of the intonation section consist of sentences that end in tag questions (i.e. "That's a great idea, isn't it?"). The sentences end in either rising or falling intonation. The learners are presented with the sentence on the screen and listen to a recording of the sentence. They are then asked to determine, based on the intonation, if the speaker is "sure" (falling intonation) or "unsure" (rising intonation) about the answer he/she will receive in answer to the question. In other words, is the speaker looking for information ("unsure" about the answer), or making a comment ("sure" about the answer). If the learner is unable to tell the difference between the two choices, he/she is allowed to select the choice labeled "I don't know." To help avoid any misunderstandings in the possible choices for each item, the choice marked "sure" is presented with the image of an exclamation point (!), and the choice marked "unsure" is presented with the image of a question mark (?). An example item from this part of the intonation section can be seen in Figure 5.

It should be noted that some items for this section of the POSE test were taken from a similar unpublished test originally developed by Brent Green and Amber Pauga of Brigham Young University, Hawaii Campus. The remaining items were original items developed by the author.

The Sentence Stress Section

This section is divided into two parts. The first part focuses on the stressed word in a sentence. In English, stress is placed on content words, or words that carry meaning. Sometimes extra emphasis is placed on a specific word to indicate its importance in the

Figure 5. *Example Item from the Intonation Section (Part 2)*

meaning of the original utterance. For example, if a man and a woman were both standing next to each other, and you wanted to indicate the man in your utterance, you would stress the word *man* in the sentence to distinguish the man from the woman. In this first part of the sentence stress section, learners are presented with a set of minimal pair sentences with a different word underlined in each sentence. The underlined word indicates the stressed word in that sentence. Then the learners listen to a recording of a native English speaker and are asked to select which sentence they hear according to the stressed words. If a learner cannot distinguish between the two sentences, he/she can select the option labeled "I don't know." An example item from this part of the sentence stress section is shown in Figure 6.

The second part of this section focuses on "thought groups" (Gilbert, 1993, p.77). In this part of the POSE test, learners are presented with twenty minimal pair sentences in which the meaning of the sentence differs based on where the speaker pauses during the utterance. As with the first part of this section, learners are presented with both sentences and then listen to a recording of a native speaker. They are then asked to select which sentence the native speaker said. If they cannot distinguish between the possible choices, they are allowed to select the option labeled "I don't know." An example item from this part of the sentence stress section can be seen in Figure 7.

For each item in the sentence stress section, parenthetical phrases, called "rejoinders" (Henrichsen, Green, Nishitani & Bagley, 1999, p. 14), are displayed on the screen. These parenthetical phrases help clarify meaning and indicate the difference between the two sentences. The rejoinders are not included in the recording. The

Figure 6. *Example Item from the Sentence Stress Section (Part 1)*

Figure 7. *Example Item from the Sentence Stress Section (Part 2)*

rejoinders were included solely to help learners distinguish between the two sentences visually.

Technology

The *Perception of Spoken English* (POSE) *Test* is computer-based, allowing the test to be conducted outside of the language classroom and providing accurate diagnostic results without the required presence of the teacher. In order to accomplish the computer-based delivery of the POSE test, several different technologies were utilized and implemented. These include combinations of Hypertext Markup Language (HTML), PHP Hypertext Preprocessor (PHP), and database server software based on the Structured Query Language (SQL) called MySQL. Each of these technologies is explained below.

Hypertext Markup Language

While an in-depth tutorial of Hypertext Markup Language and its related technologies is not appropriate for this setting, a short explanation of how this technology is implemented in this project should be discussed. Hypertext Markup Language, commonly called HTML, is a set of codes used to "markup" or produce a hypertext document. These codes are interpreted by a piece of software called a "browser" which displays the information on a computer screen according to the coded instructions contained in the hypertext document (HTML Overview, 2004). This browser resides on a computer other than the server, and in most cases the creator of the hypertext document has no control over which browser the user utilizes when accessing the document from a remote computer.

It should be noted that HTML is not a programming language. The codes provide instructions to a browser in order to display information contained in the document, but the codes cannot be used to perform calculations, read or write to computer storage, or do many of the other tasks the are characteristic of a programming language.

PHP Hypertext Preprocessor

As the name of this technology suggests, it is designed to work in combination with hypertext documents. There are some key differences between HTML and PHP Hypertext Preprocessor (PHP), however. One major difference is that PHP can be used to perform calculations, store values, read and/or write to computer storage, and many other tasks that are characteristic of programming or scripting languages. Whereas HTML codes are interpreted by a browser and displayed on the screen accordingly, PHP programs cannot be read by a browser. Instead they are processed by a computer (usually called a "server") that has access to the correct interpreters. These interpreters process the PHP instructions and perform the desired actions. These actions are performed before the information is sent to the remote computer.

PHP is useful when working with Internet-based applications because it was designed to be used in combination with HTML codes. This purposeful interaction allows a programmer to create HTML coded documents based on input from users or other variables. It also allows programmers to separate specific data from the codes used to display that data on the computer screen. For example, a programmer may have a dictionary that he/she would like to display in different settings, i.e. a desktop computer, a personal handheld device, or an Internet-enabled cellular telephone. The words and

definitions in the dictionary do not change, yet the way that those words and definitions need to be displayed in the three different examples is different for each type of technology. A programmer could produce a document for each type of technology that contained PHP instructions, telling the computer to insert the appropriate information in the appropriate place. Where and how the information was inserted in the document would depend on the PHP instructions and the specific situation of the user accessing that information.

In the POSE test, the individual sentences, image files, audio files and other information specific to an individual item in any one section of the test are stored in a database (see the next section on database structure for an explanation) and a document, or template, instructs the computer to insert the appropriate information based on the particular item the user is trying to access. In this way, instead of creating a separate document for each item, which would then have to be updated each time changes were made to the overall design of the test, one template was created for each item type, and the information can be inserted at the appropriate time. Changes made to the template in turn change the way each item is displayed on the screen without having to change each individual item.

Database Structure

The last piece of technology utilized in the POSE test is a database server based on the Structured Query Language, or SQL, called MySQL. This technology is used to create and maintain databases of information. A database consists of sections, called tables; each table is much like a table or spreadsheet used in many modern office

applications. Each table contains information that can be accessed by using SQL commands. Often the tables contain information that relate to information in other tables. This type of database structure is called a relational database.

Returning to the dictionary example offered earlier, a relational database structure could be used to store the words and definitions contained in our dictionary. For example, one table could be created to hold each word entry in the dictionary. Each word in the table would be assigned a unique identification marker (usually a number). The table would contain the unique identification marker, the word, its etymology, and other specific information. However, if we were to place the definitions in this table as well, we would run into a problem. Some words have many different definitions while others have only one. If we place the definitions in the same table as the words, we would need to provide space for every possible number of definitions, creating a table with a lot of unused space.

Instead, we create a second table. In this table we place the unique identification marker of the word we are defining, the definition, and possibly an example sentence that illustrates the given definition. By placing these definitions in a separate table and using the unique identification marker to refer back to the word (instead of the word itself), we can place any number of definitions in the table without wasting space and without causing confusion.

In the POSE test, a table was created for each section, and each table contains the individual items for that section. Tables were also created to store the demographic and personal information of the participants in the piloting stage of the project, as well as to store the individual responses of each participant to each item in the test. Special PHP

instructions are used to insert, retrieve, and otherwise manipulate the data stored in these tables.

Implementation

The three technologies mentioned in this section—HTML, PHP, and MySQL—were designed to be used in conjunction with each other. The knowledge needed to utilize these technologies was obtained from coursework at Brigham Young University and through personal research on the Internet, in reference manuals, and in conversations with other programmers.

Each of these technologies is available to the public free-of-charge, and so no special licenses or copyright permissions are required. It should also be noted that while these technologies are widely used and accepted around the world, other technologies exist that could have been used in similar fashion to create a similar product. The POSE test utilizes these technologies for two main reasons. First, the author had previous knowledge and experience using these technologies and these technologies proved to be the most convenient in terms of learning new applications and methods for accomplishing different tasks. Second, as these technologies are widely utilized in many different ways by many different people, there is a plethora of resources readily available to offer support and help in resolving specific programming problems.

Audio Recordings

The audio files used in the POSE test are stored in MP3 format, which requires less disk space, allowing them to be transmitted over the Internet fairly quickly. The

acronym MP3 stands for MPEG Layer 3. MPEG stands for Moving Picture Experts

Group and denotes a file format for digitally storing video and audio data. This format

(MP3) is widely used and accepted around the world for transmitting and storing audio

data. The audio files were recorded digitally in WAV format (an audio format developed

by Microsoft Corporation) and then converted into MP3 format later on in the

development process. This use of different file formats was done to preserve the quality

of the recordings during the editing and finalization stages. Most of the audio was

recorded in the author's office using a program called *Sound Studio*, which runs only on a

Macintosh computer. Some of the audio files were recorded in a recording studio at

Brigham Young University, also using a Macintosh computer and a sound editing

program called *Peak*. Both audio programs are proprietary and were used according to

the designated user's license obtained by the university.

Visual Design

The visual design for the POSE test was taken from a template the author

downloaded from a web site. The template was designed by JSB Web Templates

(http://www.jsbwebtemplates.com) and was made available free of charge. A statement

attributing this fact is at the bottom of every page of the POSE test. The original template

contained a hyperlink to the template designer's website; however this hyperlink was

removed so that students are not able to access this hyperlink during the test.

CHAPTER 5

Piloting the *Perception of Spoken English* (POSE) *Test*

A major part of any materials development project is its piloting stage. This section will outline the methods used in the piloting of the *Perception of Spoken English* (POSE) *Test* and then look at the demographics of the participants involved.

Methods

The POSE test was piloted at the English Language Center (ELC) at Brigham Young University. The POSE test was designed to be administered using the Internet as the primary means of delivery. The POSE test was housed on the ELC's web server and students accessed the POSE test via the ELC's multimedia computer lab. Students were asked to provide information about their gender, nationality, native language, and English language learning experience. Then students were taken to a page that allowed them to check their audio system and ensure that audio problems would not hinder their ability to complete the test.

The test was split into two different forms to facilitate a shorter overall test time, and thus allow more students to take the test while placing less burden on the ELC's facilities. This approach also allowed for comparison between forms and examination of the overall reliability of the POSE. As students began the test, the computer randomly assigned them a form, either A or B. Table 4 shows the number of items per section in each form of the test.

Table 4. *Number of Items per Section in the Piloting Stage*

Section	Form A	Form B	Total Unique Items[a]
Vowel	22	22	38
Consonant	50	50	84
Word Stress	20	20	40
Intonation	21	19	40
Sentence	20	20	40

[a]Some items were repeated in both forms, so the number of unique

items is *not* equal to the sum of the number of items in each form.

Generally, in the vowel and consonant sections there were four items for each

contrast, or two items per contrast per form of the test. However, generating minimal pair

sentences for some of the contrasts was extremely difficult. Thus, for a small number of

vowel and consonant contrasts, only two items were available. Rather than have only one

item per form for these contrasts, the two items available were repeated on both forms.

Thus the total number of unique items for each section overall may be less than the sum

of the items in the same section of both forms.

Students completed the test in two stages, with each stage lasting no longer than

one hour. The first stage consisted of the vowel and consonant sections. The second stage

consisted of the word stress, intonation, and sentence stress sections. Students were asked

to sign up for a specific time to complete each stage. Most of the participants completed

both stages in one sitting. However, some completed both stages over the course of two

days. The computer assigned each student a form only once, and that same form was used

in both stages of the piloting. Students were required to complete stage one before they

could continue on to stage two. Also, once students began one of the stages, they were

required to finish that stage. Upon completion of each stage, a page showing the individual results appeared on the screen and students were given the option of having these results sent to their listening/speaking teacher.

Upon completion of both stages of the test, the computer issued each student a number. The students used these numbers to enter a drawing for movie tickets at a local movie theater. Participants were also offered candy as incentive for completing the test. The author provided these incentives at his own expense.

Participants

The participants in the piloting of the POSE test were students currently enrolled at the English Language Center at Brigham Young University. Participants ranged from 18 to 43 years of age, with the majority of the participants being below the age of 30. There were a total of 66 participants from ten different native language backgrounds. The largest language groups were Korean (23), Spanish (14), Japanese (10), and Chinese (7). The participants ranged in English language ability from intermediate to high-intermediate with the majority of the participants being at an intermediate level. A breakdown of the participants can be found in Table 5.

