



Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

An analysis of English majors speech perception problems

Shujing Zhang

Faculty of International Studies, Henan Normal University, Xinxiang, Henan, China

ABSTRACT

For English learners, listening has its special importance among the four language skills. While it is the basic form of communication and foundation for developing other language skills, listening is usually regarded as the most difficult part among all kinds of English examinations. Dictation tests are now used as an economical and valid method to investigate English learners' listening difficulties. It can test learners' ability in phoneme discrimination, mastery of vocabulary and grammar, etc. In the process of analysis, each sentence of the students' version is compared with the original version to find out the differences which can be seen as errors. Then all the errors are collected and classified based on their distinctive features. Through the explicit description of learners' speech perception error patterns and the application of SPSS to identify the correlation of listening and pronunciation, some suggestions for English listening teaching and learning are put forward based on the above analyses.

Key words: Speech Perception, Error Analysis, Error Patterns, Listening Teaching

INTRODUCTION

Listening is of primary importance in communication. According to two famous American language teaching experts W.M. Rivers and M.S. Temperly, in a person's whole life, one spends approximately 45% of his lifetime listening, 30% speaking, 16% reading, and 9% writing. For those who want to engage in any form of communication, it is quite necessary for them to understand and react to what has been said [1-5]. Listening is not only the basic form of communication between people, but also the foundation for developing other language skills. In the late 1960s and early 1970s, applied linguistics recognized that listening was the primary channel by which the learner could gain access to L2 data, and it therefore acted as the trigger for acquisition [6-9]. Vandergrift (1996; 1998) said that a key difference between more successful and less successful acquirers relates in large part to their ability to use listening as a means of acquisition [10-13]. According to Ebel, the strength of many developmental reading programs appears to lay their success in improving auditory discrimination of language sounds. Thus Morley thinks that "the importance of listening can't be underestimated; it is imperative that it not be treated trivially in second and foreign language curricula" [14-17].

However, listening is also one of the most difficult skills for language learners. Belasco once said "I was rudely jolted by the realization that it is possible to develop so-called speaking ability and yet be virtually incompetent in understanding the spoken language...(students) were learning to audio-comprehend certain specific dialogues and drills...but couldn't understand the language out of the mouths of native speakers". According to a survey by Shanghai Foreign Languages University in 2001, among the investigated 1457 college students, only 4.40% can understand normal speed English news, 4.74% can understand VOA Special News, 12.30% can understand the daily conversation of native English speakers, and 52.60% can understand their teachers' classroom English and 27.30% can just understand little daily English. And some statistics show that nearly 50% students regard English listening comprehension as the most difficult part among all kinds of examinations. The writer has also made a questionnaire to inquire into 32 Grade Two English Majors' difficulties in learning English. The result conforms to the above investigation discovery. 24 students of this group responded that they had problems in listening.

In spite of its importance and difficulty, listening is a Cinderella skill, and it has been alternatively overlooked. In different language teaching approaches, emphasis on listening was rare and in language teaching it was usually taught after speaking in most cases. In China, the situation of EFL (English as a foreign language) listening teaching is not promising, either. After years' study of English from junior middle school (and now from primary school almost all around the whole country), learners are still poor at listening. About English listening teaching in colleges, it is pointed out that listening instructors only play the tape-recorder back and forth, without checking students' comprehension and giving them help in the process. That is why listeners are often called as "tape-recorder". In the past few years, most attention was paid to how to overcome learners' difficulty in speaking. With the application of communicative language teaching, learners' oral ability has been significantly improved. Listening, on the other hand, is still seldom mentioned compared with other hot-discussed language skills. In addition, the government's inconsistent policy towards listening is another factor causing its state of being neglected. English listening gained its right of being a required and independent subject in College Entrance Examination several years ago. However, this regulation was cancelled last year. The compulsory stress on improving students' listening ability has disappeared. To help them get high marks in examinations, English teachers only care whether their students have mastered all the grammatical points in the texts or whether they have good writing skills. Unavoidably, the nickname "deaf English learners" is still used nowadays.

SPEECH PERCEPTION ANALYSIS: THEORETICAL CONSIDERATIONS

The definition of speech perception

In Harley's (2000) words, speech perception and spoken word recognition are two different categories. The first is about how we identify or perceive the sounds of language while the latter is about recognizing the words which are composed of sounds. Actually, concerning their difference or order in listening discrimination, there is still no consensus on whether speech identification should necessarily be prior to spoken word identification. Savin and Bever (1970) tried to prove that syllable is the fundamental and meaningful unit in speech perception, while based on some experimental evidence, Marslen-Wilson and Warren (1994) held that phoneme classification and lexicon activation share certain simultaneous process, because they argued that lexical representations can be directly gained on the basis of the featural information in the sound signal (cited in Harley, 2000). In this paper, the term speech perception or sound perception is used to denote the process of both phoneme and word comprehension and may be used changeably with word recognition for the convenience of the context.

The process of speech perception

According to Miller & Eimas (1994), there is no theory in speech perception, because the only detailed evidence is the evidence from psychophysical studies. To have a better understanding of speech perception, this section intends to expatiate on its psychological identity.

Generally, listening is a process of sensing, interpreting, and evaluating aural stimuli (Steil, Barker & Watson, 1983). Clark and Clark (1977) described the process of comprehension simply:

- 1) (Listeners) take in the raw speech and a certain representation of it in "working memory".
- 2) They immediately attempt to organize the phonological representation into constituents, identifying their content and function.
- 3) As they identify each constitute, they use it to construct underlying propositions, building continually onto a hierarchical representation of propositions.
- 4) Once they have identified the propositions for a constitute, they retain them in working memory and at some point purge memory of the phonological representation. In doing so, they forget the exact wording and retain the meaning. (cited in Zhu Ping, 1998, p.6)

Compared with Clark and Clark's interpretation of listening comprehension, Lingren (1997) focused on listeners' mental process of translating sounds into visual images. His presentation involved (1) sound stimuli coming into the listener's mind; (2) the creation of mental images of the sound stimuli, that is, visual presentation of the sound in the mind; (3) the association of the sound stimuli with the visual representation of the sound in the mind. It should be emphasized that listeners do not receive the coming sounds passively. The addition or elision of a sound that was not actually spoken shows that they are engaging in an active and creative process of comprehension. Cases of adding and eliding sounds or words in the study of this thesis are good evidence to support Lingren's speculation of the mental process in listening comprehension.

