

# The Real Trouble Spots in the Perception of English

Midori IBA

## Abstract

*Second or foreign language (L2) auditory perception tends to be problematic because first language (L1) transfer leads to numerous auditory mistakes. These mistakes have complicated origins but learners seem to ascribe them above all to speech rate.*

*On the other hand, in the field of second language pedagogy, the trend since the 1990s has been to think of listening in terms of something that occurs in combination with reading, writing and speaking. Yet, the exercises of many listening materials are still essentially traditional in design. Task-based activities seem to be appropriate to increase learners' listening skills but actually they are not enough. Teachers should understand the underlying reasons for auditory mistakes made by learners. Error analysis, the procedure of describing and explaining errors, has played an important part in creating a sounder perspective of the learning process.*

*In this article, I have analyzed the results of a dictation test given to 56 students at Konan University in Japan. I will mainly discuss some of the trouble spots related to suprasegmental features but will also mention some other findings derived from the analysis. I will also argue that this type of analysis is helpful in teaching listening for better communication and suggest some activities effective to improve listening skills.*

## Introduction

It is true that the sound structure of the L1 manifests itself in the learning of the L2 in a more obvious way than the syntactic structure of the L1 does. Research has shown that L1 structure affects the developing L2 sound structure under certain conditions (Goldinger, 1997). Patterns of articulation and perception, sound sequencing, and stress and intonation inevitably carry over from the native to the foreign language. L2 listening is assumed to parallel L1 listening in terms of processing procedures and strategies to the extent that it is influenced by L1

processes of knowledge transfer.

Auditory mistakes may have several origins: stress pattern misassignments, pitch misplacements, or phonemic segment confusions. Yet, L2 listeners often report L2 speech as being “too fast.” This might reflect their affective discomfort in managing the uncertainty of phonological perception. In particular, Japanese learners of English are said to be unfamiliar with the speech simplification patterns that occur in English. Mishearings, however, are not precisely predictable. Sometimes the errors predicted by contrastive analysis do not occur at all, while many of the errors that were not predicted actually occur. In that case, it is natural to reverse the procedure by starting from authentic materials. Describing the errors first and then analyzing their possible causes will allow us to say something about the role of the L1 in the second language acquisition process.

The limitation of error analysis is that focusing on errors provides a one-sided approach to learner language. Description and analysis of errors may give a misrepresentative picture of L1 influence. The term ‘error’ may also have a negative connotation. It implies that there are some undesirable faults in the learner’s performance. Therefore, the purpose of error analysis should be clearly recognized. Error analysis has played an important part in creating a different and more reasonable perspective. Errors should be regarded as normal and inevitable features in L2 learning. The developmental pattern of an individual learner should be taken as evidence not of failure but of achievement.

In this paper, the procedure and error analysis of a dictation test will be described in the first and second sections, and then suggested activities will be given as concluding remarks in the third section.

## **1. Participants and dictation test procedure**

### **1-1. Participants**

A total of 56 native speakers of Japanese participated in this study. All were undergraduate student (15 male/41 female) at a medium-sized university, and all were enrolled in the first semester of their second or third year. They were taking an intermediate-level listening course, and two classes were combined for the test. Level placement was determined by testing. None had studied or lived abroad at that point. From teaching them during the semester, I concluded that the participants were representative of the intermediate college level (i.e., high beginner-low intermediate). All participants were motivated to improve their listening skills of English. All participants were offered the opportunity to obtain feedback on testing performance when data analysis was completed.

## 1.2. Procedure

- (1) Students watched a video clip of American tourists interviewed in San Francisco. (Material: *Viva! San Francisco*, Unit 3 “Face the Camera”, Macmillan Language House. See appendix for transcript.)
- (2) They listened to their own MD version of the video clip. Recording was made automatically while they were watching the video.
- (3) They were asked to transcribe every word they heard on a sheet of paper. They were allowed to listen as many times as they liked.
- (4) Transcriptions were collected and analyzed by the writer.