The piloting of the POSE test complied with the regulations of the Institutional Review Board (IRB) of Brigham Young University. As such, participation was strictly voluntary. Teachers encouraged their students to participate, but no form of punishment or coercion (academic or otherwise) was used. The dates for the pilot of the POSE test were announced in listening/speaking classes and students were informed of the different incentives available for participation (see the section on methods above).

Table 5. *Breakdown of Participants by Language and Level*

Native Language	Proficiency Level	Number of Participants	Total per Language
Bambara	Intermediate	1	1
Cantonese	High-Intermediate	1	1
Chinese	Intermediate	3	7
	High-Intermediate	4	
French	Intermediate	2	2
Japanese	Intermediate	8	10
	High-Intermediate	2	
Korean	Intermediate	18	23
	High-Intermediate	4	
	Advanced	1	
Mongolian	Intermediate	3	3
Portuguese	Intermediate	2	3
	Advanced	1	
Russian	High-Intermediate	1	2
	Advanced	1	
Spanish	Intermediate	11	14
	High-Intermediate	2	
	Advanced	1	

Note. Total number of participants equals 66.

Also, in order to comply with IRB regulations, students were asked to consent to participate in the piloting stages of the POSE test. These consent forms were presented on the screen as students began the first stage (see the section on methods above). Students checked either "Yes" or "No" after reading the information regarding participation. Students who checked "No" were allowed to leave without any negative consequences whatsoever.

Data Collection

The individual responses of each learner were collected and stored in a database on the ELC's web server. The data stored for each response consisted of each learner's unique id number (this number was assigned to each eligible participant at the beginning of the piloting stage), the item number of the particular item, and the answer the student had chosen in response to that item. These data were then compared to the correct answers stored in the database for each item. The items that did not match were output to a screen for the student to see as an indication of problem areas in that student's speech perception. The same comparisons were later used to determine the reliability and validity of the POSE test as described in the next chapter.

CHAPTER 6

Results and Evaluation of the Pilot

The data collected during the piloting stage were analyzed to determine the effectiveness of the POSE test in accurately diagnosing speech perception problems. Part of the analysis included determining the reliability and validity of the POSE test. This, however, proved to be more difficult than originally anticipated. The next two sections of this report will discuss these two features and the problems encountered.

Reliability

If two groups of people took the same test at different times, and the two sets of results were compared, we would be able to see how closely the two sets of scores resembled one another. If the scores closely resembled each other, the test could be said to be reliable. Reliability can be especially crucial when the results of a test or assessment are used to make important decisions. One example is the Test of English as a Foreign Language (TOEFL). Students pay a sizeable amount of money to take the TOEFL, and the results are used by universities in the United States to determine if students are eligible for admission. Certainly a test of this magnitude must be reliable. If the same student took the TOEFL on Friday, and again on Monday of the following week, and the scores varied greatly, those scores could not be used in making such important decisions.

The importance of reliability in evaluating the usefulness of any assessment will depend on how the assessment is used. The POSE test was designed solely as a diagnostic test and was not intended for use in making hefty decisions such as placement in a certain level within a language program, or admission to a university. However,

before learners can begin to correct problems in their speech perception, they must first know what those problems are. If the POSE test were not reliable, then learners could not use the results of the POSE test to determine real problems in speech perception. Teachers could not use the results of the POSE test to plan their courses to help students overcome their speech perception problems. Clearly then, a certain amount of reliability is desirable.

Reliability is usually expressed in terms of a reliability coefficient, or number between zero and one that represents the reliability of the assessment being evaluated. Depending on the type of assessment, and the demands of the assessment being evaluated, the desirable reliability coefficient can vary. Lado (1961) suggests that vocabulary, grammar, and reading comprehension tests should have a reliability coefficient above 0.90. Lado also indicates that aural comprehension assessments should have a reliability coefficient above 0.80 before they can be considered reliable. Oral production assessments are even lower than that, requiring a coefficient above 0.70. Hughes (1989) suggests that the reliability coefficient desired will depend on the decisions being made based on the results of the assessment being evaluated.

Having looked at the relevance of reliability and the desirable reliability coefficient, we can now turn to evaluating the reliability of the POSE test. One difficulty in assessing the reliability of the POSE test was its nature. In general, tests assess one specific skill or area of knowledge, and an overall score is generated. This overall score can then be compared with other scores achieved by the same person or groups of people and a reliability coefficient obtained. This method is usually called the "test-retest" method (Hughes, 1989, p. 32). Another method of generating a reliability coefficient is

the "alternate forms method" (Hughes, 1989, p. 32), in which the assessment is divided into two or more forms and those forms are administered to the same subjects. The scores can then be compared. Thus, comparing an entire test to itself in some form or another is acceptable. However, the POSE test is not designed to result in a final score that can be ascribed to the user's overall speech perception ability. Instead, it is designed to discover specific contrasts and patterns that learners have problems perceiving. Hence, the POSE test is not simply one test, but five smaller tests, with each test assessing a specific area of speech perception. It could even be argued that since each item, or group of items, examined the learners' ability to perceive different contrasts, each group of items form a single test. In this case, the POSE test is actually over 100 smaller tests combined into one larger instrument. Therefore, normal reliability measures proved inadequate in determining the overall reliability of the POSE test as a whole.

Therefore, instead of reporting a specific reliability coefficient, it seems sensible to present the data obtained from certain participants during the piloting stage and show how this helps establish the reliability of the POSE test. During the piloting stage of the project, several participants took the POSE test two times each. The computer randomly assigned students a form each time they took the POSE test. Some students were assigned the same form both times. Others were assigned alternate forms. This methodology affected the ability to determine the reliability of some sections of the POSE test. These effects and the methods of data analysis used for each section are discussed below.

Before presenting the data, however, it should be mentioned that future research in the area of reliability is needed. Indeed, one of the future plans discussed in Chapter 7

of this report is the necessity of future piloting and reliability studies. This need for future research should be remembered as we discuss the data in the following paragraphs.

The Vowel and Consonant Sections

The random assigning of forms did not particularly affect these sections. Two items for each contrast were included in each form. The column labeled "No. Incorrect" in Table 6 and Table 7 shows the number of contrasts where the particular student marked at least one item incorrect. The column labeled "No. Matched" shows the number of contrasts that appeared in the results of a particular student in both administrations of the POSE test.

While the data in Table 6 and Table 7 do not present any clear estimate of overall reliability for these sections of the POSE test, we can see that many of the students incorrectly distinguished a similar number of contrasts in each administration of the POSE test, and that the number of contrasts incorrectly distinguished by the same students in both administrations of the POSE test was 50% or higher of the total number of possible matches 17 out of 19 times. In other words, if a student incorrectly distinguished five contrasts on the first administration and three contrasts on the second administration, the total number of possible matches for both administrations for that student is three. In 17 out of 19 times, the number of contrasts that appeared on both sets of results was over 50% of the possible number of matches. While we cannot calculate a reportable reliability coefficient, these data give us some indication that these two sections are somewhat reliable.

Table 6. *Reliability Data for the Vowel Section*

Student	Form	No. Incorrect	No. Matched
1	B / B	2 / 6	2
2	B / A	4 / 3	2
3	A / A	5 / 5	3
4	A / A	5 / 2	2
5	A / B	6 / 4	3
6	A / A	3 / 2	0
7	B / A	5 / 5	4
8	A / A	6 / 7	6
9	A / B	5 / 7	3
10	B / B	8 / 6	5
11	B / B	4 / 3	2
12	A / B	5 / 6	4
13	A / B	6 / 4	4
14	B / A	3 / 2	1
15	A / B	8 / 7	5
16	B / A	8 / 3	3
17	B / A	6 / 5	4
18	A / A	3 / 2	0
19	A / B	4 / 4	3

Note. The columns labeled "No. Incorrect" and "No. Matched"

refer to the number of contrasts and not the number of items. Each

form of the test had at least two items for each contrast.

Table 7. *Reliability Data for the Consonant Section*

Student	Form	No. Incorrect	No. Matched
1	B / B	7 / 7	3
2	B / A	1 / 1	0
3	A / A	10 / 13	6
4	A / A	2 / 2	2
5	A / B	6 / 5	4
6	A / A	4 / 3	2
7	B / A	3 / 1	1
8	A / A	8 / 7	6
9	A / B	6 / 6	3
10	B / B	10 / 8	5
11	B / B	3 / 3	1
12	A / B	6 / 6	3
13	A / B	4 / 4	2
14	B / A	4 / 3	1
15	A / B	4 / 5	4
16	B / A	3 / 5	3
17	B / A	6 / 7	4
18	A / A	0 / 2	0
19	A / B	4 / 3	2

Note. The columns labeled "No. Incorrect" and "No. Matched"

refer to the number of contrasts and not the number of items. Each

form of the test had at least two items for each contrast.

The Word Stress Section

This section proved to be a little more difficult to analyze. The items in this

section were designed to test the participant's overall ability to perceive the stressed

syllable of any given word. However, the items were not as easily broken into categories

as were the items in the vowel and consonant sections. Thus, rather than arbitrarily assign

categories, it seemed more reasonable to analyze only the data from students who

completed the same form of this section. These data are presented in Table 8.

Table 8. *Reliability Data for the Word Stress Section*

Student	Form	No. Incorrect	No. Matched
1	B / B	1 / 2	1
2	B / B	6 / 9	5
3	A / A	2 / 8	1
4	A / A	14 / 2	2
5	A / A	0 / 0	0
6	B / B	10 / 9	7
7	B / B	1 / 0	0

Note. The columns marked "No. Incorrect" and "No. Matched"

refer to the number of items and not the number of contrasts.

It should be noted that the columns labeled "No. Incorrect" and "No. Matched"

refer to the number of items, and not to the number of contrasts. This is different than the

vowel and consonant section data presented in Table 6 and Table 7.

While seven cases is hardly an appropriate number for analysis, these data may

help to shed some light on the overall reliability of this section. Again, it should be

remembered that future plans include more reliability studies. For a detailed description of these plans, see Chapter 7.

From the data presented in Table 8, we can see that four students out of seven incorrectly distinguished similar numbers of items in each administration of the POSE test. Of those four, three of the students incorrectly distinguished the same items on both of the administrations over 50% of the time. While not significantly reliable, this follows the trend set in other sections of the POSE test.

The Intonation Section

This section differed slightly from the others in the way the items were categorized. This section consisted of two types of items, but each type focused solely on the intonation at the end of an utterance, either rising or falling. Therefore, only two categories, or possible contrasts were needed: one for each item type. With only two categories possible, it can reasonably be assumed that the data would show a higher level of reliability. These data are presented in Table 9; however, it is important to remember that the data do not accurately reflect the true reliability of the POSE test and more research is needed in this area (see Chapter 7 for a discussion of future reliability studies). Because the two categories of items are the same for both forms of the POSE test, we can use the data from all the students, rather than limit the analysis to only the data from the same form.

Table 9. *Reliability Data for the Intonation Section*

Student	Form	No. Incorrect	No. Matched
1	B / B	0 / 0	0
2	B / B	1 / 1	1
3	A / A	1 / 1	1
4	A / A	2 / 0	0
5	B / A	1 / 1	1
6	A / A	0 / 2	0
7	A / B	0 / 0	0
8	B / B	1 / 1	1
9	B / B	0 / 0	0
10	A / B	1 / 1	1
11	A / B	1 / 0	0
12	B / A	0 / 0	0
13	A / B	0 / 0	0
14	B / A	0 / 0	0
15	B / A	1 / 1	1

Note. Only two categories of items existed in this section. The columns labeled "No. Incorrect" and "No. Matched" refer to the categories and not the individual items.

Examining the data in Table 9, we can see that most of the students either did not incorrectly perceive any of the items presented in this section, or incorrectly perceived one item on one or both of the administrations of the POSE test. Again, only two categories were available for analysis. Yet, the trend shown in other sections of the POSE

test holds true here also. A majority of the students incorrectly perceived a similar number of categories on both administrations of the POSE test. This may indicate some degree of reliability for this section; however, because the number of comparable categories or contrasts was so low, these data cannot be regarded as completely indicative of reliability. Further research is needed in this area.

The Sentence Stress Section

This section was similar to the word stress section mentioned above in terms of determining reliability. The items in this section focused on the ability to perceive pauses and stresses in any particular utterance. However, they were not easily broken apart into categories. Therefore, only the data from participants who took the same form in each administration of the POSE test could be analyzed to determine the reliability of this section. The data are presented in Table 10.