The above review introduces how a sound is perceived. The activation/competition model is often used to account for the psychological process of spoken word recognition. When a listener hears the coming word, "the arrival of incoming speech information calls up an array of potential word candidates which form at least temporarily a partial match to the speech input" (Cutler & Broersma, 2005, p.65). Once that speech information is stimulated, words

sharing some common features will pop up. As Maddieson claimed they are constructed using only a handful of phonetic categories (on average around 30), although vocabularies contain tens of thousands of words (cited in Cutler & Broersma, 2005). Cutler & Broersma further gave an example to show how the given word is finally selected and recognized by the listener. As is pointed out above that words with certain common features may be activated at the same time and shorter words will be embedded with longer words. When listeners begin to determine the intended word, they often feel confused. For example, the intended word star may be heard as start or stark or starve or startling. The input star activates all the words with formal similarity. Then all the candidates come to a process of competing. The incoming speech information plays a vital role in settling this competition. Ellis' spreading activation networks is similar to the activation/competition model. When listeners hear a word, at first they have no cues about this word in their mind. Ellis (1995) declared that the more information listeners can associate with the missing term, the more knowledge they could activate to determine the given word and that is how the network of association spreads. At last the target word will be searched. However, this is only one aspect of spoken word recognition. Some other researchers hold that listeners do not process speech sounds linearly. Liverman (cited in Miller & Eimas, 1994) reported that a single segment of the acoustic signal does not contain limited information for its own, instead, it gives useful hints for more than one phonetic segment, and conversely, the information for a given phonetic segment is often distributed across more than one acoustic segment. In this way, speech perception is largely context-dependent. If the incoming sound is /k/, the word stark is preferred more than all the other candidates, but listeners are prone to mishear or misinterpret phonetic symbols in continuous speech. Like a circle, the later speech information is responsible for revising the previous decision. Actually the above two views about speech perception reflect the discrepancy between the two chief models of speech recognition--- the Cohort Model of word recognition and the TRACE Model. Experimental data from the Cohort Model suggests that context only plays a minor and limited role in recognition and speech signal itself carries enough information to decide one lexical entity, while TRACE lays much stress on claiming that word processing is directly influenced by the top-down process.

Studies on speech perception

Many phonologists have conducted parallel researches on speech perception by both adults and infants. Historically, researches on speech perception only concentrated on how adults identified and discriminated phonetic information in the acoustic input (Goodman, Lee & Degroot, 1994). The belief was that infants processed speech in a way that was similar to adults. The research by Goodman, Lee and Degroot (1994) suggested how the infants' impressive innate perceptual abilities were modified by experience. Miller and Eimas described "the nature of mapping between acoustic and phonetic structure in adults and the origins for this mapping in infancy, coupled with the relative lack of process in discovering the nature of perceptual mechanisms that underlie speech perception" (Miller & Eimas, 1994, p.38). Pater and others (2004) did an experiment on young infants' perceptual acquisition of phonological contrasts. The result was that when infants first acquire a phonological contrast like place or voicing, they may lose this ability under the processing demands of word learning. In his article Cross-Language Speech Perception, Werker (1994) reviewed the history of cross-language speech perception research, and three periods have been identified. Most achievements in the first period were from the early seminal work. The comparison between adults and infants in their speech perception ability showed that young infants were better at discriminating nonnative contrasts. In fact, it is wrong to state that "age-related differences in cross-language speech perception result from an absolute loss of perceptual discriminability due to lack of listening experience" (Werker, 1994, p.93). The second and the third period began to doubt the argument in the first period and gradually realized that adults can be trained to discriminate the nonnative contrasts. His review partially showed the reasonability of strategy training in teaching speech perception. These findings are significant in listening teaching for they inspire researchers to think about the case of adults learning a second or foreign language: whether it is a universal problem for adult learners to lose the ability of discriminating certain speech sounds or it is just an individually-related issue. About word-level processing, Gilliam Brown and others (1994) reviewed some related influential connectionist models. One of them was the TRACE Model of spoken words identification by McClelland and Elamn. The importance of the TRACE Model lies in its predictive findings which include: categorical perception of phonemes; trade-offs between different features in phoneme identification; lexical effects on phoneme identification; preference for phonologically regular patterns; co-articulation effects; preference for short words (e.g., 'cat' vs 'caterpilla'); the strong, though not decisive influence of word beginnings; use of lexical information to identify word boundaries; the need, in some cases, for use of right context to identify words; the ability to cope with elision at word boundaries (Brown, 1994, p.108).

Specific researches on learners' speech perception have also been conducted. A study by Riney and others (2005) on how native-speakers and Japanese listeners perceived degree of accent in English showed that untrained Japanese listeners used non-segmental parameters like intonation, fluency, and speech rate to make perceptual judgments, while untrained American listeners just did the opposite: they depended heavily on segmentals (especially /r/and/l/) to perceive the difference. About the factors influencing learners' speech perception, Rubin (1994) identified five factors affecting listening: text characteristics, interlocutor characteristics, task characteristics, listener

characteristics and process characteristics. Hasan (2000) pointed out that it was not speed but other factors like pronunciation, hesitation, pauses and varied accents that hindered listeners' speech perception. Cauldwell (2002) wrote that traditional phonology failed to describe the characteristics of the mess stream of everyday speech. The messy products in connected speech affected listeners' understanding and interpretation. He also gave some examples of the messy products which easily cause listening difficulties, for example, /tu: bi: luki is often read as /luki.

