論文

The 'Cornfield Test' : A Study into the 'Katakana Effect' and the Intelligibility of Japanese Pronunciation of English

Alison Kitzman
(Kinki University)

"The significance of a rare interview with Yoko [Ono] was dissipated by the difficulty of deciphering her mauling of the English language."

- Adam Sweeting, reporter for the Guardian newspaper

0. Introduction

Historically, pronunciation has held a highly regarded, theoretically based, position in language teaching. However, unlike continental Europe, today, many English language teachers in Great Britain, America and Japan, tend to ignore explicit pronunciation teaching. There is a common belief that students' pronunciation is too difficult to change. Moreover, when pronunciation is taught, the teaching often focuses on sound-level phonemes rather than what students need to make their language actually communicative. Though Lenneberg's (1967) Critical Period Hypothesis (CPH) posited post-pubescent difficulties in learning pronunciation, this concept has more recently been shunned for being merely an excuse for not learning. Fledge (1987) found CPH to be counter-productive to research because varying factors among individual learners are not taken into account. Nevertheless, newer research supports a weaker form of Lenneberg's ideas. Archibald (2002) posited that instead of creating all new phonological parameters for the second language (L2), learners simply modify existing parameters. Learners' pronunciation errors are therefore not considered to be just random attempts to produce new and unusual sounds. Instead, these pronunciation errors are "reflections of the sound inventory, rules of combining sounds and the stress and intonation patterns of their native languages" (Swan and Smith, 1987).

The study of pronunciation involves the perception, production, and prediction of both segmentals (i.e., phonemes or sounds) and their interactions with the suprasegmentals found in natural speech (i.e., rhythm, stress, length, intonation, syllabification, and tone). Since the mid 1980's, traditional methods of teaching pronunciation through the drilling of discrete items have given way to methods that focus more on how pronunciation affects communication at the discourse level. This is noted by pragmatists (c.f., Thomas, 1995; Jenkins, 2000; and Shockey, 2003), who have shown that breakdowns in communication are most often due to how meaning is interpreted in the context. By incorporating pronunciation practice at the discourse level and by concentrating on the meaning in context, pronunciation training has made a minor comeback and sentence-level suprasegmentals have gained the spotlight. However, teachers should not necessarily take this as a mandate

to teach pronunciation, since there may be valid reasons for learners to not perfect their L2 accents. Daniels (1997) argues that learners might resist adopting a native-like accent due to individual or social pressures.

Nevertheless, teachers and students should be aware that faulty pronunciation can lead to communication breakdown. Jenkins (2000) points out that before pragmatic elements even come into play, communication will fail if faulty pronunciation has impeded a word's intelligibility. Studies by Magen (1998) and Anderson-Hiseh, et al (1994), that had native English speakers judge the intelligibility of non-native speakers' utterances, found that prosodic features (i.e., aspiration, place or manner of phoneme, etc.) rather than phonemic features were more significant barriers to listener comprehension.

0.1 Interlanguage and pronunciation teaching in Japan

Due to socio-cultural and historical factors, both current and traditional models of 'proper' pronunciation worldwide still prefer established varieties, such as British English, to newer *nativized* varieties, such as South African English. In Japan the tendency is toward American English, possibly because of the emphasis placed on the TOEIC test for school and job placement. According to the TOEIC Report on Test-Takers Worldwide 2005 (2006), 65% of all TOEIC takers were Japanese. As important as this test may be, the Japanese do not formally start to learn English until their first year of junior high school, nor do they have a significant bilingual population that interacts daily at a national or international level. With few models of Standard English, the Japanese variety of English (JVE) still incorporates large elements of the first language (L1). Corder (1971) termed this phase of language learning *interlanguage*.

Besides phonological differences, cultural and affective factors contribute to the development of the JVE. One issue specific to the Japanese language is its written system and how it affects the pronunciation of transcribed English words. Recently, through music, television, movies, and the Internet, the Japanese language has experienced a large influx of foreign words, mostly English. These words are first transcribed into the Japanese *katakana* syllabary, or the Japanese *romaji* alphabet, and subsequently enter the Japanese language in a phonologically transformed state. This transcription crutch can be seen wherever foreign words appear such as on signs, in *karaoke* music, on menus, and even in some language textbooks. More accurate transcriptions can be derived from a standard phonetic symbol system, such as the International Phonetic Alphabet (IPA), which remains true to a language's phonemes. Dickerson (1989), however, advocates a simplified Roman alphabet transcription method because systems like the IPA add a layer difficulty by forcing learners to learn new symbols. *Katakana* is an easier method of phonetic transcription, yet it leads Japanese learners to believe that English employs the same phonetic system as Japanese; this, of course, can develop into a significant barrier to natural English pronunciation.