Table 10. *Reliability Data for the Sentence Stress Section*

Student	Form	No. Incorrect	No. Matched
1	B / B	1 / 2	0
2	B / B	0 / 0	0
3	A / A	1 / 1	1
4	A / A	2 / 3	2
5	A / A	2 / 1	1
6	B / B	1 / 0	0
7	B / B	1 / 0	0

Note. The columns labeled "No. Incorrect" and "No. Matched" refer to the number of items.

The data show that all the participants incorrectly perceived a very few number of items in each administration. Coupled with the low number of cases used in this analysis, using these numbers to determine an accurate reliability coefficient is impossible. However, as with other sections in the POSE test, several trends appear to be present in the data. First, while all the participants incorrectly perceived only a few number of items in each administration, the scores from each administration differ by only one in all cases. Another trend that can be seen is that in all but one of the cases, the items incorrectly perceived on the first administration match almost all those incorrectly perceived on the second administration and vice verse. This would suggest some degree of reliability; however, more research and piloting is necessary before any claims of reliability can be made.

Validity

Due to the nature of the POSE test, general validity was difficult to assess. One reason for this difficulty lay in the fact that no overall score was given when users completed the test. The POSE test output a list of contrasts and suprasegmental patterns that learners had trouble distinguishing, but no quantitative score was calculated. Even if such a number were made available, it would not be relevant to the actual purpose and design of the POSE test as a diagnostic tool.

Another reason for this difficulty in assessing the general validity of the POSE test is that there are no equivalent measures with which the POSE test can be compared. Therein lay the irony. On the one hand, the POSE test fulfills a strong need in the ESL community, as shown in Chapter 3 of this report. Yet, at the same time, because there are

no other equivalent measures, the validity of the POSE test is difficult to properly determine.

One method derived to combat this problem included analyzing the research on contrastive analysis and looking at contrasts that have proven to be problem areas for specific language groups, and then analyzing the data collected from the piloting stage of this project to see if both means point to the same end. However, very little research is available on the contrast-specific problems of each language group. Currently only two plausible sources exist. The next few paragraphs look at the information contained in these sources and how it compares with data obtained during the piloting stage of this project.

Pronunciation Contrasts in English (Nilsen & Nilsen, 2002)

The first plausible source of contrastive analysis information was produced in 1971, and then reissued in 2002, by Don L. F. Nilsen and Alleen Pace Nilsen, called *Pronunciation Contrasts in English*. This book lists the common vowel and consonant contrasts in English and provides examples of minimal pair words and sentences for each contrast. Also included with each contrast is a list of languages that can be expected to have problems in perception and production of that particular contrast. The data from each language group was examined and then compared to the lists in this resource. This data is shown in Table 11 and Table 12.

Table 11. *Validity Data for Vowel Contrasts*

Contrast	Language	% Incorrect[a]	On List[b]
/i/ - /ɪ/	Korean	53.5%	Yes
	Spanish	46.5%	Yes
	Japanese	8.25%	Yes
	Chinese	17.5%	Yes
/ɪ/ - /ɛ/	Korean	6.75%	No
	Spanish	16.25%	No
	Japanese	29.25	No
	Chinese	33.75%	No
/e/ - /ɛ/	Korean	8.25%	Yes
	Spanish	11.75%	Yes
	Japanese	23%	Yes
	Chinese	31.25%	No
/ɛ/ - /æ/	Korean	36.75%	Yes
	Spanish	39.75%	Yes
	Japanese	33.25%	Yes
	Chinese	27.5%	Yes
/æ/ - /ɑ/	Korean	3%	Yes
	Spanish	5.75%	Yes
	Japanese	11.75%	Yes
	Chinese	17.5%	Yes
/ɑ/ - /ɑr/	Korean	29%	------

Contrast	Language	% Incorrect[a]	On List[b]
/ɑ/ - /ɑr/	Spanish	14.5%	------
	Japanese	26.25%	------
	Chinese	10%	------
/ɑ/ - /ʌ/	Korean	44.75%	No
	Spanish	33.75%	Yes
	Japanese	24.25%	Yes
	Chinese	40%	Yes
/ɑ/ - /ow/	Korean	26.75%	No
	Spanish	46.75%	Yes
	Japanese	37.5%	Yes
	Chinese	27.5%	Yes
/ʌ/ - /ɚ/	Korean	3.25%	------
	Spanish	0%	------
	Japanese	13.75%	------
	Chinese	6.25%	------
/ʌ/ - /ʊ/	Korean	15%	No
	Spanish	23.25%	Yes
	Japanese	11.75%	Yes
	Chinese	6.25%	Yes
/ʊ/ - /uw/	Korean	35.5%	No
	Spanish	50%	Yes

Contrast	Language	% Incorrect[a]	On List[b]
/ʊ/ - /uw/	Japanese	38.5%	No
	Chinese	22.5%	Yes

Note. The numbers of responses for each language are:

Korean, *n* = 31; Spanish, *n* = 18; Japanese, *n* = 14; Chinese, *n*

= 9. These numbers include the responses of participants who

took the POSE test more than once. A dash (-----) indicates

that data for that contrast was not available.

[a]The percents given in this column are the average percent

marked incorrect by each language group. [b]Source: Nilsen, D.

L. F. & Nilsen, A. P. (2002). *Pronunciation Contrasts in*

English. Long Grove, IL: Waveland Press. The list refers to

the list of languages that are expected to have problems with

that particular contrast given in the book.

Upon examining the data in Table 11 and Table 12, several general patterns in the

data can be seen. First, in many cases, the lists given in *Pronunciation Contrasts in*

English (Nilsen & Nilsen, 2002) correspond with the data obtained from the piloting of

the POSE test. In several cases, the data show that one particular language group had

difficulty correctly perceiving a certain contrast, while the other language groups did not.

The lists in *Pronunciation Contrasts in English* correspond with this data. This may

indicate that the POSE test is a valid instrument for determining problems with these

particular contrasts.

Table 12. *Validity Data for Consonant Contrasts*

Contrast	Language	% Incorrect[a]	On List[b]
/p/ - /b/	Korean	7.5%	Yes
	Spanish	6.25%	Yes
	Japanese	4.25%	No
	Chinese	0%	No
/p/ - /f/	Korean	22.5%	Yes
	Spanish	0%	No
	Japanese	2.13%	No
	Chinese	0%	No
/v/ - /b/	Korean	16.5%	Yes
	Spanish	33%	Yes
	Japanese	16.75%	Yes
	Chinese	16.25%	No
/w/ - /v/	Korean	3.5%	No
	Spanish	4.5%	No
	Japanese	0%	No
	Chinese	0%	Yes
/f/ - /v/	Korean	1.75%	Yes
	Spanish	3.5%	No
	Japanese	0%	Yes
	Chinese	0%	No
/f/ - /θ/	Korean	4.5%	Yes

Contrast	Language	% Incorrect[a]	On List[b]
/f/ - /θ/	Spanish	6.75%	Yes
	Japanese	0%	No
	Chinese	5%	No
/θ/ - /t/	Korean	7.25%	Yes
	Spanish	15.75%	Yes
	Japanese	7.88%	Yes
	Chinese	16.88%	No
/θ/ - /s/	Korean	13.38%	Yes
	Spanish	8.13%	Yes
	Japanese	14.25%	Yes
	Chinese	13.13%	Yes
/ð/ - /d/	Korean	35%	Yes
	Spanish	44.25%	Yes
	Japanese	36.5%	Yes
	Chinese	12.5%	Yes
/t/ - /d/	Korean	3%	Yes
	Spanish	10.25%	No
	Japanese	0%	No
	Chinese	15%	No
/n/ - /l/	Korean	3.38%	No
	Spanish	2.25%	No

Contrast	Language	% Incorrect[a]	On List[b]
/n/ - /l/	Japanese	3.25%	No
	Chinese	0%	No
/l/ - /r/	Korean	31%	Yes
	Spanish	0%	No
	Japanese	52%	Yes
	Chinese	0%	Yes
/s/ - /z/	Korean	36.25%	Yes
	Spanish	65.75%	Yes
	Japanese	12.75%	No
	Chinese	58.75%	Yes
/s/ - /ʃ/	Korean	14%	Yes
	Spanish	0%	Yes
	Japanese	22%	Yes
	Chinese	0%	No
/ʃ/ - /tʃ/	Korean	5.88%	Yes
	Spanish	15.5%	Yes
	Japanese	11.63%	No
	Chinese	6.25%	No
/tʃ/ - /dʒ/	Korean	3.25%	Yes
	Spanish	15.25%	Yes
	Japanese	3.25%	No

Contrast	Language	% Incorrect[a]	On List[b]
/tʃ/ - /dʒ/	Chinese	0%	No
/dʒ/ - /y/	Korean	0%	No
	Spanish	37.25%	Yes
	Japanese	0%	No
	Chinese	0%	No
/k/ - /g/	Korean	12.5%	Yes
	Spanish	15.63%	No
	Japanese	2.13%	No
	Chinese	5.63%	No

Note. The POSE test focuses on both syllable-initial and syllable-final consonant contrasts. The language lists referred to in the "On List" column do not differentiate based on syllable position. Therefore, the data for these items was combined. The numbers of responses for each language are: Korean, $n = 31$; Spanish, $n = 18$; Japanese, $n = 14$; Chinese, $n = 9$. These numbers include the responses of participants who took the POSE test more than once. A dash (-----) indicates that data for that contrast was not available.

[a]The percents given in this column are the average percent marked incorrect by each language group. [b]Source: Nilsen, D. L. F. & Nilsen, A. P. (2002). *Pronunciation Contrasts in English.* Long Grove, IL: Waveland Press. The list refers to the list of languages that are expected to have problems with that particular contrast given in the book.

Other patterns in the data, however, do not support the language lists found in *Pronunciation Contrasts in English* (Nilsen & Nilsen, 2002). In several cases, the data show that two or more language groups had similar difficulties in correctly perceiving certain contrasts, but the lists in *Pronunciation Contrasts in English* do not correspond with these results. Indeed, in some cases where the data show a particular contrast proved to be a difficult problem for a particular language group, that particular language could not be found on the list in *Pronunciation Contrasts in English* for that particular contrast.

In short, sometimes the data obtained during the piloting stage of the POSE test and the lists in *Pronunciation Contrasts in English* (Nilsen & Nilsen, 2002) correspond and sometimes they do not. There are several possible reasons for this. One obvious possibility is that the POSE test is not a valid test for some of the contrasts it assesses. A second possibility is that the lists in *Pronunciation Contrasts in English* do not accurately reflect potential problem areas for some of the contrasts in the book. In the introduction to the book, the authors explain that the lists were compiled with the help of a group of over fifty linguists. They also point out, "that the specialists consulted could not reasonably be expected to anticipate individual digressions or to analyze all difficulties with uniform consistency" (p. xiii). This may account for some of the discrepancies between the data in Table 11 and Table 12, and the language lists presented in the book. Based on the general patterns, however, it is possible to tentatively conclude that the POSE test has some degree of validity in accurately diagnosing problems in the perception of the vowel and consonant contrasts included in the test.

Learner English (Swan & Smith, 2001)

A second possible source of this kind of contrastive analysis information is a book edited by Michael Swain and Bernard Smith (2001) called *Learner English*. This resource provides a chapter on each of a number of commonly spoken languages and contrasts each language with English. The information, however, is not complete in that it does not consider every contrast or suprasegmental pattern assessed in the POSE test. Still, we may be able to draw some conclusions as to the validity of the POSE test by comparing the results of the piloting of the POSE test and the data found in *Learner English*. We will examine each language group separately. The language groups examined below are this with a relatively high number of participants: Korean ($n = 31$), Spanish ($n = 18$), Japanese ($n = 14$), and Chinese ($n = 9$). The data are presented in Table 13, Table 14, Table 15, and Table 16.

Korean

    Vowels and Consonants.

The chapter on the Korean language was written by Jung-Ae Lee (2001). The data concerning typical problems Korean speakers might encounter in learning and using English was informative; however, it was difficult to extrapolate specific problems and compare them with the POSE test because much of the information presented in this chapter of *Learner English* was general and not specific. Still, some comparisons could be made. Table 13 shows the vowel and consonant contrasts assessed in the POSE test that the author of this chapter claims to prove problematic for Korean speakers.