SPEECH PERCEPTION ANALYSIS: METHODOLOGICAL CONSIDERATIONS

The validity of Contrastive Analysis: weak version

In the mid-twentieth century, applied linguists began to pursue the comparative study of two languages. Eventually the stockpile of the comparative and contrastive data on a multitude of pairs of languages yielded what commonly came to be known as the Contrastive Analysis Hypothesis (CAH). The advocates claimed that the principal barrier to second language acquisition was the interference of the first language system with the second language system, and contrasts between the two languages would enable linguists to predict the difficulties a learner would encounter (Brown, 2002).

CAH had its appeal in its early years and in some aspects it did work very well. English learners, after years' of studying English, still speak English with some foreign accents. In this way, supporters of CAH hold that the negative transfer of the first language is the most important factor preventing foreign language learners from acquiring a native-speaker pronunciation and thus CAH has the function of predicting the incoming difficulties in foreign language learning. Actually, it is said that CAH only plays a certain role in phonological aspect. Ellis (1994) claimed that it had been recognized widely that transfer was more prominent at the level of sound system. Some experts stated that negative transfer of the native language accounted for students' errors made in speech sounds production or discrimination, that is, the phonological difficulties were mainly caused by differences between the two phonological systems. For example, in accounting for the phoneme problem of Japanese speakers, /r/ and /l/ is commonly believed to be caused by the lack of /r/ and /l/ contrast in the Japanese sound system. And several studies have already shown that Japanese speakers have difficulty in perceiving this distinction (Sheldon & Strange, 1982; Flege, NaoYuki & Mann, 1995). Best (1988) in her studies also stated that the relationship between the native and the nonnative phonologies helped to predict the difficulty in discriminating a nonnative contrast (cited in Goodman, 1994).

Originally linguists who were in favor of CAH held that this theory can predict all difficulties caused by first language transfer in foreign language teaching and learning, and Robert Lado (1957) was one among them. He cited Fries' *Teaching and Learning English as a Foreign Language* (1945) to support his assertion: "the most efficient materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner" (cited in Wardhaugh, 1970, p.7).

This standpoint received criticism from some opponents, one of whom was Ronald Wardhaugh. He called the above claim as the strong version of the CAH and dubbed it as "unrealistic and impracticable". With the example of contrastive analysis of an English p and a French p, he pointed out that linguists had made many contrastive statements from the aspect of phonological problems ideally, without considering whether the phonemic systems of two languages were possible to contrast or not. Then he moderated the observational use of contrastive analysis as the weak version of the CAH, about which he claimed that:

The weak version leads to an approach which makes fewer demands of contrastive theory than does the strong version. It starts with the evidence provided by linguistic interference and uses such evidence to explain the similarities and differences between systems. (Ronald Wardhaugh, 1970, p.10)

While the strong view of CAH states that predictions are made based on a comparison between L1 and L2, the weak view starts with the learners' errors and attempts to account for them by comparing L1 and L2, so the starting point of the strong view and the weak view is quite different. That is why Wardhaugh consistently mentioned evidence of interference and stated that any analysis of phonological problems must "ultimately rest on phonetic evidence", in other words, the analysis must be bottom-up and data-driven. Theoretical assumptions or conclusions can not be made until details of some problem analyses come out first. This thesis is not based on an ideal analysis of the two languages' phonological systems with a conclusion stating that students will meet phonological problems in the process of listening because of a different language system from English. Instead, true experimental data must be elicited as good evidence for further analysis in that they are students' genuine listening productions which could reasonably serve as the first-hand materials to probe into their listening difficulties and problems. This thesis will adopt the propositions of the weak version of CAH, as Wardhaugh described it as helpful and undoubtedly would continue to be so as linguistic theory develops. Many previous researches used CAH as a model to predict learners'

speech production errors, since it is still useful in predicting a substantial portion of L2 learners' phonological errors. Contrarily, this thesis will use CAH as a framework to account for learners' listening errors after making an analysis of their phonological performance.

Identification procedures in Error Analysis

In dealing with errors, most teachers suggest correcting them immediately. Some others might choose to ignore them and only a few would find ways to help learners based on an analysis of their errors. Corder (1967) vividly described the picture of how learners' errors were ignored. He stated that "it almost seems as if they are dismissed as a matter of no particular importance, as possible annoying, distracting, but inevitable by-products of the process of learning a language about which the teacher should make as little fuss as possible" (Corder, 1967, p.163). However, learner's errors are significant because they reveal how far he has progressed and provide useful evidence of the processes by which a language is learnt. According to Corder, they are significant in three different ways:

First to the teacher, in that they tell him, if he undertakes a systematic analysis, how far towards the goal the learner has progressed and, consequently what remains for him to learn. Second, they provide to the researcher evidence of how language is learned or acquired, what strategies or procedures the learner is employing in his discovery of the language. Thirdly (and in a sense this is their most important aspect) they are indispensable to the learner himself, because we can regard the making of errors as a device the learner uses in order to learn. (Corder, 1967, pp.168-169)

Early critics of contrastive analysis hypothesis Banathy and Madarasz (1969) held that error analysis should form a complement to contrastive analysis. They declared that if one intended to bring findings of contrastive analysis into language teaching, error analysis must play a role in this process: "contrastive linguistics-no matter how refined-can only point toward a potential learning problem or difficulty. On the other hand, error analysis can tell us the intensity of the difficulty or the size of the problem." (Banathy & Madarasz, 1969, p.78)

A study by Libuse Duskova (1969) also manifested that contrastive analysis might be particularly useful in preparing teaching materials if profitably supplemented by the results of error-based analyses (cited in Robinett, 1970), because error analysis may be carried out directly for pedagogic purposes (Ellis, 1994). One example was given by Ellis was about the listening errors that a learner misunderstood the sentence 'Pass me the paper' as 'Pass me the pepper', because he was unable to distinguish the sounds /ei/ and /e/. Based on this phonological evidence, a researcher or teacher can use contrastive analysis to find out the causes of learners' inability in this aspect. In turn, results of these analyses can be applied into further listening teaching and may have a particular function of arousing learners' awareness as well. Listening is a complex process in which listeners have to decode the incoming speech and use their phonological knowledge to deal with many variances due to assimilation, elision or blending in connected speech and so on. According to Rost (2002), EA has been used to focus directly on phonological coding and unveil word-recognition difficulties.