In addition to the detrimental effects of *katakana* transcription on Japanese learners of English, educational factors play a role. Outdated teaching beliefs and methodologies about pronunciation, a lack of pronunciation education in school curricula, and poor models of Standard English pronunciation contribute to communicative difficulties. Textbooks

at any level of English education in Japan rarely contain practice on the sentence-level (i.e., suprasegmental) phonetic phenomena that most commonly disrupt intelligibility. Even when included, practice often contains incorrect or vague explanations, and single specific examples are shown rather than broad and productive rules. Furthermore, teachers competing for valuable class time with other core subjects can rarely devote time during an already busy English class for pronunciation.

Finally, personal identity can also inhibit phonological mastery. English, as a whole, though highly regarded, is not considered essential to daily life in Japan to the degree it is in multicultural Europe, Africa, or other parts of Asia. These places see a greater number of cross-cultural interactions. Pronunciation proficiency must be high enough to overcome the L1 phonetic influences of both the L2 listener and the L2 speaker. However, without a need to interact in English, people of a near monolingual and monocultural island nation such as Japan may not perceive a need to acquire near-native pronunciation to function within their local communities or at work. While acknowledging that these conditions exist, let us also acknowledge the unlikelihood of these issues being changed significantly by the classroom teacher. Instead, let us focus on what changes can be affected.

0.2 Prosody and the *Katakana* Effect

As sound-level errors can usually be understood from context, it is the sentence-level suprasegmental features and errors that need closer scrutiny. Suprasegmentals are the individual features and phenomena associated with, and between, whole syllables, words or phrases such as syllable boundaries, assimilation, and linking. A more general term used at the sentence-level of pronunciation is prosody, which refers to the rhythm of spoken language including intonation and stress. One aspect of this, important in the discussion of the JVE, is isochrony, or rhythmic speech. Isochrony is the tendency for stressed syllables to recur at regular intervals. Jusczyk (1997) showed that babies could recognize the difference between the prosody of their mother's language and other languages. Though some differences among languages exist, there seem to be universal aspects of rhythm, arising in early childhood, which appear to be linked to some unknown cognitive support. A cyclical repetition of a consonant (C) vowel (V) syllable in a CVCV pattern, like '*dada*' '*mama*' and '*baba*', can be observed in babbling babies across languages according to Mahler, et al (1986). Support for the hypothesis that the metrical patterns of speech are innate in the human brain, and that they are maintained by "cognitive oscillatory structures" was shown by Port, et al. (2002).

One facet of isochrony is the rhythmic timing of language. English is thought of as a *stress-timed* language, wherein the words most important to the context are spoken with greater volume and duration (i.e., stress), and those words are what dictate the language's rhythmic intervals. In English, the unstressed parts of the speech stream change phonologically at the syllabic level to accommodate these rhythmic intervals. Most often this is done by the reduction to schwa /ə/ of vowels in unstressed syllables. Unfortunately, neither this phenomenon nor the phoneme /ə/ exist in Japanese. In contrast, the Japanese language is *mora-timed*, wherein equal intervals are dictated by the *mora*, a Japanese equivalent of

the English syllable. Nazzi, Bertoncini, and Mehler (1998) showed that French newborns were able to discriminate among Japanese mora-timed, English stress-timed, and their own syllable-timed language rhythms. Bertinetto (1980) pointed out that the inter-stress intervals of syllable-timed languages, such as French, maintain their original phonetic features. Though prosodic phenomena exist in Japanese, such as vowel elision and consonant shifting, those phenomena common in English, like blending, trimming, and especially vowel reduction, are not present in mora-timed Japanese.

Recent studies have shown many conflicting theories regarding the division between stress and syllable timing. Roach (1982), Dauer (1983), and Cauldwell (2002) have all found that English does not, in fact, adhere to a strictly timed system of intervals. This is especially true for speech production. On the other hand, comprehension of non-native rhythmic patterns by native listeners does interfere with intelligibility as shown by Cutler's (1993) study on rhythms and word stress. Cutler also points out that this system of dissecting speech is so imbedded in the brain since infancy that it tenaciously remains in L2 speech patterns forever. Though a quantifiable confirmation of a strong form of isochronic theory continues to elude researchers, linguists do agree to a soft form of this called rhythmic speech. Jenkins (2000) still includes 'stress-timing' in her list of phonological core material. Vanderplank (1993) argues that the concept of isochoric speech has a "useful psychological reality," especially for advanced learners.