Table 13. *Validity Data for Vowels and Consonants*

(Korean Speakers)

Vowel		Consonant	
Contrast	% Incorrect[c]	Contrast	% Incorrect[c]
/i/ - /ɪ/[a]	53.5%	/p/ - /b/[b]	7.5%
/e/ - /ɛ/[a]	8.25%	/p/ - /f/	22.5%
/æ/ - /ɑ/[a]	3%	/v/ - /b/	16.5%
/ɑ/ - /ar/	29%	/f/ - /v/[b]	1.75%
/ɑ/ - /ʌ/	44.75%	/θ/ - /s/	13.38%
/ʌ/ - /ɚ/	3.25%	/s/ - /ʃ/	14%
/ʊ/ - /uw/[a]	35.5%	/l/ - /r/	31%
		/s/ - /z/[b]	36.25%
		/ð/ - /d/	35%
		/t/ - /d/[b]	3%
		/tʃ/ - /dʒ/[b]	3.25%
		/k/ - /g/[b]	12.5%

Note. Responses: *n* = 31.

[a]These contrasts were included on the grounds that Korean speakers have problems distinguishing between tense and lax vowels, or as Learner English calls the "long/short vowel distinction" (p. 326). [b]These contrasts were included on the grounds that Korean speakers have problems distinguishing between voiced and unvoiced consonants (p. 326). [c]The percentages in these two columns are averaged from all the items assessing that particular contrast.

The number of responses for the Korean data is only 31; not enough for any accurate validity analysis. However, we can still see that for many of the contrasts presented in Learner English as potential problems for Korean speakers, the data from the piloting stage of the POSE test show that a good percentage of Korean speakers incorrectly perceived these contrasts. This suggests some degree of validity for diagnosing the problems experienced by many Korean speakers of English.

Word Stress, Intonation, and Sentence Stress.

For these three sections, only general information was given. Lee (2001) points out that the Korean language does not employ either syllable stress or word stress, indicating that these areas may be problematic for Korean speakers of English. She also points out that "particular words in Korean sentences are not stressed in relation to other words in the sentence. The differences that stressing one word can make to the meaning of a sentence are completely foreign to the Korean learner, and require concentrated attention to be perceived or produced" (p. 328).

In terms of intonation, Lee (2001) says that Korean statements and questions, other than *yes/no* questions, generally end with falling intonation, while *yes/ no* questions and requests generally end with rising intonation. As this is somewhat similar to English, it could be expected that this characteristic would not prove problematic for Korean speakers of English. The data concerning these three sections of the POSE test are shown in Table 14.

Table 14. *Validity Data for Word Stress, Intonation and Sentence Stress (Korean Speakers)*

Section	% Incorrect[a]
Word Stress	9.43%
Intonation (Part 1)	0.7%
Intonation (Part 2)	11.2%
Sentence Stress	2.75%

Note. The intonation section was divided into two parts. The first part focused on the difference between statements and questions using rising and falling intonation. The second part of the intonation section focused on the rising and falling intonation of tag questions. Responses: $n = 29$.

[a]The percentages in this column are averaged from all the items in a particular section of the POSE test.

According to the data in Table 14, the section on word stress proved to be the most difficult of the three sections, which agrees with the data presented in *Learner English* (Lee, 2001). The intonation section proves to be the most interesting because it was divided into two parts. The first part of the intonation section focused on the rising and falling intonation at the end of a question or a statement in English. Lee (2001) points out in *Learner English* that the Korean language has this characteristic and therefore should not be a problem for Korean speakers of English. The data show that indeed it was

not. However, Lee also points out that the Korean language uses falling intonation for every utterance except *yes/no* questions and requests. The second part of the intonation section focused on the rising and falling intonation of tag questions, something that is foreign to native Korean speakers. Thus, we would expect them to have a higher rate of error in this part of the intonation section, and this can be seen in the data shown in Table 14. These data then provide some evidence, small though it may be, of validity.

Spanish

    Vowels and Consonants.

The information presented in this chapter (Coe, 2001) of *Learner English* is usually general and not contrast specific. However, comparison with some contrasts can still be made. These data are shown in Table 15.

Examining the data in Table 15, we can see several similarities between the POSE test and *Learner English*. First, Coe (2001) points out that the Spanish language does not have the phoneme /z/. The data from the POSE test show that over half the Spanish-speaking participants incorrectly perceived the contrast /s/ - /z/. Also, Coe points out that European Spanish speakers often pronounce /s/ closer to /ʃ/. There were no European Spanish-speaking participants in the pilot of the POSE test and the data show that this contrast was not a problem. Finally, Coe points out that Spanish speakers often mispronounce /y/ as /dʒ/, which is consistent with the data in Table 15, showing that over thirty-five percent of the participants could not correctly perceive this contrast.

Table 15. *Validity Data for Vowel and Consonants*

(Spanish Speakers)

Vowel		Consonant	
Contrast	% Incorrect[b]	Contrast	% Incorrect[b]
/i/ - /ɪ/	46.5%	/p/ - /b/	6.25%
/æ/ - /ɑ/	5.75%	/b/ - /v/	33%
/ɑ/ - /ɑr/	14.5%	/s/ - /ʃ/[a]	0%
/ɑ/ - /ʌ/	33.75%	/ʃ/ - /tʃ/	15.5%
/ʊ/ - /uw/	50%	/s/ - /z/	65.75%
		/ð/ - /d/	44.25%
		/t/ - /d/	10.25%
		/tʃ/ - /dʒ/	15.25%
		/dʒ/ - /y/	37.25%
		/k/ - /g/	15.63%

Note. Responses: *n* = 18.

[a]According to Learner English, this contrast is most problematic for speakers of European Spanish. [b]The percentages in these two columns are averaged from all the items assessing that particular contrast.

These data help provide limited evidence that the POSE test is able to accurately pinpoint certain problematic contrasts for Spanish speakers of English.

Word Stress, Intonation, and Sentence Stress.

Coe (2001) emphasizes that Spanish is a syllable-timed language. This means that each syllable is given the same amount of stress and time in a sentence. Thus, we could reasonably expect that word stress in English might prove to be a problem for Spanish speakers. Coe also says, "Spanish…learners find variable stress intractable, and they cannot usually either recognise or produce the difference in English expressions like: the black bird/the blackbird, [or] the green house/the greenhouse" (p. 95). The second part of the sentence stress section contains items that assess this feature of English. Coe also says that Spanish tends to place important words at the end of a sentence, and thus Spanish speakers have trouble distinguishing stress when it is placed on words in the middle or beginning of a sentence in English. This feature of English is covered in the first part of the sentence stress section. Coe makes no mention of rising or falling intonation in Spanish, and so these data are not presented for comparison. The validity data for word stress and sentence stress are shown in Table 16.

It is difficult to draw any accurate conclusions with so few responses to consider. However, the data do show that word stress and sentence stress did prove problematic for some Spanish speaking participants. It is expected that with a larger group of participants and more data to consider, this percentage would increase.

Table 16. *Validity Data for Word Stress*

and Sentence Stress (Spanish Speakers)

Section	% Incorrect[a]
Word Stress	18.43%
Sentence Stress (Part 1)	5.8%
Sentence Stress (Part 2)	5%

Note. Responses: *n* = 16.

[a]The percentages in this column are averaged from all the items for that particular section.

Japanese

Vowel and Consonants.

This chapter, written by Ian Thompson (2001), provides only limited data on specific vowel and consonant contrasts. Some of the specific contrasts or phonemes discussed in this chapter were not included in the POSE test due to their low functional load (Catford, 1987; see also Chapter 4 of this report). The data available for comparison are shown in Table 17.

With so few contrasts and so few responses, it is almost impossible to draw any valuable conclusions about the validity of the vowel and consonant sections with regards to Japanese speakers of English. However, examining what data we do have shows that for the contrasts shown in Table 17, a relatively high number of Japanese participants incorrectly perceived these contrasts. This is consistent with the information presented by Thompson (2001).

Table 17. *Validity Data for Vowels and Consonants*

(Japanese Speakers)

Vowel		Consonant	
Contrast	% Incorrect[a]	Contrast	% Incorrect[a]
/æ/ - /ɑ/	11.75%	/v/ - /b/	16.75%
/ɑ/ - /ʌ/	24.25%	/s/ - /ʃ/	22%
		/l/ - /r/	52%

Note. Responses: *n* = 14.

[a]The percentages in these two columns are averaged from all the items assessing that particular contrast.

Word Stress, Intonation, and Sentence Stress.

Thompson (2001) says that Japanese speakers are generally very adept at "hearing and repeating stress and intonation patterns" (p. 299). He mentions, however, that only a limited number of suprasegmental commonalities exist in English and Japanese. He points out that "Japanese does not have the equivalent of 'weak' unstressed forms of words" (p. 299). This may indicate that perceiving differences in word stress might be a problem for Japanese speakers of English. The data for this section of the POSE test, shown in Table 18, support this conclusion. Thompson also points out that Japanese utilizes pitch change on new or important ideas in a sentence, as well as rising intonation at the end of questions and tag questions where the speaker is looking for information. Japanese also utilizes falling intonation at the end of statements and tag questions where the speaking is making a comment. These features are similar to the features in English

Table 18. *Validity Data for Word Stress, Intonation, and Sentence Stress (Japanese Speakers)*

Section	% Incorrect[a]
Word Stress	18.38%
Intonation	5.43%
Sentence Stress	3.98%

Note. Responses: $n = 14$.

[a]The percentages in this column are averaged from all the items for that particular section.

assessed by the POSE test in the intonation and sentence stress sections. The data in Table 18 show that these features were not difficult for Japanese speakers to correctly perceive, indicating that these sections may have some validity in accurately diagnosing the problems Japanese speakers have in perceiving the suprasegmental patterns of English.

Chinese

Vowels and Consonants.

This chapter was written by Jung Chang (2001). Chang points out a fairly high number of consonant contrasts that are problematic for Chinese speakers, but very few vowel contrasts. Some of the vowel and consonant contrasts Chang discusses were not included in the POSE test, so they are not presented for comparison. The data showing the available vowel and consonant contrasts are shown in Table 19.

Table 19. *Validity Data for Vowels and Consonants*

(Chinese Speakers)

Vowel		Consonant	
Contrast	% Incorrect[a]	Contrast	% Incorrect[a]
/i/ - /ɪ/	17.5%	/p/ - /b/	0%
/ɑ/ - /ʌ/	40%	/w/ - /v/	0%
/ʊ/ - /uw/	22.5%	/f/ - /v/	0%
		/θ/ - /f/	5%
		/θ/ - /t/	16.88%
		/θ/ - /s/	13.13%
		/ð/ - /d/	12.5%
		/t/ - /d/	15%
		/n/ - /l/	0%
		/l/ - /r/	0%
		/s/ - /z/	58.75%
		/k/ - /g/	5.63%

Note. Responses: *n* = 9.

[a]The percentages in these two columns are averaged from all the items that assessed that particular contrast.

For five of the consonant contrasts shown in Table 19, none of the Chinese participants incorrectly perceived these sounds. Chang (2001) indicates that these sounds prove difficult for speakers of some Chinese dialects, but not universally for all speakers of Chinese. Chinese participants were not given the opportunity to indicate a particular

dialect of Chinese as their native language, and so it is not known if any of the Chinese participants spoke the dialects Chang mentions. This may account for the discrepancies shown in the data.

Word Stress, Intonation, and Sentence Stress.

Chang (2001) states that, while fairly common in English, Chinese does not have reduced syllables. This would seem to indicate that word stress might be a problem for many Chinese speakers of English. The data for the word stress section of the POSE test, shown in Table 20, indicate that word stress was a problem for a good number of the Chinese speaking participants. Chang also says that Chinese "sentence intonation [or sentence stress] shows little variation. The English use of [sentence stress] patterns to affect the meaning of a whole utterance is therefore difficult for Chinese [speakers] to grasp" (p. 313). The data in Table 20 indicate this was a problem for some of the Chinese speaking participants, but not for a majority. It is expected, however, based on the information provided by Chang, that with a larger group of Chinese speakers, the results would show a higher number of participants who have trouble distinguishing meaning based on sentence stress. During the presentation of this project at a national convention, the author received a request to pilot the POSE test with a group of Chinese learners at Princeton University. Plans for this research are in the works, as discussed in Chapter 7 of this report. It is hoped that this research will shed more light on the validity of the sentence stress section.