## 2. Results

The total number of mistakes was 1,191. This number includes spelling mistakes, and although, precisely speaking, these should not be categorized as a type of auditory mistake, it is often hard to judge whether a mistake is caused by a learner's consonant confusion (e.g. /r/ and /l/) or inaccurate memory.

The duration of the video clip is one minute and 10 seconds, and the number of words in the script is 214. Considering the duration and the number of words, the total number of mistakes can be said to be reasonable and was slightly less than my prediction. The mistakes that students made have a clear tendency and many of them are found in the same parts of the material. The following are the most common perception errors that students made.

### **Most common perception errors:**

#### **No.1 “What are you going to do today?”**

- (1) No correct sentences out of 56 students.
- (2) 11 students didn't write anything.
- (3) 25 students could recognize the word “today”.

### **Possible rationale for the errors: Lack of knowledge of speech simplification patterns**

- (1) In this case, the phrase ‘what are you going to do’ is phonologically changed to sound like /wəryəgənədu/. The consonant /t/ in ‘what’ is elided, the vowel /a/ in ‘are’ is reduced, ‘are’ and ‘you’ are linked, ‘going’ and ‘to’ are combined and changed. Students were most surprised to learn how the speech sound of such basic words could change so drastically.
- (2) I asked the students who wrote nothing at all the reason for their responses. To them, this set of words sounded like one word which they had never heard of.
- (3) Other students perceived ‘what,’ ‘do,’ or ‘today.’ This is because English is

a stress-timed language. When spoken faster, a stress-timed language usually shortens, obscures, or drops vowels to carry more syllables between two stresses without changing its rhythm much. Since Japanese is mora-timed, Japanese learners are not familiar with the sound changes that occur in English when it is spoken.

### **No.2 “Tell me some of the things you’ve liked so far.”**

- (1) No perfect sentences out of 56 students.
- (2) Typical mistakes:
  - some of the things → something (no correct answer)
  - you’ve liked → you liked ( 8 students gave the correct answer)

### **Possible rationale for the errors: Unknown expression and lack of grammatical parsing**

- (1) The students didn’t know the expression ‘some of the things,’ although everyone knew the word ‘something.’ They substituted the unknown expression with a familiar one, ignoring the semantic difference. This phenomenon is a typical example of a word or an expression that is chosen from the acquired vocabulary in a learner’s mental lexicon. According to Rost (1999), influenced by ‘lexical effects,’ listeners tend to ‘hear’ words that they already know, even if the input contains unknown words. The expression ‘some of the things’ seems common enough, but actually it is not taught at school in Japan.
- (2) Most of the students knew the present perfect as in ‘you’ve liked’ but didn’t apply their knowledge when they heard the words. In actual pronunciation, the consonant /v/ in ‘you’ve’ is so weak that it is more difficult to perceive than it seems. In such a case, a grammatical explanation by an instructor is needed.

### **No.3 “...and then tour around the waterfront district a little bit...”**

- (1) No perfect transcription out of 56 responses.
- (2) 10 out of 56 perceived ‘waterfront.’
- (3) 15 out of 56 perceived ‘a little bit.’

### **Possible rationale for the errors:**

- (1) Students find it hard to perceive ‘and then’ because the words are glided over and hardly pronounced. These words are generally non-stressed and considered as function words. The phrasal verb ‘tour around’ is unfamiliar to students, although they knew ‘tour’, which is recognized as a loanword ‘ツアー’ in

Japanese. However, they can't perceive the original phrasal verb since the function of the loanword is as a noun and its pronunciation is quite different from the original. Moreover, adding the word 'around' to 'tour' makes students confused in perception.

- (2) The reason why 10 students perceived 'waterfront' is probably because they had already known it as a loanword. Yet, the following word 'district' was perceived by only six students. Two of them misspelled the word as 'destrict', but they apparently perceived the word phonologically. In counting the errors in this study, an apparent misspelling like this is not categorized as a perception error.
- (3) The expression 'a little bit' was not clearly pronounced in the interview but 10 out of 56 students perceived it. It can be said that a familiar expression is perceived comparatively easily even if it is in an unstressed-syllable.

