

Differences in Comprehension: Visual Stimulus vs. Auditory Stimulus

Tomokazu Nakayama

Josai University

Akira Iwata

Hokkaido Otaru Choryo Senior High School

Abstract

The purpose of this study is to investigate whether different modalities of the same text cause differences in the comprehension of the text. Following prior studies, this study used the sentence *I won't leave the room* as the visual and the auditory stimulus, and asked 61 participants to take reading and listening comprehension tests separately. The comparison of the results of the two comprehension tests revealed that the reading comprehension was better than the listening comprehension of the sentence ($p < .05$). The detailed analysis revealed that the irrelevant phonological information of L2 that was influenced by L1 caused the gap between the reading and listening comprehension. The findings of this study suggest that training related to updating L2 sounds could help bring up the listening comprehension level of a learner to the level of his/her reading comprehension.

Key words: *listening comprehension, reading comprehension, L1 influence, interlanguage*

Phonological differences between English and Japanese are among the factors that affect the listening comprehension of learners of English as a Foreign Language (EFL). This paper tries to demonstrate how phonological differences affect listening comprehension.

Influence of L1 on the learning patterns of L2 sounds: The case of Japanese EFL learners

The influence of L1 cannot be avoided when learning L2 (e.g., Best, 1994, 1995; Hancin-Bhatt, 1994). The L2 sound information is usually mapped on to the categories of the native sounds and the influence of L1 cannot be avoided (Fledge, 1995; Hancin-Bhatt, 1994; Suzuki & Okuno, 2009). According to Shirahata, Wakabayashi, and Murano (2010), when a vowel is represented as V, and a consonant as C, the basic English syllable structure is CVC (e.g., *put*, *book*) or CCVCC (e.g., *stand*, *crush*). On the other hand, the basic Japanese syllable structure is CV (e.g., *kumo*(*cloud*), *sora* (*sky*)). Especially at the beginner level, Japanese EFL learners tend to add a vowel to the end of English words that end in a consonant, which often makes the learners misunderstand the text.

The differences in the rhythmic systems of languages also affect the listening comprehension of learners. The Japanese language is considered a syllable-timed language, in which all the syllables are pronounced at equal intervals. On the other hand, English is defined as a stress-timed language, in which the intervals between a stress and another stress, i.e., interstress intervals (ISI) are preserved equally. To preserve the ISI, the words that are not stressed are often pronounced weakly. English words can be divided into two categories: function words and content words. Function words, which include pronouns, articles, conjunctions, relatives, auxiliary verbs, and prepositions, mainly refer to the functions of sentences (Koike, 1993; Roach, 2000). Content words refer to those words that are directly concerned with meanings, including nouns, verbs, adjectives, adverbs, and exclamations. Since content words are often stressed, they are rarely weakly pronounced (Roach, 2000). However, function words are often pronounced weakly to preserve the ISI (Sudo, 2010). This difference between English and Japanese is often said to affect the listening comprehension of Japanese EFL learners (Koike, 1993; Sudo, 2010).

Influence of phonological differences on the cognitive processes of listening

The phonological differences between L2 and L1 are among the factors that lead to an unbalanced use of the cognitive processes (Lynch, 1998, 2002; Mendelsohn, 1998). Two kinds of cognitive process are often referred to in the context of listening comprehension. One is called top-down processing, in which learners try to understand the text based on prior knowledge and context. The other is called bottom-up processing, in which learners try to understand the text from scratch; i.e., learners try to

understand the text by decoding it from phoneme to discourse, without depending on prior knowledge or context. For L1 comprehension, these two processes are automatically used; however, the limited L2 knowledge often leads to the unbalanced use of these two processes (Lynch, 1998, 2002; Mendelssohn, 1998). Even advanced learners are found to rely on top-down processing (Tyler, 2001). Effective methods or approaches that facilitate each process separately are needed to help the learners effectively use these cognitive processes (Goh, 2002).

Differences in the degrees of reading and listening comprehension

Mecartty (2000) investigates whether the size of vocabulary knowledge correlates to reading and listening comprehension abilities. The results suggest that both reading and listening comprehension abilities correlate to the size of vocabulary knowledge; however, the correlation between the size of vocabulary knowledge and the listening comprehension ability is weaker ($r = .38, p < .05$) than the correlation between the size of vocabulary knowledge and the reading comprehension ability ($r = .50, p < .01$). This study suggests that visual texts have more chances of being comprehended than auditory texts.

The differences in the degree of the correlations could be caused by the gap between the phonological information that the learners have in their prior knowledge and what is present in the actual presented auditory stimuli. That is, even if the learners had the wrong phonological knowledge of a particular word, the possibility exists that they would match the word and the meaning in a reading comprehension exercise. However, since the wrong phonological knowledge does not match the auditory text, comprehension of the text would not be attained.

