

# RELC Journal

<http://rel.sagepub.com/>

---

## Identifying Problematic Segmental Features to Acquire Comprehensible Pronunciation in EFL Settings: The Case of Japanese Learners of English

Kazuya Saito

*RELC Journal* 2011 42: 363

DOI: 10.1177/0033688211420275

The online version of this article can be found at:

<http://rel.sagepub.com/content/42/3/363>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *RELC Journal* can be found at:**

**Email Alerts:** <http://rel.sagepub.com/cgi/alerts>

**Subscriptions:** <http://rel.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

**Citations:** <http://rel.sagepub.com/content/42/3/363.refs.html>

>> [Version of Record](#) - Dec 14, 2011

[What is This?](#)

# Identifying Problematic Segmental Features to Acquire Comprehensible Pronunciation in EFL Settings: The Case of Japanese Learners of English

RELC Journal  
42(3) 363–378  
© The Author(s) 2011  
Reprints and permission: [sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)  
DOI: 10.1177/0033688211420275  
[rel.sagepub.com](http://rel.sagepub.com)  
The SAGE logo consists of the word 'SAGE' in a bold, sans-serif font, preceded by a circular emblem containing a stylized 'S'.

**Kazuya Saito**

Simon Fraser University, Canada

## Abstract

The present study examines how to identify problematic pronunciation features for particular EFL learners, namely native Japanese speakers (NJs) learning English, to acquire comprehensible pronunciation, and tests the appropriateness of the selection. The study comprises two phases. In the identification phase, eight English-specific segmentals, /æ, f, v, θ, ð, w, l, ɪ/, were selected as the most problematic for NJs by drawing on various cross-linguistic analyses (i.e. a remedial approach) as well as a survey in which the advice of 48 experienced NJ English teachers was examined (i.e. an expert judgment approach). In the experimental phase, the relative influence of these sounds on comprehensibility and accentedness was analyzed. Twenty NJ participants read two types of sentences: sentences containing eight English-specific segmentals and sentences without them. Four native English speakers (NEs) subsequently rated all speech stimuli on a rubric of accentedness and comprehensibility. Significant differences were found between NEs' ratings of the two types of sentences both in the domain of comprehensibility and accentedness. The results indicate that the eight segmentals determine NEs' speech perception to a great degree, which in turn provides some support for the validity of the identification procedure (i.e. the combination of the remedial and expert judgment approaches).

## Keywords

Intelligible pronunciation, comprehensibility, teacher questionnaire, segmental

---

## Corresponding author:

Kazuya Saito, Department of Linguistics, Simon Fraser University, 8888 University Drive, Burnaby, BC, Canada V5A 1S6.

Email: [kazuyas@sfu.ca](mailto:kazuyas@sfu.ca)

## Identifying Problematic Features to Acquire Comprehensible Pronunciation

Given that foreign accent is a normal characteristic of L2 speech (Piske et al., 2001), Derwing and Munro (2005: 385) claimed that L2 pronunciation needs to be assessed at two different levels: (a) accentness ('a listener's perception of how different a speaker's accent is from that of the L1 community') and (b) comprehensibility ('a listener's perception of how difficult it is to understand an utterance'). In particular, Derwing and Munro emphasized the importance of comprehensible pronunciation for the purpose of successful L2 communication (see also Levis, 2005).<sup>1</sup> Consequently, recent accent studies have begun to investigate what characterizes comprehensible pronunciation within L2 communication between native speakers of English (NEs) and non-native speakers of English (NNEs) as well as NNE-NNE (e.g. Derwing and Munro, 1997; Field, 2005; Hahn, 2004; Munro and Derwing, 1999; Munro, Derwing, and Morton, 2006). However, it remains controversial (a) which pronunciation features are particularly crucial for acquiring comprehensible pronunciation; and (b) what pronunciation problems teachers should teach as main concerns. Such information is crucial if practitioners set teaching priorities and design effective syllabi in L2 classrooms (Celce-Murcia, Brinton, and Goodwin, 2010; Levis, 2005).