#### **No.4 "...and then we are heading north"**

- (1) No perfect transcription out of 56.
- (2) Four out of 56 perceived 'north.'

#### **Possible rationale for the errors:**

- (1) It must have been hard for students to perceive 'and then we are heading' because this set of words is unstressed and pronounced quickly and without pauses. The verb 'head' and its present participle 'heading' were not familiar to students though they knew the noun form of the verb.
- (2) The word 'north' was perceived because it was stressed. Yet, only four students understood it, probably because the consonant / θ / in 'north' is in word-final position and is usually not clearly pronounced. It is therefore hard to perceive. In that case, the word becomes less intelligible and sounds like /nɔr/ or /nɔ/.

#### **No.5 "...and so many things to do that are very easy to walk to"**

- (1) No perfect transcription out of 56.
- (2) The noun phrase 'so many things to do' was perceived by six students.
- (3) The sequence of words 'that are very easy to walk' was not perceived by any students.
- (4) The last word 'to' was perceived by 11 students. Four other students wrote 'too' instead of 'to.'

**Possible rationale for the errors:**

- (1) This set of words ‘...and so many things to do that are very easy to walk to’ must have been very difficult for students to perceive because the number of words is 13 but the stressed words were only two (‘things’ and ‘to’) when pronounced.
- (2) Six students out of 56 are not many, and yet the number of the students who didn’t write ‘to do’ but wrote ‘so many things’ was 36. This suggests that the phrase ‘so many things’ is common to the majority.
- (3) The speaker of this utterance was a middle-aged man with a low-key voice. The set of words ‘that are very easy to walk’ was pronounced in a low pitch without any stress or pauses. It must have sounded unclear to students.
- (4) The word ‘to’ gained prominence because it was stressed. That could be the reason why it was perceived. Although four students confused ‘to’ with ‘too,’ they were able to perceive the sound /tu/. If there had been no prominence on ‘to,’ the number of the students who perceived the word would have been fewer.

**No.6 “You’ve had a lot of fun, lot of fun”**

- (1) 30 students wrote ‘you had’ instead of ‘you’ve had.’ Eight students wrote ‘you’ve had’ or ‘you have had.’
- (2) 45 students perceived ‘a lot of fun.’

**Possible rationale for the errors:**

- (1) The students who wrote the present perfect form are the same who wrote it in No.2 (‘tell me some of the things you’ve liked so far.’) That means they have already acquired the form grammatically and phonologically. As mentioned in No.2, some of the other students knew the present perfect as in ‘you’ve had’ but didn’t apply their knowledge when they heard the words. In actual pronunciation, the consonant /v/ in ‘you’ve’ is so weak that it is hard to recognize for most of the students.
- (2) The expression ‘a lot of fun’ is so common to students that most of them can recognize it even when it is not pronounced clearly.

**No.7 “We had sushi the first night actually.”**

- (1) 49 students wrote ‘we had sushi,’ and seven students wrote only ‘sushi.’
- (2) No perfect transcription of the sequence of words ‘the first night actually.’

**Possible rationale for the errors:**

- (1) The majority of students wrote the correct transcription although the speaker of this utterance was a middle aged woman and her pronunciation was not very clear. It suggests that students were familiar with the past form of the verb 'have' and its meaning. The word 'sushi' is also known to every student.
- (2) The reason why students didn't transcribe the sequence, 'the first night actually' is probably because there was no stress on it and no pauses between the words. To students, it sounded like an unfamiliar word. The combination of 'the first night' and 'actually' was not familiar with students as well.

**No.8 "We're from Washington, D.C."**

- (1) No perfect transcription of the sequence of words 'we' re.' Eight students wrote 'we come.'
- (2) 10 students wrote 'Washington, D.C.'