As was mentioned above, needs analysis plays a crucial role in determining a successful application of a course, as Corder stated that "the suggestion that we should take more account of the learner's needs in planning our syllabuses is not new, but has not apparently led to any investigations, perhaps because of the methodological difficulties of determining what the learner's needs might actually be" (Corder, 1967, p.167). EA embodies a shift from research on teaching towards a study of learning and offers a way on how to analyze learners' errors and meet their needs. Brown (1994) and Ellis (1995) gave a detailed account of and exemplified a model for error analysis offered by Corder (1974). Corder outlined five steps in EA research: collection of a sample of learner English; identification of errors; description of errors; explanation of errors; evaluation of errors. Concerning what kind of errors should be analyzed, Duskova offered the following principle: "we are of the opinion that an error analysis should be based primarily on recurrent, systemic errors that are made by a number of learners and that can be readily traced to their sources, no matter whether they reflect defects in knowledge or whether they result from inadequate habit formation" (Duskova, 1969, p.219).

The purpose of error analysis is to explore the systematic error patterns. For methodological consideration, this thesis will conform to the first four steps of Corder's model of EA research, since the fifth one is usually considered as a separate and complicated issue by many researchers. EA in this paper will be made in the following procedure: first each word in the subjects' products will be compared with that in the standard reference text and anything which varies from the original will be counted and listed. Then all the errors including phonemes missing, replaced, added and so on will be categorized and put together according to their different characteristics. The last part is an analysis of the causes of the above errors based on contrastive analysis.

EXPERIMENTAL SECTION

Research questions

There are three main research questions in this paper. English majors' speech perception problem is the center of this paper. Causes and ways to solve these problems will be provided as well. Specific research questions to be

addressed here are:

- 1) What are English majors' error patterns in speech perception?
- 2) Why do they manifest these systematic errors?
- 3) What is the implication of this research in listening learning and teaching?

Subjects used in this research

Participants in this research are 30 Second-Year English majors in the Faculty of International Languages in Henan Normal University. They are from a natural class, and this class is chosen by chance. At first, all the 30 students participated in the test, but only 24 students' test results were valid, because 6 of them failed to complete all the five dictation passages leaving one or two totally blank. Among the 24 subjects, 4 subjects are from the city of Xinyang, a dialectal area which is between the border of Henan and Hubei Province. 10 subjects are from the city of Xinxiang where the University is situated. There are altogether 8 subjects from the city of Zhoukou, the city of Zhengzhou, the city of Zhumadian, and the city of Nanyang which are all in Henan Province. At last, there are two subjects from the province of Shanxi and the province of Anhui respectively. This study tries to give an account of these subjects' errors in speech perception.

Data-collecting procedure

In order to find out their speech perception problems, subjects are required to do five dictation tests in this empirical research. The dictation materials are selected from the instructional book TEM-4 Express Train compiled by Shen Fuying and Lv Xin in 2002. Five passages of different topics are selected covering the field of geography, politics, and transportation and so on. These materials are recorded by an American teacher who teaches Oral English in Xinxiang Teachers' College. The recording process conforms to the requirements in TEM-4. The dictation test is accomplished in twice. At the first time subjects are required to listen to the first three passages and the second time the last two ones. The overall procedure in the test conforms to the requirements in TEM-4. Altogether each passage is read to the subjects four times. During the first reading, which is read at normal speed, subjects are supposed to listen and get a general idea of the test. For the second and third readings, the passage is read sentence by sentence with an interval of 15 seconds for a pause. The last reading is back to normal speed again and during this time subjects are expected to refine their written work. Finally, two minutes is given for them to check it once more. After the subjects complete all the five passages, their written materials are collected. Altogether 120 pieces of materials (24 subjects x 5 passages) are used as data-base for further analysis. For each sentence in a dictation, there are 24 versions. In the process of analysis, all versions of the five passages are typed into the Microsoft Word file and altogether there are 120 pieces of materials. For each sentence, there must be 24 versions which are used to compare with the standard reference sentence to find faults. All the errors are then collected and classified based on their distinctive features. Each student is marked according to the types and the number of errors they made. Their pronunciation proficiency is tested as well. SPSS (Statistical Package for the Social Sciences) helps to find out the relativity between pronunciation and listening. After the test, a questionnaire is handed out to investigate subjects' overall impression on the five passages, which helps to identify the correlation between topic familiarity and the difficulty level of a text.

RESULTS

The research materials are five short passages selected from an instructional book of TEM-4. Procedures of how to collect and process the data have already been discussed in detail. In the final analysis, punctuation errors are not taken into consideration, for the sake of a main focus on subjects' linguistic competence. Altogether errors revealed by these subjects fall into six main categories: misperceived sounds, misperception of morphological changes, adding and missing sounds, misperception of unstressed elements, paraphrase and assimilation errors. The error of misperceived sounds and morphological errors are further sub-categorized into several types and each of them will be talked elaborately below. Distribution of the six categories is presented by the graphic representation. Frequency of each error type is measured in the ratio of the number of each error type to the total number of the errors. The total number of errors of the six categories is 193. The number of errors in each category will be mentioned when detailed analysis is under discussion. The distributional figures reveal that misperceived sounds, omission and addition of sounds and morphological errors occupy 84% of all the errors. These typical error patterns are meaningful to language teachers in pedagogical teaching. The following figure is the distribution of the six error types.