Table 20. *Validity Data for Word*

Stress, Intonation, and Sentence

Stress (Chinese Speakers)

Section	% Incorrect[a]
Word Stress	38.3%
Intonation (Part 1)	0%
Intonation (Part 2)	11.63%
Sentence Stress	6.83%

Note. Responses: $n = 8$.

[a]The percentages in this column are averaged from all the items in that particular section.

One verifiable trend shown in the analysis of the collected data was that those learners who possessed a greater level of proficiency generally responded correctly more times than those with lower levels of proficiency. This is consistent with the findings of Flege and Eefting (1987) which showed that higher proficiency led to more accurate perception. The data showing this trend are presented in Table 21.

Table 21. *Percentage of Incorrect Responses by Language and Level*

Section	Language	Level	No. of Responses	% Incorrect[a]
Vowels	Korean	5	1	9.09%
		4	4	12.5%
		3	26	23.5%
	Spanish	5	1	4.55%
		4	2	7.95%
		3	15	28.11%
	Japanese	4	2	12.5%
		3	12	23.91%
	Chinese	4	4	26.14
		3	5	18.93
Consonants	Korean	5	1	3%
		4	4	5.5%
		3	26	12.18%
	Spanish	5	1	4%
		4	2	5.5%
		3	15	14.3%
	Japanese	4	2	3%
		3	12	9.36%
	Chinese	4	4	8.5%
		3	5	7.16%

Section	Language	Level	No. of Responses	% Incorrect[a]
Word Stress	Korean	5	1	0%
		4	4	0.63%
		3	26	11.13%
	Spanish	5	1	2.5%
		4	2	3.75%
		3	15	21.7%
	Japanese	4	2	2.5%
		3	12	19.65%
	Chinese	4	4	30%
		3	5	41.88%
Intonation	Korean	5	1	0%
		4	4	1.25%
		3	26	7.35%
	Spanish	5	1	12.5%
		4	2	7.5%
		3	15	12%
	Japanese	4	2	1.25%
		3	12	5.4%
	Chinese	4	4	2.5%
		3	5	16.25%

Section	Language	Level	No. of Responses	% Incorrect[a]
Sentence Stress	Korean	5	1	0%
		4	4	3.75%
		3	26	2.13%
	Spanish	5	1	2.5%
		4	2	0%
		3	15	5.43%
	Japanese	4	2	1.25%
		3	12	4.18%
	Chinese	4	4	2.5%
		3	5	8.13%

[a]The percentages in this column are averaged from all the items in a particular section.

CHAPTER 7

Limitations, Subsequent Changes, and Future Plans

Obviously, there is still much that can be done to make the *Perception of Spoken English* (POSE) *Test* a more effective tool for diagnosing speech perception problems. There are several aspects of speech perception that are not assessed by the POSE test. Some of these aspects were not incorporated because of limitations in time and resources for the completion of this project. Other aspects fall outside the design of the POSE test, and as such were not incorporated into this project. By evaluating the results of the pilot of the POSE test, some changes and adjustments were decided upon and made to different areas of the test. Other changes will have to wait. The next three sections discuss some of these limitations and subsequent changes, and then discuss plans for future research.

Limitations

While the POSE test fills a void in speech perception diagnostic materials, there are certain aspects of speech perception that are not assessed. Some of these aspects were not included because of time constraints involved in the completion of this master's project. These aspects should be included in future versions of the POSE test to help ensure a useful and accurate diagnostic tool. Other aspects of speech perception did not fall within the scope of this project, and so were not incorporated. These are discussed below.

One of the aspects of speech perception not included in the POSE test due to time constraints is the perception of reduced syllables in English. English is a stress-timed

language, which means that unimportant syllables or sounds are reduced, or shortened, while important syllables or sounds are given more time in the utterance. Many non-native English speakers have difficulty producing and/or recognizing these reduced sounds and syllables. Some of this aspect may be present in the word stress section; however, this section does not provide an adequate diagnosis of this important feature. Future versions of the POSE test will hopefully include this important aspect of speech perception.

Another aspect of speech perception that will hopefully appear in future versions of the POSE test is the perception of syllabic consonants in English. Syllabic consonants function as "weak syllable[s]" (Celce-Murcia, Brinton, & Goodwin, 1996, p. 67) without a separate vowel sound. These include [n̩] and [l̩], as in [tʌn̩l̩] or [ʃʊdn̩t]. Many times, other consonant sounds are changed based on the sounds that follow them. In the case of syllabic consonants, sounds like /t/ and /d/ may change to become a glottal stop (/ʔ/). The segmental sections of the POSE test focus mainly on phonemic contrasts in English utilizing minimal pair sentences. Items that diagnose this unique aspect of English phonology should be included either in the consonant section, or perhaps in a future section on reduced speech (see previous paragraph). These items will be added to future versions of the POSE test.

Some aspects of speech perception did not fall within the scope of this project, and did not lend themselves to being tested in the manner utilized in the POSE test. For example, some pronunciation errors are not based on the learner's inability to hear those sounds or patterns. Instead, these errors are influenced by the learner's knowledge of English spelling and other linguistic features. One example of this is the *–ed* ending, as in

passed or *blessed*. Many non-native English speakers see the word *passed* and pronounce

it /pæsəd/, instead of /pæst/. Whether or not learners hear the difference when listening

to native speakers may be irrelevant in determining the source of this type of

pronunciation error. It is not known if this orthographic influence affects the perception

of this ending, but this type of perception skill is not one that can be assessed using

minimal pairs. As such, the possibility of including this aspect of perception in the POSE

test is not likely.

Another limitation of this project, and not particularly a limitation of the POSE

test itself, was the lack of sufficient data in determining the reliability of the POSE test.

This was due not to the lack of available participants because the total number of

participants was 66. This limitation was due to the random assigning of either form A or

form B to participants who took the POSE test more than once. Thus, the number of

participants who took the same form both times was very few. Hence, the data available

for determining the reliability of some sections of the POSE test was grossly limited.

Future research studies will seek to include sufficient numbers of participants whose

responses to the items in the POSE test can be analyzed to provide more accurate data

regarding the reliability of the POSE test.

Subsequent Changes

The version of the POSE test described in Chapter 4 of this project report differs

from the original version used in the piloting stage. These differences are the result of

adjustments made in response to data collected during the piloting of the POSE test and

feedback from participants. The version of the POSE test utilized during the piloting

stage of this project existed on a web server, and the very structure and design of the past and current versions POSE test require technologies that can only be implemented using the Internet as a means of delivery. For this reason, a working version of the current version of the POSE test could not be included in this project report. In lieu of a working version of the POSE test, the items for each section of the POSE test are included in Appendices A through E. The changes described in the next few paragraphs have been added to the description of the POSE test in Chapter 4, and any future research that utilizes the POSE test will be done using this newer version. However, this newer version has not yet been used by any learner.

One major adjustment made to every section was the addition of an "I don't know" option to each item in the test. Originally, students were forced to choose one of the two or more possible responses in each minimal-pair set. This allowed factors such as guessing to affect the final results. In other words, in the original version of the POSE test, a student had a fifty-percent chance of selecting the right answer for most of the items. The addition of an "I don't know" option helps avoid this problem by allowing the students an opportunity to indicate that they simply cannot distinguish between the responses in the minimal-pair set for any particular item. This change was suggested by several people after examining the results of the piloting stage and presentations about this project made at national and local conventions.

Another change made to the POSE test was the combining of both forms into one complete assessment. In the piloting stage of this project, time was an issue, and as such, the test was divided into forms to shorten the time required to complete each section. In the newer version of the POSE test, the items from each form have been compiled into

one tool, providing more items for each contrast or aspect of speech perception, and thus leading to a more accurate diagnosis of speech perception problems.

Although this change necessitates using more time to complete each section of the POSE test, this is offset by the final adjustment. In the pilot version of the POSE test, every student proceeded through the POSE test in the same order; starting with the vowel section, then the consonant section, and on to word stress, intonation, and sentence stress. In the newer version of the POSE test, students are free to choose the section they would like to complete. They have the option to complete all the sections at once, or do each one separately. This allows teachers and learners more freedom in using the POSE test to diagnose problems in English speech perception. This change was made based on the original design of the POSE test. The mandatory ordering of sections of the POSE test during the piloting stage was required to ensure that each section of the POSE test was piloted and received roughly the same amount of consideration by the participants involved. The original designs of the POSE test included this feature of being able to choose and complete any particular section without regard to the other sections of the test.

Currently, the amount of time required to complete each section of the POSE test is not known because no one has used the newer version of the POSE test; however estimates range from about one to two hours, depending on the section being completed.

Future Changes

Other changes will be made in the future to help improve the POSE test. One future change is improving the quality of the audio recordings used in the POSE test. Currently, while audible and understandable, many of the audio recordings were not

produced on professional equipment. Rerecording the audio portions of the POSE test in a sound-proof recording studio with professional equipment would greatly improve the quality of the POSE test and possibly its ability to accurately diagnose speech perception problems.

Also, changes need to be made in the way the POSE test reports problem areas in speech perception. Due to some time constraints in the development and pilot of the POSE test, feedback and data from teachers and learners regarding the usefulness of the results in designing specific instruction, as well as the comprehensibility of the results generated were not available. Currently, the POSE test outputs a list of items the students marked incorrectly, but this list does not provide adequate indication of specific contrasts or suprasegmental patterns. Different output formats should be tested for their usability by teachers in designing pronunciation instruction, as well as their comprehensibility. The appropriate format should then be chosen to help teachers and learners make the most of the diagnostic results output by the POSE test.

Another possible future addition to the POSE test involves the integration of other pronunciation and perception oriented instructional materials with the results of the POSE test. This would enable learners to complete a section of the POSE test, and their results would generate a list of materials or exercises learners could study to help improve their perception in problem areas. One example of a possible suggested material is the book *Pronunciation Matters* (Henrichsen, Green, Nishitani, & Bagley, 1999). Based on the results of the POSE test, teachers and/or learners could be directed to a specific unit or lesson in this book. *Pronunciation Matters* is more suited to this type of

integration because some of the items in the POSE test are based on text and illustrations from this book.

This does not mean, however, that the results of the POSE test could only be used in this way. Other materials or suggestions may be just as appropriate for integration with the POSE test in this manner. It is even possible that a series of perception and/or production exercises could be added to the POSE test, independent of any other material or text. Learners could then use those exercises to practice and improve their skills in problem areas identified by each section of the POSE test.

Future Research Plans

During the course of this project, and the documentation of this project report, several plans for future research were considered. One very important plan for future research involves piloting the newer version of the POSE test with a group of non-native speakers of English. These future pilots would be conducted specifically to obtain data that could help establish the overall reliability and validity of the POSE test. With a larger group of participants, and more complete data, more accurate statistical analyses could be performed. For example, one pilot could include multiple administrations of the POSE test to the same group of participants. The data collected could then be analyzed to help determine the reliability of each section of the POSE test, as well as the overall reliability of the POSE test as a whole. Another example of a future pilot might be to administer the POSE test to a large homogenous group of non-native speakers of English. With enough participants in the pilot group, more definite conclusions could be drawn as to the ability of the POSE test to pinpoint problematic contrasts for each language group. Certainly

further work needs to be done in this area. The author has received requests from other institutions for permission to help with further research and piloting, but no definite plans have been made. It is expected, however, that another pilot of the POSE test will be conducted in the near future.

CHAPTER 8

Conclusions and Implications

The completion of this project involved the design, production, piloting, evaluation, and revision of the *Perception of Spoken English* (POSE) *Test*. The POSE test is based on research that indicates a complex and important relationship between speech perception and production (see Chan, 2001; Flege & Eefting, 1987; Ingram & Park, 1997; Llisterri, 1995). The research also shows that by isolating and identifying specific problem areas in speech perception, and then focusing specific training on those areas, speech perception abilities can be improved (Borden, Gerber & Milsark, 1983; Cenoz & Lecumberri, 1999; Underbakke, 1993). Because of this relationship between speech perception and production, and the importance of self-monitoring in overcoming speech production problems (Acton, 1984; Morley, 1991), teachers and learners can benefit from knowing the specific sounds and patterns learners have trouble perceiving correctly. Teachers can use the results of the POSE test to then design specific curriculum to help their learners overcome the problems identified. Learners can use the results of the POSE test to raise their awareness of speech perception difficulties and focus on these difficulties and overcoming them in their own study and interaction in English.