Another prosodic feature separating the standard and the Japanese varieties of English is the phonetic unit. In English, a phonetic unit, or *syllable*, is comprised of an obligatory vowel, which may be surrounded by consonants or by consonant clusters before or after the vowel. This results in syllables with consonant and vowel combinations including V, CV, CVC, CCVC, and CCVCC. Unlike English, where there is an obligatory vowel, the Japanese phonetic unit, or *mora*, consists of either a vowel, possibly preceded by a consonant, the *moraic consonant*, ん , or the beginning of a geminate (e.g., double consonant <tt>, etc.). This constrains Japanese to V, ん , CV, and CVCV patterns. So, an English speaker would count two syllables in 'plas-tic' (CCVC-CVC), but a Japanese speaker of English would count five morae 'p(u)-ra-s(u)-ti-k(u)', with superfluous vowels making a CV-CV-CV-CV-CV pattern for the same word. Japanese listeners were found to have difficulty in both the perception of, and the counting of, English syllables by Erickson, et al. (1999), which the researchers attributed to the fact that much of the listeners' negative language interference resulted specifically from originally having learned English pronunciation by way of *katakana*. Ohata (2004) states that Japanese speakers of English use a vowel insertion strategy to conform to the Japanese open-syllable pattern. This seems to be a "natural reaction to the difficulties in pronouncing consonant clusters," but Ohata also argues that learners are probably unaware of doing so. Ishiikawa (2002) found that participants trained to recognize English syllabification were able to segment English words and even non-words in a more syllabic than moraic manner, suggesting that such phonological tendencies can be overcome.

0.3 Purpose

Though mistaken usage of either segmental or suprasegmental features can inhibit intelligibility, it is the suprasegmental features that do so most. The differences between Japanese and English rhythmic timing, and the all-open nature of Japanese morae, contribute to negative L1 transfer, which seriously impedes the prosodic features of natural sounding speech in English. Other factors, such as using *katakana* to transcribe foreign words and relying on it as a crutch in the early stages of learning English, have caused a 'Katakana Effect' in pronunciation; superfluous vowels after word-final consonants and between consonant clusters. This Japanese tendency of adding extra vowels between consonants in consonant clusters and after word-final consonants to create an L1-like CVCV patterns in English is casually referred to as *katakana eigo*. With the addition of extra vowels, the natural prosodic features native speakers use between words to keep the stressed-timed rhythm (primarily vowel reduction and linking) become impossible, which seriously affects comprehensibility. This Katakana Effect may be what interferes most with the intelligibility of the JVE to the average native speaker of English. For this study the researcher specifically set out to determine:

1) whether average native speakers of English, in this case Americans, having little to no experience with the Japanese language, culture, or people, could comprehend the Japanese variety of English.
2) which type of pronunciation errors, superfluous vowels resulting from the Katakana Effect or individual phonemes, caused the most miscomprehension.
3) whether incomprehensibility could be overcome with additional context clues.

1. The Present Study

1.1 Hypotheses

1) A typical American with little or no knowledge or experience with Japanese learners of English will have more difficulty in understanding the Japanese variety of English than a person who has such experience.
2) Incomprehensibility at the sentence level is due more to Katakana Effect errors (i.e., superfluous vowels), based on negative L1 transfer, than to problems with individual phonemes (i.e., /l/ vs. /r/, /v/ vs. /b/, etc.).
3) An increase in the number of contextual clues provided to the listener improves intelligibility (i.e., whole phrases vs. single words).

1.2 Subjects

There were 25 volunteer participants, 22 of whom were Americans. These subjects were divided into two groups. One group of five subjects, who had studied the Japanese language and were familiar with the Katakana Effect, served as a small control group. This group was called the Studied Japanese group (SJ). The SJ was comprised of three females (one American, one Australian, and one bi-lingual naturalized American of Japanese descent), and two American males. The group's average age was 42.8. They had lived, on average, 17.3 years in Japan and, except for the naturalized American female, had formally studied

Japanese or another L2 for an average of 3.9 years. Three were living in Japan, and two had moved from Japan to the American state of Maine five years previously. The naturalized American female living in the States had not returned to Japan in five years. Three of this Japanese-experienced SJ group listened to Voice One and two listened to Voice Two.

The experimental group consisted of 20 subjects with very limited experience with Japan. This group was called the No Japanese group (NJ). All had lived in more than one area in the US, and all but four had traveled abroad. At the time of the study, they were all living in the states of Idaho, Oregon, or Illinois. Of this group, 14 were female and six were male. Their average age was 51.8. Nineteen of them had studied at least one foreign language, other than Japanese, for an average of 3.3 years. Sixteen of these participants had traveled abroad, but only five had been to Japan. The average total time of their stay in Japan was three weeks, and none of these travelers had had much direct interaction with the Japanese during their stays. One male in this group was a Spanish national who had studied English and had lived in the United States more than 11 years. Of the 20 NJ group members, half listened to Voice One and half to Voice Two.