If we accept that a primary goal in pronunciation teaching is to help L2 learners acquire comprehensible pronunciation, it is natural to ask which pronunciation features need to be prioritized on the basis of their potentially different contribution to comprehensibility. For this purpose, suprasegmental-based instruction (i.e. teaching intonation, speech rate, stress patterns, and phonotactic knowledge) has generally been employed in most pronunciation teaching classrooms in the past 25 years (Kang, 2010). Levis (2005) claimed that this suprasegmental-supremacy could reflect the fact that it is easier to teach suprasegmentals when not considering students' diverse linguistic backgrounds especially in ESL classrooms. While some research has identified crucial suprasegmental features in NE and NNE perception (e.g. Field, 2005 for lexical stress; Hahn, 2004 for sentence stress), previous L2 speech studies do note the role of segmentals. In NNE-NNE communication (i.e. English as an International Language [EIL]), Jenkins identified a range of affecting pronunciation features for communication breakdown (i.e. Lingua Franca Core [LFC]), most of which constitute segmentals (see Jenkins, 2000, 2002). Some L2 pronunciation research has reported that particular EFL learners such as Japanese learners of English (NJs) whose L1 phonetic system greatly differs from that of English tend to have salient L1-L2 transfer problems at a segmental level (e.g. Lambacher, 1999; Riney and Anderson-Hsieh, 1993). For example, Riney and his coworkers (Riney, Takada, and Ota, 2000; Riney and Takagi, 1999) offered detailed documentation of the link between NJs' accurate production of segmental sounds (a non-native contrast /i/-/I/, word-initial stops /p, t, k/) and foreign accent scores judged by NE listeners. Their results showed that the number of NJs' erroneous production negatively influenced foreign accent ratings (the more errors, the worse rating scores). Accordingly, they claimed that NJs need segmental-level pronunciation instruction in order to consciously reduce the level of their L1-L2 transfer problems and dilute their marked deviation from NEs' norms. However, as Riney et al. (2000: 733) acknowledged, 'the findings reported are based on a small sample and must be interpreted with caution'. Importantly, the results

of these previous studies are still at a micro level in terms of implications for L2 classrooms. For example, although English /ɪ/-/I/ contrast or word-initial stops /p/, t/, and /k/ should be taught to NJs, the question of other features remains. In sum, teachers and learners in EFL classrooms, such as English education system in Japan, have yet to receive a sufficient amount of information and practical advice from these lines of L2 speech research. Acknowledging the importance of suprasegmental features in comprehensible pronunciation especially for ESL learners as well as LFC features for EIL learners, the current study further examines which pronunciation features can be relatively crucial for NJs in EFL settings to acquire comprehensible pronunciation.

## Current Study

In order to make pedagogical implication for L2 classrooms, the present study first identified eight English-specific segmentals /æ, f, v, θ, ð, w, l, ɹ/ as problematic features for NJs based on cross-linguistic analyses as well as the expert judgments by 48 experienced EFL teachers (the *identification* phase). Subsequently, the impact of these segmentals on NE listeners' perception was assessed on a rubric of accentedness and comprehensibility (the *experimental* phase). The research questions to be concerned are as follows:

1. How can we identify problematic English segmentals for particular L2 learners such as NJs?
2. To what degree are these segmentals related to NE listener ratings of accent and comprehensibility?

## Identification Phase

As trustworthy ways to identify the relative difficulty, learnability, and teachability of target linguistic features for a particular group of L2 learners, Ellis (2006: 30) recommended two options: a *remedial* approach and an *expert judgment* approach. In the remedial approach, the target feature can be chosen 'based on previous empirical findings that have demonstrated the feature is problematic to learners.' In contrast, in the expert judgment approach, researchers elicit experienced L2 teachers' opinions to determine learners' problematicity. In fact, Robinson (1996) turned to the expert judgement of experienced ESL teachers for determining easy-hard grammar rule distinctions for NJs by asking them to rate grammar complexities on a 7-point scale. He pointed out several advantages of eliciting experts' advice:

1. It is empirical.
2. It is replicable.
3. It has high face validity.
4. It brings to light psycholinguistically appropriate targets because teachers usually take into account information-processing load for learners.