**Possible rationale for the errors:**

- (1) In the contracted form of 'we are,' the verb 'are' is no longer stressed and unclearly pronounced. As for the expression 'to come from,' it is quite common among students. Eight students might not have been confident in choosing the words 'we come from,' from their mental lexicon.
- (2) Proper nouns are usually hard to perceive but 10 students out of 56 transcribed 'Washington, D.C.' correctly because it is the name of the capital of the United States and they had already learned it in school.

**No.9 "Tell me where you folks are from?"**

- (1) No perfect transcription out of 56 students. No students perceived the word 'folks.'

**Possible rationale for the errors:**

- (1) In the phrase 'you folks,' 'folks' is in apposition to 'you.' Some of the students must have known the word but they didn't know the use. If the sentence were "Tell me where you're from?" or "Tell me where you guys are from?," the number of correct answers would probably increase.

**No.10 'Napa Valley,' 'Sonoma,' 'Alcatraz'**

- (1) No correct translation of 'Napa Valley' and 'Alcatraz.'
- (2) Five students wrote 'Sonoma.'

### **Possible rationale for the errors:**

- (1) Proper nouns are difficult to perceive, except for some common words (see No.8.)
- (2) Although Sonoma is a popular destination for wine lovers, students didn't know the place. The reason why five students wrote the correct answer is that the spelling of the word happens to be the same when written in the Roman alphabet, which is known to students.

## **3. Concluding remarks and suggestions**

### **3.1 General tendency of errors**

Judging from the errors from No.1 to No.8, mishearings take place when the target sentences or phrases are relatively longer. These especially emerge in the weak forms of English stress patterns. As mentioned earlier, the Japanese language is recognized as mora-timed, Japanese learners of English are not familiar with stress-timed languages. Monolingual Japanese listeners tend to process another language with their own speech segmentation pattern; mora-timed. To them, the rhythm of English is quite different, and they are sometimes uncomfortable with it. In the speech stream of English, there are some stable segments that are clearly articulated, typically stressed syllables. On the other hand, there are the less clearly articulated parts, weak forms. Especially in weak forms, several phonological phenomena could appear and they are problematic for Japanese listeners: consonant assimilation; vowel reduction; and elision.

Mishearings are caused by lexical effects and schematic effects of language processing known to operate in both L1 and L2 listening. With lexical effects, listeners tend to hear words that they already know even if the input contains unknown words. They are also influenced by schematic efforts. They tend to hear words that are more common or expected in a given context, even when less common or less expected words are actually spoken.

As for grammatical parsing, errors occur if words are assigned to the wrong grammatical class, if semantic relationships between lexical items are not properly computed, or insufficient syntactic information is given. As grammatical knowledge provides top-down information in word recognition, errors are often crucial and listeners will not understand the meaning of words at all.

In checking 56 sheets of students' transcription, several students stopped writing in the middle of the discourse. They seemed to give up transcribing. I asked them the reason why they stopped writing. They explained that they suddenly felt the discourse was 'too fast.' Actually, the speech rate of the discourse is not that fast (3.06 words/sec). This 'mental block' might reflect



affective discomfort in managing the uncertainty of phonological perception.

### **3.2 Suggestions for teaching: Make students notice the importance of weak forms**

Teachers of English should increase student awareness of the characteristics of stress-timed language and encourage them to focus on their errors in the weak forms when they practice. This may contradict with the often-used listening strategy “Try to catch the content words (the strong forms) when you listen.” Of course, this is true for beginners or in the case of listening comprehension tests, yet if the students are in the second stage of learning, they should also be aware of the ‘minor’ parts of the target language as well. For this purpose, language transcription is a good option. There are mainly three ways for Japanese learners of English regarding transcription: using the phonetic alphabet, using katakana and transcribing English directly. The following are the merits and demerits of each way. Teachers could apply one of them or a combination of two or three ways in their classes.