1.3 Measure: The *Cornfield Test*

A *Cornfield Test* is a term commonly used among native (NS) and non-native (NNS) English speaking Japanese language and culture teachers and ESL teachers at the University of Illinois, Urbana-Champaign. It describes an informal method aimed at determining whether an average American citizen, such as a corn farmer, with little international experience, can understand the English spoken by a non-native speaker. This test was intended to mimic a brief interaction between an average native speaker of English and a Japanese speaker using *katakana eigo,* as might happen on the street or in a shop in Middle America.

For this study, the researcher created a four-part listening dictation (see Appendix) as a more formalized *Cornfield Test.* Two female native Japanese speakers were digitally recorded speaking English with a heavy *katakana*-affected accent. Both speakers' English is normally accented with *katakana*, and they were instructed to read the way they would naturally speak, rather than to concentrate on proper pronunciation. Two speakers were used in the event that other phonological features affected listener comprehension.

Part A consisted of ten one or two syllable words spoken twice, which gave the subject-listeners no context to aid in comprehension. The subjects were asked to write down what they heard in the space provided on the test answer sheet. Part B was a similar word list dictated twice, though context was suggested by providing two or three written multiple-choice options from which the subject listeners were asked to select. Answer choices tested whether the subject heard a phonetic or a Katakana Effect variation in the recorded word list. For example, the voice said, /sports/, and the choices were, a. physical activity (/sports/ correct answer), b. to encourage (/supports/ testing for consonant cluster vowel insertion), and c. small round mark (/spats/ testing for the phonetic replacement of /r/ with /a/). Part C consisted of context-rich single sentences spoken twice. For example, 'The woman hurt her thumb on the hot fry pan.' This was to enable the subject to guess at the individual

words better from the context. The subjects were asked to write down what they heard in the space provided on the test answer sheet. The Part D dictation provided the most context in the form of a context-inclusive short dialog dictated twice in which the first speaker (the researcher) was a NS and the second speaker was one of the two native speakers of Japanese. The words for both the Part A and B word lists and the Part C and D sentences were chosen specifically for their commonly mispronounced phonemes and or for their word-final consonants and consonant clusters. The nine most commonly mistaken phonemes by Japanese were included: /l/, /r/, /th/, /æ/, /a/, /wu/, /v/, /f/, and /si/.

1.4 Procedures

Subjects listened once to the recorded dictations in a quiet room either through an iPod with ear-buds or through a Macintosh computer sound system. Before taking the test, participants were informed only that there were four parts to the test and were instructed to write down the list of words, Part A, or the sentences, Parts C and D, they heard, in the space on the answer sheet. Part B was multiple-choice so the participants were asked to circle the best choice.

A few problems are inherent in a study like this. First, due to geography, such tests may not be carried out with absolute consistency. Distractions, sound quality, subject affect, and other factors may vary, though not to a degree that should cause major variation in a participants' answers. A further problem may be that listening to a dictation is not a normal form of communication. There is no body language, facial expression, opportunity for repetition, nor time between items for the listeners to reflect on what the meaning might be. A further issue is that even though the contextual clues between parts increase, the sentences are isolated so subjects have little recourse to guess by way of their background knowledge. One difficulty of this study arose wherein spelling by the subject was not always clear. For example, /u/ was spelled 'oo', 'ow' or 'ou', so the benefit of the doubt was given to the subject. Only twice was handwriting illegible, so it was determined by a consensus of the researcher's peers.

1.5. Analysis

When analyzing the data, two things were examined; whether additional context aided in comprehensibility, and whether errors in comprehension came from mistaken phonemes or from superfluous vowels caused by the Katakana Effect. In order to determine the degree to which the subject-listener understood the NNS's dictation, the researcher compiled data not only on the number of correctly understood responses, but also the number of responses that were 'real' English words, even if they were not the ones intended by the NNS. For example, instead of the correct word 'that' (Part A, number 1) some subjects wrote 'desktop,' 'bathtub,' and 'data.' These real words, attributable to some degree of communicative competence, were counted separately because the subject-listener presumably believed that was the word they heard. In an authentic situation, these 'real' words could provide a base for both parties to negotiate meaning. Non-words such as 'thatthough,' zatzo,' and 'deuto,' however, were not counted as 'real' or 'correct.'

These were considered to be wild guesses or phonetic transcriptions indicating a total lack of comprehension. In an authentic situation, this could be the root of a breakdown in communication. The results of the SJ (experience with the JVE) were then compared with the NJ (little or no experience with the JVE) to determine whether having experience with the JVE aided in comprehensibility.

Out of the 200 total possible words in the Part A (i.e., no-context word list) the 20 NJ participants only comprehended and wrote four 'correct' words, a total of two percent. The five SJ participants correctly understood and wrote 32 words out of the 50 total possible words for a total of 64% 'correct' words. Both the NJ and SJ wrote about 25% total 'real' words indicating that both groups believed that they understood the correct word at equal rates. When scores of both 'correct' and 'real' words are totaled, the NJ had only 27% but the SJ had 90%. The NJ scored 63% lower than the SJ, presumably because the subjects had previous experience with the phonological errors inherent in the JVE.