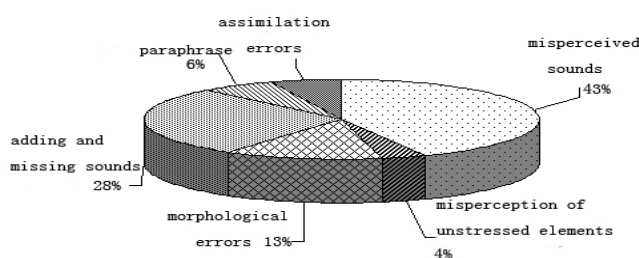


Figure 5.1 Frequency of each error type

Analysis of misperceived sounds

Misperceived sound is the biggest error type in speech perception tests. Tables are used to present errors appeared in each dictation material. Sources of the errors, subjects and frequency (frequency is counted in the ratio of the number of the subjects who made the specific error to the total number of the subjects) are also marked clearly in each table. The presentation of the raw test results is a preparation work for later analysis (see appendices). For example, in transferring speech sounds into letter symbols, there are 20 subjects who did not write the word crust correctly. 16 of them spelled crust as crast and 4 wrote it as cross. In another case, 4 subjects misspelled divided as devided. Although both errors involve formal misspelling, specifically they can not be classified into the same error type. Mishearing is attributed to the misspellings of crust, whereas the error of divided is caused by subjects' unfamiliarity of this word. The principle is that if the misspelled word can be pronounced the same as the correct one, then this error can be regarded as an error of sound substitution but not mishearing. The principle and validity of classifying each error type will be discussed in the following sections. Misperceived sounds can be further sub-categorized into 5 types according to the causes of these errors. They are: mishearing, sound substitution, misperception of multi-words sound sequence, phone insertion and sound omission. Figure 5.2 helps to show frequency of these five error types:

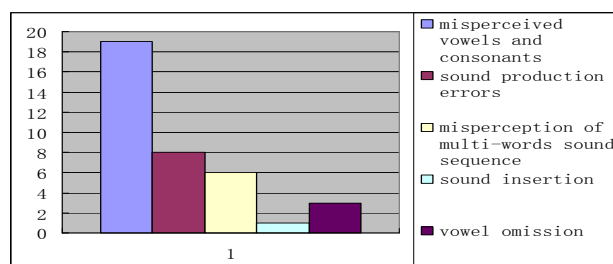


Figure 5.2 Distributions of the subcategories of misperceived sounds

Among the five types, the number of mishearing ranks first. L2 listeners have more difficulties in the part of sound discrimination, because any similarities between two phonemes are possible to cause their misunderstanding. Besides, their articulating habits can also have a huge and negative impact on their phoneme-determining judgments. This interfering factor will be elucidated in a later discussion. The error of sound substitution comes in a second place. In listening comprehension, if listeners fail to retrieve the target sound(s) by referring phonological cues, most probably they will substitute the target sound(s) with a phonological approximant or a phonologically-related non-word. The perception of multi-words sound sequence also seems troublesome for these subjects. Comparatively, subjects make fewer mistakes of sound omission and phone insertion. However, in the process of transcribing what they heard by handwriting, they would have corrected some errors on their own if they had have used sentential or semantic information.

Analysis of sound substitution

Another kind of error relevant to misperceived sounds is sound substitution when dictation-takers have to decide a word that they don't know exactly. The rule is that if they do not remember every letter which a word is composed of, they prefer to spell it according to its pronunciation or syllable, resulting that the target letter may be substituted by a phonologically-identical or phonologically-approximate letter, since many letters have the same pronunciation in English. Sound substitution can be accounted for by overgeneralization of English spelling rules by non-native English speakers. The orthographic form of a word may be in disparity with the sequence of segments which is heard. There are 4 subjects who wrote boundaries as boundaries, because both e and a can be pronounced as the

central vowel. Another 4 subjects wrote architects as architects and 2 more wrote engineers as engineers. Sometimes letter e and i have the same pronunciation as the front vowel [i]. Letter y has two pronunciations: semi-vowel [j] and the closing diphthong [ai]. The second pronunciation of letter y is the same as i in motorcycling. In the words motorcycling and nevertheless, both or and e are pronounced as the central vowel. The letter er can also be pronounced as the central vowel. Unfamiliarity of the exact spelling of the two words causes the substitution of letters er for e and or based on their same pronunciation. Both letter s and c have the pronunciation as the fricative [s]. This is an error which appears in 10 subjects' dictation handwriting. The addition of letter c before k in earthquake shows that 4 subjects were not clear about the exact spelling of this word or maybe ck and k also have the same pronunciation as plosive [k]. Another 4 subjects wrote engineer as engineer. Besides plosive [g], letter g can also be pronounced as, the same as letter j. The following table shows the frequency of sound substitution that occurs in all the five dictation tests.

Table 5. 1 Frequency of sound substitution

phoneme (s)	Letter(s) substituted	Frequency
[i]	i	16.67%
[ai]	y	16.67%
[E]	e	25%
	or	16.67%
[s]	c	41.67%
[k]	ck	16.67%
[δç]	g	16.67%

Sound substitution is very common in dictation tests or learners' written work. There is no simple sound-symbol correspondence system in English. This characteristic is elucidated clearly in the book *Teaching Pronunciation* written by M.Celce-Murcia and others. The following statement is quoted from the book and helps to explain the above errors more elaborately:

“In the case of English, the use of a phonemic transcription system is especially important because the language has no simple sound-symbol correspondence system — that is, one letter of the alphabet does not represent the same sound all of the time, nor does a specific sound always find its representation in one letter of the alphabet. For example, the letter c has four different pronunciations in the words cat, city, ocean, and cello; and the phoneme /s/ has various spelling representations, such as s, ss, c, ce, sc, and ps (as in sit, less, city, face, descent, and psychology). In some cases, the same letter within one word can represent two different sounds, as in the pronunciation of the two c's in success: first c is pronounced /k/ and the second one /s/. Also, combinations of letters are often simply an artifact of the spelling system and do not necessarily indicate that there is more than one actual sound being produced. This is the case with the underlined consonant sounds in words like stopped, butter, and reckon; it is also the case with the vowel sounds in feast, bait, and road” (M.Celce-Murcia et al. 1996, p.38).