#### **3.2.1 Using the phonetic alphabet**

As for the trouble spot like No.1 “What are you going to do today?,” teachers can use the phonetic alphabet such as /wəryəgənədu/ for explaining the phonological change and practicing pronunciation. Phonetic transcription can make students realize that the pronunciation of a word can differ, sometimes radically, depending on the sort of environment in which the word is said. This analytical process will also help students to improve their listening skills. The types of variation in phonetic forms are not random or arbitrary. They are common to all language to economize on the amount of effort when speech sounds are produced. The utterances of speakers of all languages display this type of feature which can be explained by reference to a small number of connected speech processes, though there are differences in detail between one language and another. Once you master the international phonetic alphabet, you can apply it when you transcribe any word or expressions of not only English but also any language.

Especially if you teach a teacher training course, the phonetic alphabet should be taught to students so that they can listen to English analytically and explain the change of quality of speech sounds.

We could say that phonetic transcription is an ideal and useful learning technique for L2 learners, and yet mastering phonetic symbols takes time. For example, in RP, standard British English, there are 24 consonant sounds classified

according to the place and manner of articulation and voicing. There are nine places articulation: bilabial, labiodental, dental, alveolar, post-alveolar, palatal, velar, glottal and labial-velar. Regarding manners of articulation, there are five ways: plosive, fricative, affricate, approximant and nasal. It is impossible to learn to transcribe without getting involved with this technical terminology. For instance, the consonant /v/ can be described as a voiced labiodental fricative. As for vowels, there are 12 monophthongs and eight diphthongs. It would be time-consuming even if you focus on mastering English sounds only.

### 3.2.2 Using katakana

The transcription of English to Japanese has been done since the earliest cultural contacts between English speakers and Japanese. During the Edo period, kanji were used phonetically to write English and other foreign words, but in the modern period katakana has become the principal target script. Unlike the systems for *romaji*, there is no standard for transcribing into katakana, and methods vary. However, generally all methods attempt to preserve the pronunciation of English, not the spelling.

In a regular English class, katakana can be a substitute for the phonetic alphabet. Some students cannot realize the change of connected speech sound well. They try to listen to the words separately. In that case, showing katakana transcription would help students to become aware that they are quite different from the words they see in a script. ‘What are you going to do?’ in No. 1 would be ‘ワリヤゴナドウ’ and students can easily understand the phenomenon.

However, there are some difficulties in applying katakana for transcribing English, mainly because the Japanese language distinguishes fewer sounds than the English language. For example, Japanese does not distinguish the vowel sound of “run” and “ran”, or the consonant sound of “row” and “low”. Moreover, the rules by which sounds can be combined in Japanese are generally more restrictive than the English rules. As a result, the pronunciation of the transcribed word can differ quite considerably from the original word in English. Yet, using katakana for English transcription is easier for Japanese people than using the phonetic alphabet. It is not a perfect method, but it can be used as an effective supplementary means.

### 3.2.3 Transcribing English directly

In a strict linguistic sense, transcription is the process of matching the sounds of human speech to special written symbols using a set of exact rules so that these sounds can be reproduced later. In this article, the term is used to refer to the

conversion of a spoken language into written, typewritten or printed form.

By applying the transcription activity in class, students will be aware of English grammar. It will also help students learn the spelling of words. The spelling system of English is so complex and full of exceptions. Transcribing is a skill which needs regular practice. It is much better to do a little transcription fairly often rather than a lot all at once. As transcribing words needs concentration, a passage of 100 words or so at most is recommendable for one sitting.

As a last remark, I would like to emphasize again that errors should be regarded as normal and inevitable features in L2 learning. Especially in L2 listening, individual difference in developing listening skills will obviously appear. Some students show remarkable development without difficulty and others do not, in spite of their efforts. Developmental patterns of errors in an individual student should be taken as evidence not of failure but of achievement. Teachers' constant encouragement will help students to learn L2 actively.

### Notes

- i ) Japanese is often quoted as a typical mora-timed language. In Japanese, the coda represents one mora. In English, it is clear that the codas of stressed syllables represent a mora, but it is not clear whether the codas of unstressed syllables do. Most Japanese dialects including the standard use moras as the basis of the sound system rather than syllables.
- ii ) Basically non-stressed words are classified as determiners, auxiliary, prepositions, conjunctions and pronouns.