The present study's approach to problematicity was two-fold: First, cross-linguistic analyses were conducted on phonetic distance between Japanese and English systems, in conjunction with previous L2 speech studies, to identify the most problematic segmental

features for NJs (i.e. the remedial approach). Second, the advice of expert NJ English teachers was examined (i.e. the expert judgment approach). Finally, it was further investigated whether findings in the two approaches correspond to each other.

### *Cross-linguistic Analyses*

Although other English pronunciation models are available (e.g. British English, Australian English, New Zealand English), for the convenience of this study, cross-linguistic analyses were conducted on the phonetic difference between Japanese and General American English.<sup>2</sup> Vance (1987) noted that, compared to GA, Japanese has relatively fewer phonemes both in the domain of vowel sounds (GA: 16, Japanese: 5) and consonant sounds (GA: 24, Japanese: 14). It has been shown that NJs likely transfer their L1 phonological knowledge resulting in various pronunciation errors especially in English-specific segmentals (e.g. Lambacher, 1999; Riney and Anderson-Hsieh, 1993). Among several segmental problems indicated by previous L2 speech studies, eight English-specific segmentals were carefully selected as most problematic features for NJs.

**Vowels.** While GA has 16 vowel sounds including diphthongs, Japanese has only 5 vowel sounds in its vowel inventory. In addition to four diphthongs, English has five front vowel sounds /i, ɪ, e, ε, æ/, two central vowel sounds /ə, ʌ/ and five back vowel sounds /u, ʊ, o, ɔ, a/, while Japanese has only two front vowel sounds /i, e/, one central vowel sound /a/ and two back vowel sounds /u, o/ (Vance, 1987). Whereas NJs tend to assimilate most of these American vowels into the closest L1 counterparts (i.e. the equivalence classification: see Nishi, Strange, Akane–Yamada, Kubo, and Trent-Brown, 2008), a series of laboratory perception training studies (e.g. Lambacher et al., 2005) demonstrated that NJs tend to attain a great deal of improvement especially for perception and production of a low front vowel /æ/, arguably due to its acoustic dissimilarity from any Japanese vowel phones; thus, the present study highlights the low front vowel /æ/ in conjunction with its relative teachability.<sup>3</sup>

### *Consonants*

1. *Fricatives /f, v/ and /θ, ð/.* First of all, there are no labio-dental fricatives /f, v/, or interdental fricatives /θ, ð/ in Japanese, so those sounds should be taught explicitly to NJs. Lambacher (1999) stated that NJs likely mispronounce English /f, v/ as Japanese /ɸ, b/, and English /θ, ð/ as Japanese /s, z/ due to the influence of the Japanese phonetic system.
2. *Approximants /w, l, ɹ/.* According to Vance (1987), there are two approximants /w, j/ and one lateral apical postalveolar flap /l/ in Japanese as counterparts for English approximants /w, j, l, ɹ/. As for English /w/, Lambacher (1999) stated that NJs likely mispronounce English /w/ as Japanese /w/, so it is necessary to teach NJs to round their lips in order to pronounce English /w/ correctly. Next, in contrast with the two English liquids /w, j, l, ɹ/, there is only one, a lateral apical postalveolar flap /l/ in Japanese, which necessitates that NJs be taught explicitly how to pronounce English /l/ and English /ɹ/. It has to be noted that since American English was the target language system, the present study took into account /ɹ/ in the onset as well as coda positions.

In summary, for the purpose of the current study, eight *English-specific segmentals*, /æ, f, v, θ, ð, w, l, ɹ/, were hypothesized as important components for NJs' pronunciation problems. Next, the expert judgment was also adopted to investigate how experienced teachers identify problems for NJs and the extent to which these two approaches (i.e. the remedial and expert judgment approaches) match with each other.