Figure 1: Part A – No Context Word List

Word	The Correct Word Guessed		A Real Word Guessed	
	No JPN	Studied JPN	No JPN	Studied JPN
that	0	5	4	0
first	0	1	5	3
trouble	1	4	2	1
steam	0	5	2	0
woman	2	5	1	0
love	0	4	4	0
harm	1	3	10	2
mouth	0	1	5	4
sin	0	2	15	3
tax	0	3	2	0
Total Words Guessed Correctly	4	32	50	13
Total Words Possible	200	50	200	50
Percent of Words Guessed Correctly	2%	64%	25%	26%

To understand whether a lack of comprehensibility in Part A was due to phonemes or superfluous vowels, each type of error was counted. Out of a possible 38 potential errors, an average of 15 phonemic mistakes were made by the NJ, an error rate of 43%. The error rate for the SJ was lower at 4.6, or 13%. Interestingly, the phoneme errors were not always the nine commonly associated with Japanese speakers. Whereas subject-listeners were expected to mistake, and did mistake, the commonly dropped /w/ in 'woman' (number five) and write 'oo-' or 'u-' as spoken by the Voices, other unexpected and unexplainable phoneme errors occurred. Particularly troublesome regardless of the dictating voice was the word 'trouble' (number three), which was written by some NJ as 'flava', 'slasuru', or 'camlu'. As for superfluous vowels, the NJ heard and added an average of 7.75 extra vowels to the 12 possible consonant clusters or word-final consonants, for an error rate of 65%. The SJ were able to avoid most

Figure 2: Part A – No Context Segmental vs. Suprasegmental Errors

Word	Possible Phoneme Errors	Average Phoneme Errors Made		Possible *Katakana* Errors	Average *Katakana* Errors Made	
		No JPN	Studied JPN		No JPN	Studied JPN
that	3	1.25	0.2	1	1	0
first	4	1.75	1.4	1	.95	0.2
trouble	5	3.45	0.6	3	.85	0
steam	4	.85	0	2	.95	0
woman	5	1.60	0	0	0	0
love	3	1.40	0.4	1	1	0.2
harm	4	1.45	0.4	1	.45	0
mouth	3	1.15	0.8	1	.75	0
sin	3	.65	0.8	0	0	0
tax	4	1.45	0.4	2	1.80	0.2
Average Total Correct	38 possible	57% correct	87% correct	12 possible	35% correct	93% correct
Average Errors Made	35 total average	15 errors 43%	4.6 errors 13%	12 total average	7.75 errors 65%	0.8 errors 7%

Katakana Effect pitfalls erring an average of 0.8 times, or only seven percent. Overall, those with experience speaking and listening to Japanese, the SJ, made 30% fewer phoneme-based errors and 58% fewer Katakana Effect errors than those with no experience, the NJ.

Part B increased the level of context by suggesting two or three definitions for the words spoken so participants would not be blindly guessing. Inherent in the choices were the possibility for miscomprehension due to the Katakana Effect and or phonemic errors. The added context allowed the NJ to choose the 'correct' word an average of 3.95 out of eight times, or 49%. Those with experience in the JVE chose the 'correct' word 6.3 times, for 79% accuracy. One variation between dictating voices was found on number 2, 'fly.' Of the 12 who listened to Voice One, 11 answered incorrectly whereas of the 13 who listened to Voice Two, 10 answered incorrectly indicating an expected variation between voices. Overall, the NJ did 22% better choosing the 'correct' word with this suggested context than with no context at all. The SJ also did better on Part B than on Part A, for a 15% improvement. As expected, probably due to their experience with the JVE, the SJ scored 30% higher than the NJ group in choosing the 'correct' word.

Figure 3: Part B – Suggested Context Word List

Word	Times Word Chosen Correctly	
	No JPN	Studied JPN
vase	16	3
fry	9	3
closed	15	5
clothes	14	4
brand	3	4
sports	14	5
8-8	2	4
cracked	7	4
Average Correct	3.95	6.3
Total Percent	49%	79%

Part C's context-rich sentences were to provide clues to the subject-listener to aid in comprehension. Again, counts were broken down into 'real' versus 'correct' words. On average, the NJ wrote a total of 67% 'correct,' or 6.2 out of 9.3 possible 'real' or 'correct' words. The SJ wrote an average of nine, or a near perfect 97%. Interestingly, all subjects faired better after the first sentence. Perhaps the subjects had difficulty switching from the multiple-choice Part B to the dictation in Part C. Significant improvements were made in the understanding of word-final consonants with superfluous vowels. For example, though both voices pronounced the word 'end' as /endo/ in Sentence One, only one of the NJ wrote 'endo.' Similarly, the word 'put,' spoken /puto/, in Sentence Two was written with a final superfluous vowel only three times. Though comprehension increased in the NJ, six people did not write anything for the word 'end' and seven did not write anything for 'put.' This indicates a continued lack of intelligibility of the JVE even when some context was provided. However, with the additional context both groups' scores went up significantly. Both the NJ and the SJ improved at the same rate, 18%, making the difference between the groups a constant 30%.