The POSE test differs from other speech perception diagnostic materials in some important ways. First, the POSE test incorporates a large number of items for each aspect of speech perception included in the test. Second, the POSE test is computer based, requiring less teacher time and resources, and allowing more people to utilize the POSE test in different ways. Finally, the POSE test includes both segmental and suprasegmental

aspects of speech perception, providing teachers and learners with a much clearer picture of speech perception problems.

The pilot of the POSE test provided data that could be analyzed to help determine the reliability and validity of this instrument. The data, while insufficient to provide any significant, empirical evidence, do show several trends which indicate that were more data available, the POSE test would be shown to be a reliable and valid instrument for diagnosing problems in speech perception. Hence, plans for future research include designs to obtain more data in this area. Some of these plans include doing specific studies where the POSE test is administered to the same group of participants multiple times, and the data compared to see if the POSE test diagnoses the same problems for the same group of people. Other studies might include administering the POSE test to a large group of participants from the same language background and analyzing the data to determine if the POSE test identifies the problems that contrastive analysis tells us might exist. These types of studies will help us understand the ways the POSE test can be used to help learners communicate more effectively in English.

<div align="center">Benefits of the *Perception of Spoken English* (POSE) *Test*</div>

The POSE test offers many benefits to the language learning and research community. First, as has been explained in detail earlier in this project report, learners can benefit from knowing specific problem areas they have in speech perception. Some of these benefits include improved listening comprehension and the ability to recognize significant mistakes in one's own speech production. Other benefits include a heightened

awareness of speech perception overall and possibly heightened motivation to improve in the specific problem areas identified by the POSE test.

It is not known specifically how teachers and learners will use the results of the POSE test to improve speech perception abilities. Perhaps just being more aware of specific problem areas will foster improvement. Perhaps training in the perception of specific phonemic and suprasegmental contrasts will improve speech perception skills. Due to the complex relationship between perception and production, it is possible that teachers may be able to use the results of the POSE test to help design specific lessons and curriculum that will help learners improve both their perception and production of spoken English. One of the future changes or additions suggested in Chapter 7 involves the integration of perception and production exercises that could be suggested to learners based on the problems identified in the results of the POSE test. Another possible future change already discussed is the addition of a list of suggested pronunciation/perception instructional materials that teachers could use to help in planning lessons focused on specific pronunciation problems. These are just some of the possible benefits of using the POSE test in pronunciation instruction.

Another area where the POSE test might prove useful is in research. Many of the studies cited in Chapter 2 of this report endeavored to show a relationship between perception and production. The POSE test could be used as a research instrument in a study that examined this relationship more closely. The POSE test could also be used to provide empirical evidence regarding language specific problems identified through contrastive analysis. For example, the data from the pilot of the POSE test showed that many native Spanish speakers had trouble perceiving the /z/ sound in English. This is

consistent with the data derived from contrastive analysis of Spanish and English. The POSE test could be used to discover more of these types of problems and provide empirical evidence as to the extent of their influence of the pronunciation of English learners from that native language group. Finally, the POSE test could be used to examine the effectiveness of different teaching methods at improving speech perception abilities. The POSE test could be administered to a group of participants, followed by several different types of training sessions. Then the participants could take the POSE test a second time to measure their progress. These are just some of the benefits of the POSE test to the research community.

Another benefit not previously mentioned is the possibility of using the POSE test as a template for diagnosing speech perception problems in languages other than English. Throughout this project report, references to the benefits and principles of the POSE test as they apply to English have been plentiful. Yet, the principles mentioned in Chapter 2 regarding the relationship between speech perception and production, as well as the principles of diagnosing problem areas in speech perception utilizing minimal pairs can be applied to other languages besides English. The content of these diagnostic tests would of course be different than the POSE test, and they would be based on the features of the language being used in the diagnostic test. However, the design and techniques used in the POSE test could readily be applied. This is a major benefit for teachers of languages where pronunciation materials are scarce.

Lessons Learned

Besides the benefits to the language community, the development of the POSE test during the course of this project has provided several benefits to the author. One of those benefits was the reiteration of the importance of constant evaluation during the materials development process. Many times in a materials development project, the design process is not linear, but more circular. In other words, evaluation is done at the beginning, middle, and end of the design process, and changes are made based on the results of that evaluation at any stage along the way. This proved true with the design of the POSE test. One area where this was most evident was in the design of the word stress section. Originally, the items consisted of sentences with a single word removed. This word was then broken into syllables and presented as possible responses for the user to select. However, as was mentioned in Chapter 4 of this report, it was discovered that this item type allowed learner knowledge of grammar and other aspects of English separate from speech perception to influence or aid students in responding to the items. As a result, the items for the word stress section were redesigned to eliminate this problem.

The development and piloting of the POSE test also required knowledge of test construction and evaluation, including important concepts such as reliability and validity. The author's understanding of these concepts was renewed during the evaluation and analysis of data obtained during the pilot stage of this project. Innovative solutions were required to overcome difficulties not previously anticipated in determining the overall reliability and validity of the POSE test.

Another, less obvious benefit that completing this project provided the author was a deeper understanding of the importance of research and the necessity of well-structured

and carefully controlled research designs. With the help of colleagues and others, many common research pitfalls were avoided in this project. Completing the process of getting the pilot of this project approved by the Institutional Review Board (IRB) of Brigham Young University (BYU) reiterated the importance of obtaining permission from research participants and maintaining confidentiality of personal data. The disregard of these important research principles by researchers has the potential of hindering further research at institutions like BYU. The disregard of research principles also has the potential of providing inaccurate research results, which can mean wasted resources and time. By going through the research process, these principles were impressed upon my mind and will remain at the forefront of future research I do, both with the POSE test, and in other areas.

Most of all, the development of the POSE test, and the completion of this project required hard work and determination. These qualities, along with the lessons learned as mentioned above, helped to produce an improved teacher of English and researcher, and will provide essential experience and background for future endeavors.

REFERENCES

Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly, 18* (1), 71-85.

American Council of Learned Societies. (1952). *Structural notes and corpus: A basis for the preparation of materials to teach English as a foreign language.* Washington, D.C.: American Council of Learned Societies.

Banathy, B. H. & Madarasz, P. H. (1969). Contrastive analysis and error analysis. *Journal of English as a Second Language, 4* (3), 77-92.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research.* Timonium, MD: York Press.

Borden, G., Gerber, A. & Milsark, G. (1983). Production and perception of the /r/ - /l/ contrast in Korean adults learning English. *Language Learning, 33*, 499-526.

Catford, J.C. (1987). Phonetics and the teaching of pronunciation. In Morley, J. (Ed.) *Current perspectives on pronunciation: Practices anchored in theory* (pp. 83-100). Washington, D.C.: TESOL.

Celce-Murcia, M., Brinton, D. M. & Goodwin, J. M. (1996). *Teaching pronunciation.* Cambridge, UK: Cambridge University Press.

Cenoz, J. & Lecumberri, L. G. (1999). The effect of training on the discrimination of English vowels. *International Review of Applied Linguistics, 36*, 261-275.

Champagne-Muzar, C., Schneiderman, E. I. & Bourdages, J. S. (1993). Second language accent: The role of the pedagogical environment. *International Review of Applied Linguistics, 31*, 143-160.

Chan, C. (2001). The perception (and production) of English word-initial consonants by native speakers of Cantonese. *Hong Kong Journal of Applied Linguistics, 6*, 26-44.

Chang, J. (2001). Chinese speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed., pp. 310–324). Cambridge, UK: Cambridge University Press.

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Coe, N. (2001). Speakers of Spanish and Catalan. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed., pp. 90-112). Cambridge, UK: Cambridge University Press.

Crystal, D. (1997). *The Cambridge Encyclopedia of the English Language*. Cambridge, UK: Cambridge University Press.

de Bot, K. (1983). Visual feedback of intonation I: Effectiveness and induced practice behavior. *Language and Speech, 26*, 331-350.

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48,* 393-410.

Fayer, J. & Krasinski, E. (1987). Native and non-native judgments of intelligibility and irritation. *Language Learning, 37*, 313-326.

Flege, J. E. (1987). A critical period for learning to pronounce foreign languages? *Applied Linguistics, 8*, 162-177.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience.* Timonium, MD: York Press.

Flege, J. E. & Eefting, W. (1987). Production and perception of English stops by native

   Spanish speakers. *Journal of Phonetics, 15*, 67-83.

Ferguson, C. & Macken, M. (1980). Phonological development in children: Play and

   cognition. *Child Language Development, 18*, 138-177.

Fries, C. C. (1945). *Teaching & learning English as a foreign language* (pp. 10-26). Ann

   Arbor, MI: University of Michigan Press.

Gilbert, J. (1993). *Clear speech: Pronunciation and listening comprehension in North*

   American English (2nd ed., Teacher's Resource Book)*.* New York: Cambridge

   University Press.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds 'L' and

   'R.' *Neuropsychologia, 9*, 317-323.

Gumperz, J. (1982). *Discourse strategies*. Cambridge, UK: Cambridge University Press.

Henrichsen, L. E., Green, B. A., Nishitani, A. & Bagley, C. L. (1999) *Pronunciation*

   matters. Ann Arbor, MI: University of Michigan Press.

Hewlett, N. (1990). Processes of development and production. In P. Grunwell (Ed.),

   Developmental Speech Disorders. Edinburg: Churchill Livingstone.

Holden, K. & Hogan, J. (1993). The emotive impact of foreign intonation: An experiment

   in switching English and Russian intonation. *Language and Speech, 36*, 67-88.

Howatt, A. P. R. (1984). *A history of English language teaching.* Oxford: Oxford

   University Press.

HTML Overview. (2004). Retrieved June 16, 2004, from

   http://www.learnthenet.com/english/html/46html.htm

Hughes, A. (1989). *Testing for language teachers*. Cambridge, UK: Cambridge
 University Press.

Ingram, J. C. L. & Park, S. G. (1997). Cross-language vowel perception and production
 by Japanese and Korean learners of English. *Journal of Phonetics, 25*, 343-370.

Jamieson, D. G. & Morosan, D. E. (1986). Training non-native speech contrasts in adults:
 Acquisition of the English /ð/-/θ/ contrast by francophones. *Perception &
 Psychophysics, 40*, 205-215.

Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA
 Model of how speech perception develops. *Journal of Phonetics, 21*, 3-28.

Kelly, L. G. (1969). *25 centuries of language teaching*. Rowley, MA: Newbury House.

Krashen, S. D. (1973). Lateralization, language learning, and the critical period: Some
 new evidence. *Language Learning, 23*, 63-74.

Kuhl, P. K. (1991). Human adults and human infants show a "perceptual magnet effect"
 for the prototypes of speech categories, monkeys do not. *Perception and
 Psychophysics, 50*, 93-107.

Kuhl, P. K. & Iverson, P. (1995). Linguistic experience and the "perceptual magnet
 effect." In W. Strange (Ed.), *Speech perception and linguistic experience: Issues
 in cross-language research* (pp. 121-154). Timonium, MD: York Press.

Lado, R. (1961). *Language testing*. London: Longman.

Lane, H. (1963). Foreign accent and speech distortion. *Journal of the Acoustical Society
 of America, 35*, 451-453.

Leather, J. & James, A. (1991). The acquisition of second language speech. *Studies in
 Second Language Acquisition, 13*, 305-341.
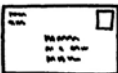
Lee, J. (2001). Korean speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed., pp. 325-342). Cambridge, UK: Cambridge University Press.

Lenneberg, E. (1967). *Biological foundations of language.* New York: Wiley.

Llisterri, J. (1995). Relationships between speech production and speech perception in a second language. In Elenius, K & Branderud, P. (Eds.) *Proceedings of the XIIIth International Congress of Phonetic Sciences Vol. 4* (pp. 92-99). Stockholm: KTH/Stockholm University. Retrieved May 24, 2004, from http://liceu.uab.es/~joaquim/publicacions/Prod_Percep.html

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. L., Jenkins, J. J., & Fujimura, O. (1975). An effect of linguistic experience: the discrimination of r and l by native speakers of Japanese and English. *Perception and Psychophysics, 18*, 331-340.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*, 481-520.

Morley, J. (1994). Multidimensional curriculum design for speech-pronunciation instruction. In J. Morley (Ed.), *Pronunciation pedagogy and theory. New views, new directions* (pp. 64-91). Alexandria, VA: TESOL.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Nilsen, D.L.F. & Nilsen, A.P. (2002). *Pronunciation contrasts in English.* Long Grove, IL: Waveland Press.