Though prior studies suggest that the gap in comprehension could depend on the differences in the modality presented to the learners, very few studies report the gap in comprehension using the same text in different modalities. This study focuses on the English words *won't* and *want*. Inuzuka (2010) points out that the word *won't* is often mistaken for *want* because of L1 influence. The results reveal that Japanese EFL learners often retained the pronunciation of *won't* /wount/ as /want/ or /wanto/, and often mistook the word *won't* for *want to* because of the L1 influence. However, the task that Inuzuka (2010) used for his study was dictation; he asked the participants to dictate the sentence *He won't listen to our advice*. Dictation is an exercise in which the participants are asked to write down exactly what they hear. This exercise does not require the participants to measure their comprehension of the text. Therefore, the

results of this study cannot confirm the differences in the degrees of the reading comprehension and the listening comprehension of the same text.

The current study utilized the sentence *I won't leave the room* as the stimulus; the sentence was presented in two different modalities, visual and auditory, to investigate whether there were any differences in the degree of comprehension. This study tried to answer two research questions.

RQ1. Does the presentation of *I won't leave the room* in different modalities affect the degree of the comprehension of the sentence?

RQ2. If yes, is the difference in comprehension caused by irrelevant phonological information in prior knowledge that is influenced by the L1?

Method

Participants and design

Sixty-one university students (male: 31; female: 30) participated in this study. They participated in my weekly English class, intended to improve the students' Test of English for International Communication (TOEIC) scores. The participants had six years' experience of learning English in Japanese junior and senior high schools. No participant had the experience of learning English overseas. The fluency level of the participants' English was at a low intermediate level, approximately. The study was conducted in the middle of July, after the conclusion of the English class.

Equipment

To present the visual stimulus, a Panasonic personal computer and a Sony projector were used in this experiment. The visual stimulus was made using Microsoft PowerPoint. The auditory stimulus was presented using the PA system in the classroom with the same personal computer.

Material for the reading and listening comprehension tests

The sentence *I won't leave the room* was used as the stimulus. Three Japanese sentences, 私は部屋を出たい ('I want to leave the room'), 私は部屋を出ない ('I won't leave the room'), and 私は部屋へ行きたい ('I want to go to the room'), were given as the choices. These three choices were used for both the visual and the auditory stimuli. Two slides were made for the test; one slide was for presenting the stimulus visually, and the other was for presenting the three Japanese sentences. Microsoft PowerPoint was used to make these slides. To make the auditory stimulus, a Canadian teacher who teaches English at a university was asked to make a recording. An

Olympus IC recorder (Voice-trek V-22) was used to make this recording. The recording was made thrice. After recording the stimulus, the teacher and I listened to the three recordings and chose the one recording that both of us thought would be the most natural for this experiment. The measured speech rate was 140 w/m. The recording was converted into a digital file so that it could be presented using the computer.

Procedure

The study was conducted after the conclusion of the TOEIC English class in mid-July 2011. The objectives of this study were clearly explained to the participants prior to the study. Only those students who agreed to the objectives participated in the study. After confirmation of participation, the stimulus was presented visually on the screen, and the participants were asked to choose the Japanese sentence that best matched the meaning of the stimulus. The participants were asked to write the number of the Japanese sentence that they chose on the paper that was given prior to the study. After confirming that every participant had written down their answer, the participants were asked to listen to the auditory stimulus and to choose the Japanese sentence that best matched the meaning of the stimulus. The participants were asked to write the number of the Japanese sentence that they chose on the same paper that had been used for answering the question related to the visual stimulus. All the answer sheets were collected after confirming that all the participants had written their answers on the paper, and the end of the experiment was announced.

Results and Discussion

RQ1. Does the presentation of *I won't leave the room* in different modalities affect the degree of the comprehension of the sentence?

In order to answer RQ1, all the scores of the test were compiled into a Microsoft Excel spreadsheet, and were analyzed. The results of the analysis are shown in Table 1.

Table 1. Results of the Comprehension Test

	<i>N</i>	<i>M</i> (%)	<i>SD</i>	<i>df</i>	<i>t</i> -value	<i>p</i> -value	<i>r</i>
Visual	61	62.30	0.49	60	3.57	0.00**	0.42
Auditory		34.42	0.48				

Note: ** $p < .01$

Of the 61 participants, thirty eight answered the visual question correctly ($M = 62.30$; $SD = 0.49$), and 21 participants answered the auditory question correctly ($M = 34.42$; $SD = 0.48$). A dependent t -test revealed that the number of participants who had answered the visual question correctly was significantly larger than the number of the participants who had answered the auditory question correctly ($t(60) = 3.57$; $p < .01$; $r = 0.42$). A comparison of the results of the two comprehension tests revealed that the participants' reading comprehension of the sentence was better than their listening comprehension, which indicates that the sentence *I won't leave the room* in different modalities did affect the degree of comprehension of the text.