Insertion of phoneme [t]

Of all the five dictation tests, there is one case of adding a phone in the syllable of a word. 10 subjects misheard layer as latter. This error is possibly to be justified by language transfer. Most frequently, a Chinese syllable structure is CV with no clustering of vowels or consonants, while English has a potential syllable structure of CCCVCCCC. It is found that Chinese English speakers prefer to add a vowel after a consonant in speech production to conform strictly to Chinese syllable structure CV. This tendency can also be transferred in speech perception. Listeners try to add a phoneme---vowel or consonant---to fulfill the established rule in Chinese. In dictation one, 10 subjects wrote layer as latter. The comparison of the two words' syllables suggests language transfer in the syllabic level. This error can be interpreted in another way. In fact, word frequency is one factor that affects lexical access. The later interview with these subjects reveals that they are neither familiar with the geographical background knowledge in that passage or the word layer, because it is not a commonly-used word and belongs to the peripheral vocabulary category. In listening test, words of this kind are hard to be activated and retrieved in limited time. The new or peripheral words can be falsely recalled or easily mixed up with the familiar ones. The insertion of alveolar plosive [t] may be caused by the subjects' conscious compliance with Chinese phonological rules or their failure to recall the correct form of word layer.

Table 5. 4 Insertion of phone [t]

Word1	Syllable1	Added phone	Word2	Syllable2	frequency
layer	/ˈleɪE(p)/	[t]	latter	/lAτE(p)/	41.67%

Omission of central vowel [E] and semi-vowel [j]

Respectively there are 50%, 41.67% and 25% of all the subjects who committed errors in spelling the three words

nowadays, specialists and lawyer. In word specialist, deleted after the letter i and e before the letter y in lawyer. Both and y are vowels (y is usually regarded as a semi-vowel). The syllable of nowadays is /naʊ. ɪ. ɔ/. Although letters before the omitted letter a are w and d, the syllable of this word shows that the diphthong is before the central vowel. The conclusion here is that omission of sounds appears in the word structure CVVCV or CVCVV where two vowels can occur continuously, which is referred as the avoidance of hiatus (Naucler, 1980). There is no evidence of consonant omission in subjects' dictation tests. Those points below are some other ways to interpret the error of omission. The following table presents the frequency of this error.

Table 5. 5 Omission of letter a and letter y

phoneme omitted	Word	Error	Frequency
[E]	nowadays	nowdays	50%
	specialists	specilists	41.67%
[j]	lawyer	lawer	25%

1) Omission of letter and letter y in nowadays and lawyer is not difficult to understand. Nowadays is composed of now + a + days. It seems that nowadays is a compound word. In speech perception, now and days could be recalled immediately as two independent parts. For a commonly-used word like nowadays, listeners are supposed to know its pronunciation and it is impossible to commit this error because of mishearing. Perhaps carelessness is the most important factor 12 subjects should be blamed for their neglect of letter a.

2) In this case, the corresponding letter y as a semi-vowel [j] is omitted. The root of lawyer is law. Suffix -er can change words into nouns by adding -er at the end of nouns denoting people or things and adjectives or verbs. Examples are cobble + (e)r, foreign + er, village + (e)r and so on. However, when law changes its form to be a noun indicating people, letter y should be added between law and -er. Apparently 6 subjects failed to detect the weak articulation of semi-vowel [j] or they may just forget the irregularity of spelling rules and omitted letter y.

3) There are 10 subjects who omitted letter a when writing specialists. The syllable for specialist is /'spe list/. Letter c is pronounced as fricative, and letters i and cooperate to pronounce as the central vowel. In English, the central vowel can be realized both by a single letter and clusters. In dictation tests, if listeners fail to hear every phoneme clearly, they will end up with omitting letters, because they have little time to check and correct the spelling of words.

Misperception of weak sounds and unstressed phonemes

In the dictation tests, subjects revealed their difficulty in perceiving the word and in reduced speech. 16 subjects wrote and as an. At the same time, 8 subjects wrote an as and. In English, some words have two forms—a strong form and a weak form—depending on their positions in a sentence. Weakening is “phonetically motivated process of sound change that leads to the reduction of sounds and, in extreme cases, to loss of segments; typically this occurs in positions where assimilation is favored or in syllabically ‘weak’ positions (e.g. in final position, in unstressed syllables)” (Bussmann, 2000, p.519). This special group includes form words and pronouns. Conjunctions and, as, but, or and some others often have weak forms in reduced speech. For example, the strong form of and is read as /And/, while its weak form is read as /En/, /En/ or simply nasal /n/. The American speaker in this dictation test reads and as /En/ without articulating plosive [d] loudly. If sentential information is not taken into consideration, listeners are likely to write an instead of and according to its pronunciation.

Some subjects showed their inability to distinguish unstressed phonemes in listening. In English, it is quite common to find that words with two or more syllables have a strong stress on one of their syllables and a weak stress on the remaining syllable or syllables. Just as its name implies, the syllable or syllables which are stressed will be read more strongly than the other syllable or syllables which are not emphasized and read weakly. The feebleness in articulating the unstressed syllable or syllables causes difficulty in hearing. In this dictation test, subjects wrote words adult, above, account, against and extends wrongly. For the first four words, stress usually falls on the syllable following the prefix a-. In speech recognition, subjects missed the unstressed central vowel [E]. About the word extends, the unstressed phoneme combination [ik] was missed. One thing in common is that all the five carrier words have weak-strong stress patterns (i.e., the first syllable is weak). The following table shows the frequency of these errors.