Figure 4: Part C – Minimal Context Single Sentences Dictation

	Sentence 1		Sentence 2		Sentence 3		Average	
Words Written	No JPN	Studied JPN	No JPN	Studied JPN	No JPN	Studied JPN	No JPN	Studied JPN
Correct Words	3.1	7.2	5.8	8	3.2	9.2	4	8.1
Real Words	1.6	1	1.6	0.8	3.4	0.8	2.2	.9
Not Real Words	0.5	0	0.3	0	0.7	0	0.5	0.1
No Word	4.2	0.6	1.7	0.2	3.5	0	3.1	0.2
Total Average Real or Correct Words	4.7	8.2	7.4	8.8	6.6	10	6.2	9
							67% correct words	97% correct words
Words Possible	9		9		10		9.3	

As with Part A, a simple count of phonetic and Katakana Effect errors was made for Part C. On average, the NJ made 2.5 phonetic errors per sentence, while the SJ made less than one. The phonetic mistakes expected in the Japanese variety were apparent. In Sentence Two, 'clean clothes' was also heard as 'green clothes,' which, though incorrect, did not interfere with rhythm or total comprehensibility of the sentence as a whole. Striking was the fact that zero Katakana Effect errors were made by the SJ, whereas the NJ heard 33 superfluous vowels. Most of these errors were common to all NJ subjects. Single syllable words were often transcribed with vowels added to their word-final consonants creating double syllable words. The word 'hot' became 'hotto,' 'hoto,' or 'hato' in Sentence Three. The word 'hot' caused further trouble as many in the NJ wrote two separate words, creating 'hot dog,' hot tub,' or 'hot box.' The word 'hurt' became 'had 2,' 'heart to,' and 'harto.' Whether the error by the NJ was due to the addition of vowels or syllables, a serious disruption of the rhythm of the sentence occurred leading to the lack of comprehension of the whole sentence. Few subjects wrote even one cohesive sentence.

Figure 5: Part C – Phonetic Effects and Superfluous Vowels in Single Sentences

	Sentence 1		Sentence 2		Sentence 3		Average	
Type of Error	No JPN	Studied JPN	No JPN	Studied JPN	No JPN	Studied JPN	No JPN	Studied JPN
Phonetic	2	0.6	1.7	0.6	3.9	0.4	2.5	0.5
Katakana	0.8	0	1.3	0	1.8	0	1.3	0
Superfluous Vowels	7	0	3	0	23	0	33	0

Part D used a standard American native speaker as the first speaker to provide a complete and easily understood context in each two-line dialog. The NJ obtained a correctness rate of 84% for 'correct' or 'real' words written. The SJ maintained their rate of 96%. The only difficulty, which both the NJ and SJ had, was with Sentence Two, which had been written deliberately to entrap all subjects through their background knowledge usually, a significant aid to contextual understanding. 'Either hoes potatoes,' pronounced /aizah hohz pohteitohz/, was written 'Idaho potatoes.' by nearly half of the subjects regardless of group or place they lived. Nevertheless, with fuller context the NJ narrowed the comprehension gap to 12% less than the SJ.

As for the type of error in Part D, both the number of phonetic and Katakana Effect mistakes decreased due to the greater context provided. The NJ made half the phonetic mistakes, added about half as many superfluous vowels, and created a

Figure 6: Part D – Contextual Short Dialogues Dictation

	Sentence 1		Sentence 2		Sentence 3		Average	
Words Written	No JPN	Studied JPN	No JPN	Studied JPN	No JPN	Studied JPN	No JPN	Studied JPN
Correct Words	8.5	11.8	0.9	1.2	4.4	5.8	4.6	6.3
Real Words	1.6	0.2	1.4	0.8	0.8	0	1.3	0.3
Not Real Words	0.3	0	0.2	0	0.2	0	0.2	0
No Word	2.3	0	0.8	1	1	0.2	1.4	0.4
Total Average Real or Correct Words	10.1	12	2.3	2	5.2	5.8	5.9	6.7
							84%	96%
Words Possible	12		3		6		7	

Figure 7: Part D – Phonetic Effects and Superfluous Vowels in Short Dialogues

	Sentence 1		Sentence 2		Sentence 3		Average	
Type of Error	No JPN	Studied JPN	No JPN	Studied JPN	No JPN	Studied JPN	No JPN	Studied JPN
Phonetic	1.6	0.2	1.2	0.4	1	0	1.3	0.2
Katakana	1.1	0	0.1	0	0.4	0	0.5	0
Superfluous Vowels	3	0	7	0	8	0	18	0

greater number of cohesive sentences compared to Part C. The SJ maintained a zero percent rate of Katakana Effect errors, and lowered their phonetic mistakes to almost zero. Even with improvements due to added context, the NJ still understood the JVE 12% less than the SJ did.