RQ2. If yes, is the difference in comprehension caused by irrelevant phonological information in prior knowledge that is influenced by the L1?

In order to answer RQ2, the proportions of the number of actual choices made by the participants were calculated and analyzed in detail. The results are shown in Table 2.

Table 2. Number and Proportion of Choices Made by the Participants

The choices	Visual		Auditory	
	N	M	N	M
1.部屋を出たい ('I want to leave the room')	10	16.39%	21	34.43%
2.部屋を出ない ('I won't leave the room')	38	62.30%	21	34.43%
3.部屋へ行きたい ('I want to go to the room')	13	21.31%	19	31.15%

Total $N = 61$

The fact that 13 students selected choice 3 for in the visual test indicates that there were students who did not know the meaning of the word *leave*. The fact that 10 students selected choice 1 in the visual test indicates that there were students who misunderstood *won't* for *want to*, even in the visual test. Interestingly, in the auditory test, the number of students who selected choice 3 ($N = 19$; $M = 31.15$) increased, compared to number of students who made the same choice in the visual test ($N = 13$; $M = 21.31$). Inugai (2010) points out that if students have to pay a lot of attention to some unclear section of speech, they tend to make unusual mistakes. The detailed analysis of this study revealed that 11 students who selected choice 2 in the visual test selected choice 3 in the auditory test; this indicates that the students selectively paid attention to the word *won't* and they could not pay attention to the word *leave*, or the information was simply lost.

The most interesting result is that the same number of students selected choice 1 ($N = 21$; $M = 34.43$) and choice 2 ($N = 21$; $M = 34.43$), and the total of these numbers almost equals the number of students who selected choice 2 in the visual test ($N = 38$; $M = 62.30$). This suggests that about half of these students recognized the meaning of the word *won't* when it was presented visually; however, the same students could not recognize the meaning of the word when it was presented orally, because they confused the auditory stimulus *won't* /wount/ as /wanto/, which is irrelevant phonological information influenced by L1. This also clearly explains why the results of visual test were better than the results of the auditory test, even though the same sentence was used as the stimulus. Thus, it can be concluded that the influence of the L1 phonological system caused differences in the comprehension of visual and auditory input.

Conclusions and implications

The purpose of this study was to investigate whether the same text presented in different modalities would cause differences in comprehension. The sentence *I won't leave the room* was used as the visual and the auditory stimulus. The results of the two comprehension tests reveal that the participants' reading comprehension of the sentence was better than their listening comprehension. A detailed analysis was conducted to verify the reasons, which revealed that listening comprehension was affected by irrelevant phonological information in prior knowledge that was influenced by the L1. The results of this study confirm the findings of prior studies that L2 sounds cannot avoid the influence of L1 (Fledge, 1995; Hancin-Bhatt, 1994; Suzuki & Okuno, 2009). Though the findings of this study suggest that irrelevant phonological information influenced by L1 could be one of the factors that caused the differences in reading and listening comprehension (Mecarty, 2010), further research is required to validate this. The results of this study indicate that training in updating L2 phonological information could bring up the level of listening comprehension at least to the level of reading comprehension.

References

- Fledge, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience* (pp. 233-277). Baltimore: York Press.
- Goh, C. M. (2002). Exploring listening comprehension tactics and their interaction patterns. *System*, 30, 185-206.
- Hancin-Bhatt, B. (1994). Segment transfer: A consequence of a dynamic system. *Second Language Research*, 10, 241-269.
- Inuzuka, H. (2010). Nihongo bogowasha niyoru eigo onsei no choukai handan. *Iwatedaigaku Eigokyouiku Ronshu*, 12, 46-51.
- Koike, I. (1993). Eigo no hiaringu to shido. Tokyo: Taishukan shoten.
- Lynch, T. (1998). Theoretical perspectives on listening. *Annual Review of Applied Linguistics*, 18, 3-19.
- Mecartty, F. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning*, 11, 323-348.
- Mendelsohn, D. (1998). Teaching listening. *Annual Review of Applied Linguistics*, 18, 81-101.
- Roach, P. (2001). *English Phonetics and Phonology: A Practical Course*. New York: Cambridge University Press.
- Shirahata, T., Wakabayashi, S., & Murai, H. (2010). Shosetsu daini gengo shutoku kenkyu riron kara kenkyu hou made. Tokyo: Kenkyusha.
- Sudo, M. (2010). Eigo no onseishutoku no seisei to chikaku no mekanizumu. Tokyo: Kazama shobo.
- Suzuki, A., & Okuno, I. (2009). Subvocalization of Japanese students when listening to English. *The Language Education Center of Josai University Bulletin*, 3, 63-72.
- Tyler, M. (2001). Resource consumption as a function of topic knowledge in nonnative and native comprehension. *Language Learning*, 51, 257-280.