### References

- Anderson A., Lynch T. (1988) *Listening*. Oxford, England: Oxford University Press.
- Anderson-Hsieh, J. (1994). Interpreting visual feedback on suprasegmentals in computer assisted pronunciation instruction. *CALICO Journal*, 11(4), 5-22.
- Bertinetto, P.-M. (1989). Reflections on the dichotomy "stress" vs. "syllable-timing." *Revue de Phonetique Appliquée*, 91-93, 99-130.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge, England: Cambridge University Press.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410.
- Di Cristo, A. (1998). Intonation in French. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp. 195-218). Cambridge, England: Cambridge University Press.
- Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19(2), 193-212.
- Goldinger, S. D. (1997). Words and voices: Perception and production in an episodic lexicon. In K. Johnson & J. W. Mullennix (Eds.), *Talker variability in speech processing* (pp.33-66). San Diego, CA: Academic Press.

- Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context and talker variability. *Applied Psycholinguistics*, 24, 495-522.
- Hardison, D. M. (2000). The neurocognitive foundation of second-language speech: A proposed scenario of bimodal development. In B. Swierzbin, F. Morris, M. E. Anderson, C. A. Klee, & E. Tarone (Eds.), *Social and cognitive factors in second language acquisition* (pp.312-325). Somerville, MA: Cascadilla Press.
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York: Newbury House.
- Hirst, D. J., & Di Cristo, A. (1998). A survey of intonation systems. In D. Hirst & A. Di Cristo (Eds.), *Intonation systems: A survey of twenty languages* (pp.1-44). Cambridge, England: Cambridge University Press.
- Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.), *New Sounds 90* (pp.72-97). Amsterdam: University of Amsterdam.
- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242-1255.
- Molholt, G. (1988). Computer-assisted instruction in pronunciation for Chinese speakers of American English. *TESOL Quarterly*, 22(1), 91-111.
- Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25(3), 481-520.
- Munro, M. J. (1995). Nonsegmental factors in foreign accent. *Studies in Second Language Acquisition*, 17(1), 17-33.
- Palmer, C., Jungers, M. K., & Jusczyk, P. W. (2001). Episodic memory for musical prosody. *Journal of Memory and Language*, 45, 526-545.
- Pennington, M. C., & Ellis, N. C. (2000). Cantonese speakers' memory for English sentences with prosodic cues. *Modern Language Journal*, 84(3), 372-389.
- Pennington, M. C., & Esling, J. H. (1996). Computer-assisted development of spoken language skills
- Rost, M. (1999). Listening in a second language. In B. Spolsky (Ed), *Concise Encyclopedia of Educational Linguistics*(pp.290-295). Oxford, England: Elserier Science Ltd.

## Appendix

The following is the script of the dictation test.

Tell me where you folks are from?

We're from Washington, D.C.

Washington, D.C. And how long are you in San Francisco for?

Just a week. We are in san Francisco for a few days, and then we are heading north to the wine country tomorrow.

Napa Valley?

Yes, and Sonoma.

You like the wine?

Yes.

Good choice. Tell me, did you come for wine or did you come for San Francisco?

Both.

Nice choice, nice choice. Tell me some of the things you've liked so far.

So far? I guess the cable car ride was a lot of fun. And the restaurants are great.

Did you have some seafood?

Uh, yes. We had sushi the first night actually.

Sushi? Crabs?

Yes, I ate lobster and crab last night. It was great.

You've had a lot of fun, lot of fun. What are you going to do today?

See Alcatraz, and then tour around the water front district a little bit; go into the parks.

Tell me, San Francisco a little bit different from Washington? Same?

Very different. The hills make this place so beautiful. There's so much culture and so many things to do that are very easy to walk to.

You feel relaxed when you are here?

Oh, yeah, it's a great place. (214 words)