Table 5. 6 Frequency of misperceived unstressed phoneme(s)

Phoneme(s) missed	Word	Syllable	subjects	Frequency
[E]	Adult	/E'dθλτ/	14	58.33%
	Above	/E'bθm/	12	50%
	Account	/E'kαυvτ/	8	33.33%
	Against	/E'genst/	4	16.67%
[ik]	Extends	/ik'stend/	12	50%

Misperception of fricative and inflectional changes

English word forms such as walks, walker, walked and walking contain the same word root walk and four other elements –s, -er, -ed, -ing. These elements are called morphemes. The definition of morpheme is “the smallest meaningful element of language that, as a basic phonological and semantic element cannot be reduced into smaller elements” (Bussmann, 2000, p.313). Morphemes can be divided into two types: free morpheme and bound morpheme. Free morphemes (sometimes also referred to as unbounded morphemes) are words which can stand alone such as walk and wake. Those which can only occur as a part of words such as –s, -er, –ed and are typically attached to a root are called bound morphemes. Most subjects committed the error of missing fricative [s] at the final position of plural words. It may be generalized that Chinese learners have certain tendency to ignore this grammatical restriction in English countable nouns. Phoneme(s) before suffix –s may influence subjects' perception to some extent. For the sake of a clear observation, the following table lists those plural words and phoneme(s) before the suffix –s.

Table 5. 7 Frequency of missing fricative

Phoneme(s)before[s]	Word1	Word2	Frequency
[E(:)]	layers	layer	8.33%
	programmers	programmer	33.33%
[t]	plates	plate	25%
	motorcyclists	motorcyclist	25%
	insects	insect	16.67%
	rights	right	25%
	specialists	specialist	41.67%
[k]	earthquakes	earthquake	33.33%
[i(:)]	boundaries	boundary	41.67%
[m]	customs	custom	33.33%
[b]	jobs	job	16.67%
[Φ(E)v]	positions	position	50%
[f]	lives*	life	41.67%

(Note: The erroneous writing of life for lives can be accounted for by two reasons. The first one is that subjects have difficulty in distinguishing fricative [f] and [v]. The second one is that subjects are poor at using sentential information to correct this error.)

In Chinese, nouns do not go through morphological changes to show their plural meaning. English is different from Chinese in this aspect. Besides those nouns whose singular and plural forms are the same, all the other nouns should be changed to their corresponding plural forms if needed. Differences of two languages in the same facet bring confusion to the non-native learners. Furthermore, morphology can affect lexical access (Jay, 2004). There is a widely accepted hypothesis about how words are stored in the mental lexicon. Root words and morphological affixes are separately stored in memory, and morphological affixes can be used to append to the root word when necessary. In this way, cognitive economy is achieved, for only root words require a storage space but not those morphemes-attached elements. In speech perception, listeners have to first access the root words to recognize the morphemes-attached elements, for example, first custom and then customs. Foreign learners have to go through a longer time course to decide a word. In other words, omission of fricative [s] can be seen as a result of learners' effort to reduce their linguistic burden in memorization. From these points of view, it is not difficult to understand why Chinese learners make this kind of mistakes both in dictation and composition. In the dictation tests, Chinese listeners frequently missed fricative [s] before the central vowel, the front vowel, fricative [t] and [k]. The collection of these phonemes shows that in most cases, fricative [s] is missing before consonants, that is to say, in terms of the phonemes before fricative [s], consonants outnumber vowels and there are only two cases of fricative [s] missing before a vowel. However, except plosive [t], no other predominance has been found of all the consonants before fricative [s]. Errors of this type might be described as mistakes in performance but not in competence, because subjects are able to correct these errors using clues of sentential information if given enough time.

The missing of fricative [s] in the above analysis belongs to the type of morphological ending errors. The inability of perceiving possessive pronouns is also a kind of this type. This section deals with inflectional changes of genitives, another case of misperceiving morphological endings.

Table 5. 8 Frequency of misperceived genitives

Word(s)	Error	Subjects	Frequency
your	you	18	75%
women's	women	4	16.67%
teacher's	teacher	6	25%
earth's	earth	16	66.67%
two-year	*two years'	12	50%

(*it is a counter example indicating that subjects have internalized this rule.)

As in the above case, this error is resulted from subjects' failure in performance rather than in competence. Most subjects wrote the possessive pronoun your as the personal pronoun you. The word your appears in the sentence "if that is the customs your friends are following the...". As a personal noun, you can not modify friends but can be used as its apposition. There are 4, 6 and 16 subjects who respectively wrote women's as women, teacher's as teacher, earth's as earth, and omitted the bound morpheme -'s which denotes the possessive feature of the nouns in question. Some subjects revealed their inability to make the right judgment to decide the right form of a verb. The last error is of the same kind. The context of two-year is "you have to get a special two-year driving license before you can start". Two nouns with a hyphen have the possessive feature of an adjective and can modify another noun following them. Another way is to add -s' at the final position of the second noun to realize this function. Subjects tried to make use of what they have learned to reorganize the sentence. Here it shows listeners' active role in listening process rather than passively receiving information. Errors of this type show the effect of schematic knowledge in listening comprehension. In a listening test, if listeners take sentential or textual information into consideration, sometimes these stereotypes could cause their confusion especially when there are several possible variations of one term.