Orion, G. F. (1997). *Pronouncing American English: Sounds, stress, and intonation.* Pacific Grove, CA: Heinle & Heinle.
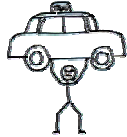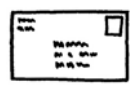
Pennington, M. C. (1993, August). Breaking the L1 phonological habit. AILA World Congress. Amsterdam, Netherlands.

Pennington, M. C. (1998). The teachability of phonology in adulthood: A re-examination. *International Review of Applied Linguistics (IRAL), 36*, 323-341.

Purcell, E. T. & Suter, R. W. (1981). Predictors of pronunciation accuracy: A re-examination. *Language Learning, 30*, 271-288.

Richards, J. C. & Rogers, T. S. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge: Cambridge University Press.

Robinett, B. W. & Schacter, J. (Eds.). (1983). *Second language learning: Contrastive analysis, error analysis, and related aspects*. Ann Arbor, MI: University of Michigan Press.

Schacter, J. (1986). In search of systematicity in interlanguage production. *Studies in Second Language Acquisition, 8*, 119-134.

Schneiderman, E., Bourdages, J. & Champagne, C. (1988). Second-language accent: The relationship between discrimination and perception in acquisition. *Language Learning, 38*, 1-19.

Scovel, T. (1969). Foreign accent: Language acquisition and cerebral dominance. *Language Learning, 19*, 245-254.

Seliger, H. W., Krashen, S. D. & Ladefoged, P. (1975). Maturational constraints in the acquisition of second language accent. *Language Sciences, 36*, 20-22.

The Sigma Institute. (n.d.) *History of Hangul.* Retrieved May 24, 2004, from http://www.sigmainstitute.com/koreanonline/hangul_history.shtml
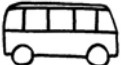
Strange, W. & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics, 36*, 131-145.

Sudlow, M. & Bischof, R. (1994). *Exercises in American English pronunciation.* Monmouth, OR: Excellence in Education.

Suter, R. W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning, 26*, 233-253.

Swan, M. & Smith, B. (Eds.) (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge, UK: Cambridge University Press.

Thompson, I. (2001). Japanese speakers. In M. Swan & B. Smith (Eds.) *Learner English: A teacher's guide to interference and other problems* (2nd ed., pp. 296-309). Cambridge, UK: Cambridge University Press.

Underbakke, M. E. (1993). Hearing the difference: Improving Japanese students' pronunciation of a second language through listening. *Language Quarterly, 31*, 67-89.

Wardhaugh, R. (1970). The contrastive analysis hypothesis. *TESOL Quarterly, 4*, 123-130.

Werker, J. F. & Polka, L. (1993). Developmental changes in speech perception: new challenges and new directions. *Journal of Phonetics, 21*, 83-101.

APPENDIX A

Items in the Vowel Section

Contrast	Item #	Sentence 1	Sentence 2
/i/ - /ɪ/	1	Don't sleep on the deck.	Don't slip on the deck.
	2	Mr. Green was beaten.	Mr. Green was bitten.
	3	You must heat it.	You must hit it.
	4	Look at the sheep.	Look at the ship.
/ɪ/ - /ɛ/	5	The spaghetti sauce is bitter.	The spaghetti sauce is better.
	6	I need a pin for the message.	I need a pen for the message.
	7	Will you please pick up that litter for me?	Will you please pick up that letter for me?
	8	They're picking the fruit.	They're pecking the fruit.
/ɛ/ - /e/	9	I'd like to sell the boat.	I'd like to sail the boat.
	10	Put that chair in the shed.	Put that chair in the shade.
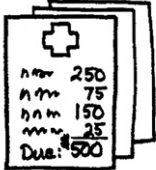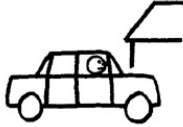
Contrast	Item #	Sentence 1	Sentence 2
/ɛ/ - /e/	11	He held the cab for me.	He hailed the cab for me.
	12	He looked at the mess on the floor.	He looked at the mace on the floor.
/ɛ/ - /æ/	13	This pen leaks.	This pan leaks.
	14	Where is the letter?	Where is the ladder?
	15	They had to pedal the boat.	They had to paddle the boat.
	16	The men will be here soon.	The man will be here soon.
/æ/ - /ɑ/	17	Jacob took good care of his axe.	Jacob took good care of his ox.
	18	That's my sack.	That's my sock.
	19	He was hurt when he hit the rack.	He was hurt when he hit the rock.
	20	He sat on his cat.	He sat on his cot.

Contrast	Item #	Sentence 1	Sentence 2
/ɑ/ - /ɑr/	21	It was a big shock.	It was a big shark.
	22	The gods were angry.	The guards were angry.
	23	Look at that cot.	Look at that cart.
	24	There was a lock on the box.	There was a lark on the box.
/ɑ/ - /ʌ/	25	Kevin ran after the boss.	Kevin ran after the bus.
	26	Doug caught the big fish.	Doug cut the big fish.
	27	I don't like the collar.	I don't like the color.
	28	I need a cop.	I need a cup.
/ɑ/ - /ow/	29	Look at the clock.	Look at the cloak.
	30	How did you like my fox?	How did you like my folks?
	31	That cot is too small.	That coat is too small.
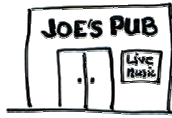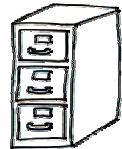
Contrast	Item #	Sentence 1	Sentence 2
/ɑ/ - /ow/	32	He slept under the ox.	He slept under the oaks.
/ʌ/ - /ɚ/	33	Shirley enjoys looking at the buds.	Shirley enjoys looking at the birds.
	34	Look at the gull.	Look at the girl.
/ʌ/ - /ʊ/	35	Steve needed two bucks.	Steve needed two books.
	36	Jeremy putts the golf ball.	Jeremy puts the golf ball.
/ʊ/ - /uw/	37	Look at this soot!	Look at this suit!
	38	The sign says, "Pull."	The sign says, "Pool."
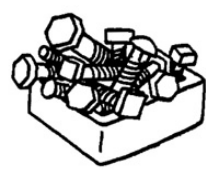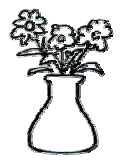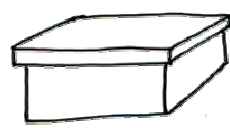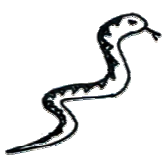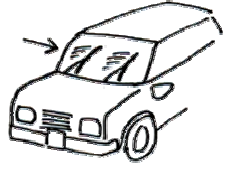
APPENDIX B
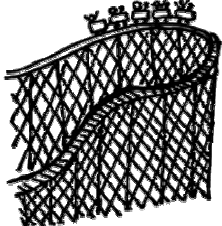
Items in the Consonant Section
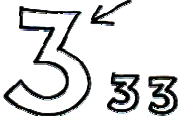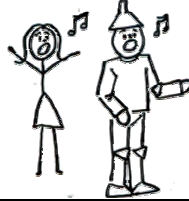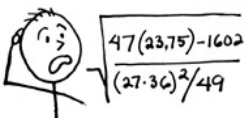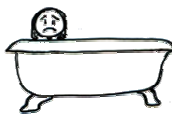
Contrast	Item #	Sentence 1	Sentence 2
/p/ - /b/ (Initial)	1	Those pills are very large!	Those bills are very large!
	2	Pete caught a fantastic pass.	Pete caught a fantastic bass.
	3	That was not a good place to park.	That was not a good place to bark.
	4	Look at that pear.	Look at that bear.
/p/ - /b/ (Final)	5	A robber stole my uncle's cap.	A robber stole my uncle's cab.
	6	I need a new rope.	I need a new robe.
	7	Put this in your lap.	Put this in your lab.
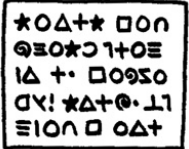
Contrast	Item #	Sentence 1	Sentence 2
/p/ - /f/ (Initial)	8	He bought a new pup.	He bought a new pub.
	9	She needed a pan.	She needed a fan.
	10	I need some pins.	I need some fins.
	11	The paper was lost in the pile.	The paper was lost in the file.
	12	He had never eaten pigs before.	He had never eaten figs before.
/p/ - /f/ (Final)	13	Your cup is dirty!	Your cuff is dirty!
	14	Look at that clip.	Look at that cliff.
	15	That is a small leap.	That is a small leaf.
	16	You cannot snip my flowers.	You cannot sniff my flowers.

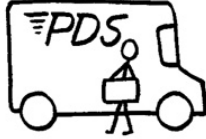| Contrast	Item #	Sentence 1	Sentence 2
/v/ - /b/ (Initial)	17	She's voting.	She's boating.
	18	That's a lot of volts!	That's a lot of bolts!
	19	She has a veil.	She has a bale.
	20	Use that for the vase.	Use that for the base.
/v/ - /w/ (Initial)	21	That's verse.	That's worse.
	22	It's in the vest.	It's in the west.
	23	That's a strong vine.	That's a strong wine.
	24	Bad vipers are dangerous.	Bad wipers are dangerous.
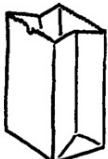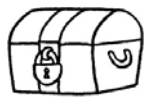
Contrast	Item #	Sentence 1	Sentence 2
/f/ - /v/ (Initial)	25	I want to buy the fan.	I want to buy the van.
	26	He asked her to change the fee.	He asked her to change the "V".
	27	It's a type of fowl.	It's a type of vowel.
	28	His farm had foals.	His farm had voles.
/f/ - /θ/ (Initial)	29	They fought about it.	They thought about it.
	30	My first is terrible.	My thirst is terrible.
	31	It was the frill of it.	It was the thrill of it.
	32	He was free.	He was three.

Contrast	Item #	Sentence 1	Sentence 2
/θ/ - /t/ (Initial)	33	Teresa thought about Abraham Lincoln.	Teresa taught about Abraham Lincoln.
	34	Please don't touch my thigh.	Please don't touch my tie.
	35	That's a big three!	That's a big tree!
	36	She sang with the thin man.	She sang with the tin man.
/θ/ - /t/ (Final)	37	It was just an old myth.	It was just an old mitt.
	38	He's just having a fifth.	He's just having a fit.
	39	The math was too hard.	The mat was too hard.
	40	She didn't like the bath.	She didn't like the bat.

Contrast	Item #	Sentence 1	Sentence 2
/θ/ - /s/ (Initial)	41	Paul didn't want to think.	Paul didn't want to sink.
	42	That's a good thimble.	That's a good symbol.
/θ/ - /s/ (Final)	43	Ken tried not to lose faith.	Ken tried not to lose face.
	44	That's a big mouth.	That's a big mouse.
/ð/ - /d/ (Initial)	45	Don't worry, they will come soon.	Don't worry, day will come soon.
	46	It is forbidden to those in class.	It is forbidden to doze in class.
/t/ - /d/ (Final)	47	We need a cart.	We need a card.
	48	Have you seen this coat before?	Have you seen this code before?

Contrast	Item #	Sentence 1	Sentence 2
/t/ - /d/ (Final)	49	He gave me his seat.	He gave me his seed.
	50	Look at the horse's feet.	Look at the horse's feed.
/n/ - /l/ (Initial)	51	Nikki likes naps.	Nikki likes laps.
	52	He just nicked it.	He just licked it.
	53	That's a nine.	That's a line.
	54	It's night outside.	It's light outside.
/n/ - /l/ (Final)	55	The dog wants its bone.	The dog wants its bowl.
	56	Where's the spoon?	Where's the spool?
	57	This is Ann.	This is Al.

Contrast	Item #	Sentence 1	Sentence 2
/n/ - /l/ (Final)	58	He likes to spin.	He likes to spill.
/l/ - /r/ (Initial)	59	Rebecca's answer was long.	Rebecca's answer was wrong.
	60	They found a lake.	They found a rake.
	61	I need a big lock.	I need a big rock.
	62	There are lamps in the tomb.	There are ramps in the tomb.
/s/ - /z/ (Final)	63	She looked at my ice.	She looked at my eyes.
	64	He likes peace.	He likes peas.
	65	He was surprised by the price.	He was surprised by the prize.