2. Results

This study has shown that native speakers, regardless of having had experience with the Japanese variety of English or not, lack comprehension of a greater percentage words with superfluous vowels inherent in the Japanese variety of English than words with only phoneme errors. However, it is difficult to say without further research which type of error makes overall intelligibility more difficult. It has also been shown here that having experience with learning Japanese raises a listener's ability to comprehend, especially problems with *katakana eigo*. Lastly, this study shows that there is a definite correlation between intelligibility and a greater number of contextual clues given to the listener.

Though this study is meant to check for intelligibility of a NNS English by a NS of English, one participant's L1 was Spanish, another Japanese. Further research on NNS to NNS intelligibility, with English being an L2 for both, is necessary. Additional research should be done to determine to what extent it is the individual phonemes or the superfluous vowels that inhibit intelligibility and to what degree context is needed to overcome this.

3. Classroom Implications and Conclusion

With so many exacerbating factors, why would the Japanese choose to pursue mastery of a standard variety of English? English is no longer just a foreign language; it is an international language. Roach (1994) posited that students in Singapore, which does have a recognized variety of English, also need to learn a standard variety of English in order to be more intelligible to all speakers of English. This may hold true for Japan, too. Keeping in mind that it may not be the goal of the L2 learner to speak effectively with a native speaker, pronunciation instruction in some form is warranted so that non-native speakers, regardless of L1, can be intelligible to each other. Though standardized pronunciation is not a main goal in the English classroom, it is the responsibility of educators in Japan to help students become aware as early as possible in the learning process of which pronunciation features, such as *katakana eigo*, inhibit communication.

Curricula that ignore teaching pronunciation and external sources that use *katakana* transcriptions of English are rarely within the power of the classroom teacher to remedy. Yet, despite the limited class time faced by teachers, several quick and easy points can be taught to enable the learner to speak English with greater intelligibility. First, the focus of pronunciation teaching at all levels of education should be on discourse and sentence-level, and to a lesser degree on word-level, features. Instead of discrete drills, prosody-rich activities, where the pronunciation instruction is incorporated into context-rich activities, should be taught. For example, singing or listening to music is not only beneficial but also more motivationally engaging to students, as found by Kanel (1997). Also, teachers should choose phonetic features to teach based on research of correlation to intelligibility.

Jenkins (2000) posited a set of 'core' pronunciation features to be taught which includes consonant clusters and syllabic structures. In addition, teachers should instruct students, on a metacognative level, that difficult to fix phonetic mistakes do not interfere with communication to the extent that Katakana Effect errors do. For example, students should worry less about their phonemic errors and should instead concentrate on one simple thing: not adding a vowel sound at the end of a word ending in a consonant. Though eliminating these superfluous vowels takes much effort and concentration, which may inhibit students, the resulting slower communication is better than miscommunication. Creation of healthy affective development and learner autonomy can also be instilled metacognatively. This can be done by encouraging an understanding and acceptance of *accent addition*, the addition of L2 pronunciation features to a learner's inventory, to overcome any concern for *accent subtraction*, a potential loss of personal identity. Learners should feel no need to deny their L1 accent, or their L1 identity, but must be prepared to make changes to accommodate intelligibility.

For Japan to maintain its economic and political prominence, English learners must be able to phonologically produce English well enough to be understood by other native or other non-native speakers of English who have not had experience with the localized Japanese variety of the language. From the deserts of the Middle East to the cornfields of the American Heartland in this age of multinational corporations, globalization and the Internet, English is the new *lingua franca*.

References

Anderson-Hsieh, J., Johnson, R. and Koehler, K. (1992). The relationship between native speaker judgments of non-native pronunciation and deviance in segmental, prosody and syllable structure. *Language Learning* 42, 4, 529-555.

Archibald, J. (2002). Parsing procedures and the question of full access in L2 phonology. *New Sounds 2000: The Fourth International Symposium on the Acquisition of Second Language Speech,* University of Amsterdam, 11-21.

Bertinetto, Pier Marco, 1980. The perception of stress by Italian speakers, *Journal of Phonetics* 8, 385-395.

Carlisle, R. S. (2001). Syllable structure universals and second language acquisition. *International Journal of English Studies* 1, 1, 1-10.