Missed and added sounds

Missing and adding "sounds" (here "word" as a unit) occur in the dictation tests very frequently. Dictation is different from the traditional listening comprehension. In listening comprehension test, what listeners need to do is get the main idea or specific information from the listening materials and then answer some related questions. In dictation tests, listeners are required to memorize every word for the sake of transcribing in the paper at the same time. Short-term memory (STM) plays an important role in this process. There have been many experiments on what and how much can be stored in STM. Researches found that listeners do not remember a sentence or text by verbatim. They can remember more words relevantly that contribute to the complete meaning than isolated. Missing and adding words are partly caused by the instantaneous feature of STM.

CONCLUSION

Relatedness between listening and pronunciation

The notion that there is a high correlation between listening and pronunciation is not new. Brown (1974) demonstrated that "those native speakers who can most accurately evaluate the pronunciation proficiency of other nonnative speakers exhibit better pronunciation themselves than do those who rate others less accurately" (cited in Celce-Murcia et al. 1996, p.244). In speech perception, listeners have to detect sound signals carried by articulator gestures. In this way, Best claims that "perception is intrinsically linked to knowledge of how speech is produced" (cited in Goodman et al. 1994, p.13). With regard to American's problems in distinguishing Spanish, Stockwell and Bowen suggest that "listeners should learn to pronounce Spanish correctly before they want to hear it correctly" (cited in Robinett et al. 1972, p.29).

In this research, specific analysis is also made to examine the relationship between perception and production. First, subjects' dictation materials are marked according to the grading standards of dictation test in TEM-4 without considering punctuation errors, simply for the sake of focus on sound perception problems. Major errors like disagreement between subject and predicate are deducted by 0.5 points. Minor errors like the wrong use of articles are deducted by 0.25 points. Secondly, subjects' pronunciation proficiencies are marked by a native English speaker. Pearson product moment correlation is frequently used to find out the relationship of two parameters, like dictation and pronunciation in this case. The following statistic shows their relatedness.

Table 6. 1 Descriptive statistic

	Mean	Std. Deviation	N
pronunciation	78.4167	7.15005	24
dictation	78.7500	8.21584	24

Table 6.2 Correlation between pronunciation and dictation

		pronunciation	dictation
pronunciation	Pearson Correlation	1	.864(**)
	Sig. (2-tailed)	.	.000
	N	24	24
dictation	Pearson Correlation	.864(**)	1
	Sig. (2-tailed)	.000	.
	N	24	24

Table 6.2 shows that the coefficient of the correlation between pronunciation and dictation is 0.864. There are two * on the right upper side of the coefficient, which means that pronunciation and dictation are in remarkable correlation. The statistic in this research supports the prevailing standpoint that speech perception and pronunciation is closed related.

Implications to pedagogical teaching

Firstly, unlike many other pure theory-centered or intuition-inspired researches, this research is data-based, which suggests its objectivity and meaningfulness. The research method can be useful to teachers who want to find out their students' problems in speech perception. Secondly, in this research, speech perception errors are classified in different types and shown by tables and graphs, which can be partially used by teachers in pedagogical teaching. They can make these systematic errors known to their students. Thirdly, besides the problems of distinguishing phonemes, this research also probes into some other facets which can influence learners' speech perception. They are mother tongue transfer, background knowledge, short memory, and contextual information. These factors are specially pointed out in order to make learners be aware of their importance. Listening teachers are expected to find ways to help their students notice contextual information, build background knowledge, and increase short-memory span. Fourthly, although researchers have not yet reached a consensus over the effects of phonetic discrimination training, laboratory results have indicated its positive role in improving learners' listening ability to some extent. The study by Jamieson and Morosan (1986) suggested that "the introduction of stimulus variability will help listeners to respond accurately to the complexity of natural speech" (cited in Pisoni et al. 1994, p.144). Teachers could design appropriate training methods based on an analysis of their students' listening problems. Last but not least, pronunciation has a direct effect on speech perception. Its strong correlation with listening has been proved by Pearson coefficient in the above table 6.2. In listening classes, listening teachers can use all kinds of resources to promote the integration of listening with other language skills such as oral English teaching. In this way, the side effects of mispronunciation on listening can be avoided to some extent.

REFERENCES

- [1] Arondt, J. *Journal of Memory and Language*, **2006**, 54(1), 113-130.
- [2] Baayen, R. H., L.B.Feldman, & R.Schreuder. *Journal of Memory and Language*, **2006**, 54(1), 290-313.
- [3] Banathy, B. H., & Madarasz, P. H. *Journey of English as Second Language*, **1969**, 4, 72-92.
- [4] Zhang B.; Feng Y.. *International Journal of Applied Mathematics and Statistics*, **2013**, 40(10), 136-143.
- [5]Zhang Shu-xue. *Journal of Nanjing Institute of Physical Education*. **1995**, 31(2), 25-27.
- [6]Pan Li. *Journal of nanjing institute of physical education(natural science)*. **2004**, 19(1), 54-55.
- [7]Li Yu-he; Ling Wen-tao. *Journal of Guangzhou Physical Education Institute*. **1997**, 17(3), 27-31.
- [8] Xu Guo-qin. *Journal Of Hebei Institute Of Physical Education*. **2008**, 22(2), 70-72.
- [9] Chen Qing-hong. *China Sport Science and Technology*. **1990**, 21(10), 63-65
- [10] Tian Jun-ning. *Journal of Nanjing Institute of Physical Education*. **2000**, 14(4), 149-150.
- [11] Zhang B.; Zhang S.; Lu G.. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(9), 256-262.
- [12] Zhang B.; *International Journal of Applied Mathematics and Statistics*, **2013**, 44(14), 422-430.
- [13] Zhang B.; Yue H.. *International Journal of Applied Mathematics and Statistics*, **2013**, 40(10), 469-476.
- [14] Xiaomin Zhang. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(12), 8-14.
- [15] Wang Bo; Zhao Yulin. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(12), 21-26.
- [16] Mingming Guo. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(12), 64-69.
- [17] Bing Zhang. *Journal of Chemical and Pharmaceutical Research*, **2014**, 5(2), 649-659.