Contrast	Item #	Sentence 1	Sentence 2
/s/ - /z/ (Final)	66	The sauce cost too much.	The saws cost too much.
/s/ - /ʃ/ (Initial)	67	Can you sip it quickly?	Can you ship it quickly?
	68	Look at that sack.	Look at that shack.
/ʃ/ - /tʃ/ (Initial)	69	He hurt his shin when he fell.	He hurt his chin when he fell.
	70	Sherry shows her software at conventions.	Sherry chose her software at conventions.
/ʃ/ - /tʃ/ (Final)	71	The captain made sure to lash the chest.	The captain made sure to latch the chest.
	72	Please wash the car.	Please watch the car.

Contrast	Item #	Sentence 1	Sentence 2
/tʃ/ - /dʒ/ (Initial)	73	He's choking.	He's joking.
	74	She's cheering.	She's jeering.
/dʒ/ - /y/ (Initial)	75	He went to jail.	He went to Yale.
	76	Please pass the jam.	Please pass the yam.
/k/ - /g/ (Initial)	77	My uncle has a new coat.	My uncle has a new goat.
	78	They looked at the coast.	They looked at the ghost.
	79	They looked at the cards.	They looked at the guards.
	80	He loves curls.	He loves girls.
/k/ - /g/ (Final)	81	We needed a tack for each picture.	We needed a tag for each picture.

Contrast	Item #	Sentence 1	Sentence 2
/k/ - /g/ (Final)	82	This dock is old.	This dog is old.
	83	Look at that buck!	Look at that bug!
	84	That's my pick.	That's my pig.

APPENDIX C

Items in the Word Stress Section

No. of Syllables	Item Number	Possible Choices
2	1	PRO·ject pro·JECT
	2	CON·flicts con·FLICTS
	3	per·MIT PER·mit
	4	pro·TEST PRO·test
	5	SUR·vey sur·VEY
	6	re·CORD RE·cord
	7	RE·bel re·BEL
	8	PRE·sent pre·SENT
	9	in·CLINE IN·cline
	10	CON·tract con·TRACT
3	11	AD·vo·cate ad·VO·cate ad·vo·CATE
	12	al·ter·NATE al·TER·nate AL·ter·nate
	13	DEL·e·gate del·e·GATE del·E·gate

No. of Syllables	Item Number	Possible Choices
3	14	IN·ti·mate in·TI·mate in·ti·MATE
	15	SEP·a·rate sep·A·rate sep·a·RATE
	16	GRAD·u·ate grad·U·ate grad·u·ATE
	17	es·ti·MATE ES·ti·mate es·TI·mate
	18	AG·gre·gate ag·GRE·gate ag·gre·GATE
	19	O·ver·flow o·VER·flow o·ver·FLOW
	20	des·o·LATE DES·o·late des·O·late
4	21	ap·PRO·pri·ate AP·pro·pri·ate ap·pro·PRI·ate ap·pro·pri·ATE
	22	AP·prox·i·mate ap·PROX·i·mate ap·prox·I·mate ap·prox·i·MATE
	23	de·LIB·er·ate DE·lib·er·ate de·lib·ER·ate de·lib·er·ATE
	24	E·lab·o·rate e·LAB·o·rate e·lab·o·RATE e·lab·O·rate

No. of Syllables	Item Number	Possible Choices
4	25	PRE·cip·i·tate pre·cip·i·TATE pre·CIP·i·tate pre·cip·I·tate
	26	re·CRE·a·tion RE·cre·a·tion re·cre·a·TION re·cre·A·tion
	27	PHO·tog·ra·pher pho·tog·ra·PHER pho·TOG·ra·pher pho·tog·RA·pher
	28	dip·lo·MAT·ic dip·LO·mat·ic DIP·lo·mat·ic dip·lo·mat·IC
	29	con·GLOM·er·ate con·glom·er·ATE CON·glom·er·ate con·glom·ER·ate
	30	ac·TIV·i·ty AC·tiv·i·ty ac·tiv·I·ty ac·tiv·i·TY
5	31	EC·o·nom·i·cal ec·o·NOM·i·cal ec·O·nom·i·cal ec·o·nom·i·CAL ec·o·nom·I·cal
	32	ap·PRO·pri·ate·ly AP·pro·pri·ate·ly ap·pro·pri·ate·LY ap·pro·PRI·ate·ly ap·pro·pri·ATE·ly

No. of Syllables	Item Number	Possible Choices
5	33	sim·pli·fi·ca·TION SIM·pli·fi·ca·tion sim·PLI·fi·ca·tion sim·pli·FI·ca·tion sim·pli·fi·CA·tion
	34	com·mu·NI·ca·tive com·MU·ni·ca·tive COM·mu·ni·ca·tive com·mu·ni·ca·TIVE com·mu·ni·CA·tive
	35	con·sid·er·A·ble CON·sid·er·a·ble con·SID·er·a·ble con·sid·er·a·BLE con·sid·ER·a·ble
	36	math·E·mat·i·cal math·e·MAT·i·cal MATH·e·mat·i·cal math·e·mat·i·CAL math·e·mat·I·cal
	37	con·sec·u·TIVE·ly con·SEC·u·tive·ly CON·sec·u·tive·ly con·sec·u·tive·LY con·sec·U·tive·ly
	38	an·NI·ver·sa·ry an·ni·ver·SA·ry an·ni·VER·sa·ry AN·ni·ver·sa·ry an·ni·ver·sa·RY
	39	OR·gan·i·za·tion or·gan·I·za·tion or·gan·i·ZA·tion or·gan·i·za·TION or·GAN·i·za·tion

No. of Syllables	Item Number	Possible Choices
5	40	IN·de·ter·mi·nate
		in·DE·ter·mi·nate
		in·de·TER·mi·nate
		in·de·ter·MI·nate
		in·de·ter·mi·NATE

APPENDIX D

Items in the Intonation Section

Item #	Sentence
1	You are teaching class today
2	He thought the movie was good
3	She can't hear me
4	They bought five gallons of juice for the party
5	John cooked the dinner by himself
6	The phone isn't working
7	It's snowing in Phoenix, Arizona
8	They missed their connecting flight to Honolulu
9	Jeff's leaving tomorrow
10	He got yelled at by his teacher
11	He passed the TOEFL exam
12	Mandi likes chocolate
13	Your parents are coming today
14	We're going camping this weekend
15	John went home already
16	We have two pages of homework
17	Mary's going to have a baby
18	Paul's taking Jessica out tonight
19	She works on campus
20	Justin liked the novel
21	You're in biology class, aren't you?
22	You didn't call me last night, did you?
23	Mom cooked dinner, didn't she?
24	Our baseball team will win their final game, won't they?
25	She won't be here next semester, will she?
26	He should be here by six o'clock, shouldn't he?
27	They see each other everyday, don't they?

Item #	Sentence
28	You can finish that sandwich, can't you?
29	They just arrived home from their trip to Europe, didn't they?
30	He is going to call me this evening, isn't he?
31	You're going to the party, aren't you?
32	You have seen the movie, haven't you?
33	He won't be mad at us, will he?
34	She can't come tonight, can she?
35	He studies very hard, doesn't he?
36	She's studying Japanese, isn't she?
37	They don't like each other, do they?
38	You like chocolate, don't you?
39	He's making lunch for us, isn't he?
40	They're leaving tonight, aren't they?

Appendix E

Items in the Sentence Stress Section

Item #	Sentence 1	Sentence 2
1	They bought three <u>shirts</u>. (not pants)	They bought <u>three</u> shirts. (not two)
2	Does <u>he</u> speak English? (no, but she does)	Does he speak <u>English</u>? (no, but he speaks French)
3	She wants to leave on <u>Thursday</u>. (not Friday)	She wants to <u>leave</u> on Friday. (not arrive)
4	I think he goes to <u>Harvard</u>. (not Princeton)	I <u>think</u> he goes to Harvard. (I'm not sure)
5	It's <u>John's</u> job to wash the dishes. (not Sara's job)	It's John's job to wash the <u>dishes</u>. (not the car)
6	We told <u>her</u> we were going to be late. (not him)	We told her we were going to be <u>late</u>. (not early)
7	<u>Mr.</u> Webb rides his bike to work. (not Mrs. Webb)	Mr. Webb rides his bike to <u>work</u>. (not to school)
8	I like <u>chocolate</u> ice cream. (not vanilla)	I like chocolate <u>ice cream</u>. (not pie)
9	I'm sorry. We ordered <u>two</u> hamburgers. (not one)	I'm sorry. We ordered two <u>hamburgers</u>. (not hot dogs)
10	The <u>cat</u> caught a bird. (not the dog)	The cat caught a <u>bird</u>. (not a mouse)
11	The <u>dog</u> is under the table. (not the cat)	The dog is under the <u>table</u>. (not the chair)
12	He bought a <u>grammar</u> book. (not a writing book)	He bought a grammar <u>book</u>. (not a grammar tape)
13	He broke his <u>computer</u>. (not his television)	He broke <u>his</u> computer. (not her computer)
14	He stole the red <u>Ferrari</u>. (not the red Cadillac)	He stole the <u>red</u> Ferrari. (not the blue one)
15	He didn't finish his <u>homework</u>. (but he finished his chores)	<u>He</u> didn't finish his homework. (but she did)
16	He doesn't like to <u>study</u>. (but he likes to play)	<u>He</u> doesn't like to study. (but she does)
17	He played <u>soccer</u> last weekend. (not basketball)	He <u>played</u> soccer last weekend. (but he didn't watch it)

Item #	Sentence 1	Sentence 2
18	He plays the piano <u>every</u> night. (not just once a week)	He plays the <u>piano</u> every night. (not the violin)
19	We couldn't find <u>Jim</u>, so we left. (but we could find Sally)	We couldn't find Jim, so we <u>left</u>. (we didn't stay)
20	I passed the <u>TOEFL</u>. (not the SAT)	I <u>passed</u> the TOEFL. (I got a good score)
21	The students like history. (The students like the subject.)	The students like his story. (The students like the story he told.)
22	Did she want a needle? (Did she want something to sew with?)	Did she want any doll? (Did she want something to play with?)
23	She likes eating pie and apples. (She likes eating two things.)	She likes eating pineapples. (She likes eating one thing.)
24	It was elementary. (It was basic.)	It was a lemon tree. (It grew lemons.)
25	Is this Europe? (Am I on the right continent?)	Is this your rope? (Can I borrow this?)
26	He sold his car stereo and golf clubs. (He sold two things.)	He sold his car, stereo, and golf clubs. (He sold three things.)
27	Shall we call, Tina? (Tina is going to call someone.)	Shall we call Tina? (They are going to call Tina.)
28	Have we met, Mr. Smith? (The person is talking to Mr. Smith.)	Have we met Mr. Smith. (The person is talking to someone else.)
29	Make sure to wash, Suzy. (Suzy is going to wash.)	Make sure to wash Suzy. (Someone else is going to wash Suzy.)
30	"Elizabeth," said John, "was late." (John was talking about Elizabeth.)	Elizabeth said, "John was late." (Elizabeth is talking about John.)
31	"Jerry," explained Susan, "was in the car also." (Susan is talking about Jerry.)	Jerry explained, "Susan was in the car also." (Jerry is talking about Susan.)
32	We are going to eat, Henry. (They will eat with Henry.)	We are going to eat Henry. (They will eat Henry.)
33	The movie is called "Fall and Summer". (Is it about the seasons?)	The movie is called "Fallen Summer". (Is it about disappointment?)
34	He fixed his houseboat and car. (He fixed two things.)	He fixed his house, boat and car. (He fixed three things.)
35	He bought ice, cream, and sugar at the store. (He bought three things.)	He bought ice cream and sugar at the store. (He bought two things.)

Item #	Sentence 1	Sentence 2
36	He placed the tea, cup, and saucer on the table. (He put three things on the table.)	He placed the teacup and saucer on the table. (He put two things on the table.)
37	I bought that horse, trailer, and saddle last week. (I bought three things.)	I bought that horse trailer and saddle last week. (I bought two things.)
38	Joan asked Mother, "What are we having for dinner?" (Joan is speaking.)	"Joan," asked Mother, "what are we having for dinner?" (Mother is speaking.)
39	Molly said, "Judy won't be going to the party." (Molly is speaking.)	"Molly," said Judy, "won't be going to the party." (Judy is speaking.)
40	Henry said, "The farmer at our chicken." (Henry is speaking.)	"Henry," said the farmer, "ate our chicken." (The farmer is speaking.)