Cauldwell, R. (2002). Grasping the nettle: the importance of perception work in listening comprehension. http://www.developingteachers.com/articles_tchtraining/perception1_richard.htm

Corder, S. P. (1971). Language continua and the interlanguage hypothesis. In S. P. Corder and E. R. Roulet (eds.), *The notions of simplification, interlanguages, and pidgins, and their relation to second language pedagogy,* 11-17.

Cutler, A. (1993). Segmenting speech in different languages. *The psychologist* 6, 453-455.

Daniels, H. (1997). Psycholinguistic, psycho-affective and procedural factors in the acquisition of authentic L2 pronunciation. In A. McLean (ed.) (1997). *SIG Selections 1997 Special Interests in ELT,* 80-85.

Dauer, R. M. (1983). Stress timing and syllable timing reanalyzed. *Journal of Phonetics.* 11, 51-62.

Dickerson, W. B. (1989). Stress in the speech stream: the rhythm of spoken English. Urbana: University of Illinois Press.

Erickson, D., Akahane-Yamamda, R., Tajima, K. and Matsumoto, K. (1999). Syllable counting and mora units in speech perception. *ICPhS99,* San Francisco, 1479-1482.

ETS (2006). *Report on Test-Takers Worldwide 2005.* [Online]. http://members.ozemail.com.au/~hamalex/toeic/ReportonTestTakers2005.pdf.

Fledge, J. E. (1987) The critical period for learning to pronounce foreign languages? *Applied Linguistics* 8, 2, 162-177.

Ishiikawa, K. (2002). Syllabification of intervocalic consonants by English and Japanese speakers. *Language and Speech* 45, 4, 355-385.

Jenkins, J. (2000). *The phonology of English as an international language.* Oxford: Oxford University Press.

Jusczyk, P. (1997). *The discovery of spoken language.* Cambridge, Massachusetts: MIT Press.

Kanel, K. (1997). Teaching with music: a comparison of conventional listening exercises with pop song gap-fill exercises. *JALT Journal*, 19, 2, 217-231.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics 5, 253-263.*

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

Magen, H. (1998). The perception of foreign-accented speech. *Journal of Phonetics* 26, 381-400.

Mahler, J., Lambertz, G., Jusczyk, P. and Amiel-Tison, C. (1986). Discrimination de la langue matternelle par le nouveau-ne. *Contes Rends de l' Academie de sciences de Paris.* 303, 637-640.

Nazzi, T., Bertocini, J. and Mehler, J. (1988). Language discrimination by newborns: towards an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance* 24, 3, 756-766.

Ohata, K. (2004). Phonological differences between Japanese and English: several potentially problematic areas of pronunciation for Japanese ESL/EFL learners. *Asian EFL Journal.* 6, 4.

Port, R., de Jong, K., Kitahara, M., Collins, D., Leary, A., and Burleson, D. (2002) *Temporal attractors in rhythmic speech.* Manuscript.

Roach, P. J. (1982). On the distinction between 'stress-timed' and 'syllable-timed' languages. In D. Crystal (ed.), *Linguistic controversies,* 73-79.

Roach, P. J. (1994). Working on the model pronunciation. Unpublished paper, Singapore Tertiary English Teachers' Society, Singapore.

Shockey, L. (2003). *Sound patterns of spoken English.* Malden, MA and Oxford UK: Blackwell.

Swan, M. and Smith, B. (eds). (1987). *Learner English: A teacher's guide to interference and other problems. Cambridge:* Cambridge University Press.

Sweeting, A. (1998). Review of the ballad of Yoko and John. *Guardian.* (Sup.) January 7. p. 19.

Thomas, J. (1995). Meaning in interaction. London: Longman.

Vanderplank, R. (1993). 'Pacing' and 'spacing' as predictors of difficulty in speaking and understanding English. *ELT Journal*, 47, 117-125.

Appendix

Transcripts Part A

1. that
2. first
3. trouble
4. steam
5. woman
6. love
7. harm
8. mouth
9. sin
10. tax

Transcripts Part B

1. a. thing you put flowers in b. foundation
2. a. bug b. way to cook
3. a. place people live b. something bad
4. a. opposite of 'opens' b. act of going over a bridge c. something you wear
5. a. actor in 'The Godfather' b. fashion maker
6. a. physical activity b. to encourage c. small round mark
7. a. 8-0-8-0 b. 8-8
8. a. free from mistakes b. broken c. bring together

Transcripts Part C

1. Taxes are collected at the end of the month.
2. The man put the clean clothes on the bed.
3. The woman hurt her thumb on the hot fry pan.

Transcripts Part D

1. A: Could you tell me how to get to the station?
   B: Go left at the first light and take city bus number eight.
2. A: Which one should I get at the garden store?
   B: Either hoes potatoes.
3. A: Which player on the field is Johnson?
   B: His team uniform number is seven.