## The Expert Judgment

### Participants

Forty-eight expert Japanese teachers of English participated in the present study. All of them belonged to a private language school located in Tokyo and they reported to have several years of teaching experiences (varying between two and ten years) with communicative language teaching methods (i.e. teaching conversational English) rather than grammar-translation methods (i.e. concentrating on reading and writing texts). Thus, it can be said that all participants were well qualified to judge pronunciation problems of NJs in communicative settings. Since the purpose of this questionnaire was to obtain teachers' own opinions without being biased by any prior information, only one open-ended question was asked: 'Which English segmental sounds do you think are problematic for NJs based on your teaching or learning experiences?'<sup>4</sup> Instead of limiting the number of answers, participants were asked to feel free to point out as many segmental problems as they felt the need to teach. The questionnaire also included one example answer (i.e. 'The non-native contrast /l/ and /r/: Most students tend to have tremendous difficulty in perceiving and producing these two sounds, resulting in many related spelling errors.') to familiarize the informants with the answering procedures. All of their responses were used for data analysis.

Importantly, since this English institute that all the participants belonged to does not usually provide teachers with any specific training for pronunciation teaching, the assumption was that they judged the pronunciation problems solely based on their teaching (and learning) experiences. Yet, given that some teachers might have had training outside of the institute (the current study could not control the amount of their experience for pronunciation teaching), it needs to be stressed here that the trustworthiness of teachers' answers in the current analysis might have been considered heterogeneous rather than homogeneous and the results and following discussion could inevitably be exploratory in nature. However, for future replication studies, I also need to note that it could be extremely difficult to control the teachers' experience in pronunciation teaching, because very few practitioners have had even opportunities to receive appropriate training for pronunciation teaching and their knowledge and attitude towards pronunciation teaching widely varies.

### Coding

All responses were coded by the author. Given that the questionnaire provided one example answer, most of the teachers' responses followed the format. Whereas some noted phonetic (or alphabetic) symbols (e.g. 'θ, ð': Many Japanese students substitute /s/ especially when they are engaged in free conversation' by Teacher 11; 'æ, ɹ': Most of my students use the Japanese /a/ without differentiating the vowel distinction

between /æ, ʌ/ and /a/’ by Teacher 20), others showed simply alphabetic letters for the problematic sounds without much description (e.g. ‘f, v, sh, y’ by Teacher 32). In the latter case, the author identified which phonetic symbols corresponded to their answers.

## Results

First, all of the responses by the 48 teachers were assessed by looking at whether they fell into any segmental categories in the English phonetic system (i.e. IPA) excluding answers that could be considered as suprasegmental errors (i.e. intonation, lexical stress, speech rate, and syllabification). Second, all answers were transformed into point scores in the following manner: If one participant identified one segmental problem, then that sound received one point. For example, if one participant reported three English sounds such as /æ, l, ɪ/ as problematic, one point was given to each segmental category respectively: each of /æ/, /l/ and /ɪ/ receives one point.

Forty-eight expert NJ teachers identified 20 problematic English segmentals for NJs: /θ, v, sɪ, j, æ, w, f, ɪ, l, ð, h, ʌ, n, ə, tʃɪ, ɪ, ʒ, dʒ/, diphthongs, and voiceless stops with 164 points in total. On average, one participant contributed 3.41 points ( $SD=1.95$ ). Details of the data are presented in Table 1. Some of the answers specific to English vowel diphthongs were grouped as one category (e.g. /aʊ, aɪ, oʊ, ɔɪ, eɪ/). Similarly, those regarding aspiration problems were grouped as voiceless stops (i.e. /p, t, k/).

**Table 1.** Problematic Segmentals for NJs based on 48 Expert NJ English Teachers

1. /θ/	21 points
2. /v/	21 points
3. /sɪ/	20 points
4. /j/	15 points
5. /æ/	13 points
6. /w/	12 points
7. /f/	12 points
8. /ɪ/	11 points
9. /l/	8 points
10. Diphthongs	8 points
11. /ð/	5 points
12. /h/	4 points
13. /ʌ/	3 points
14. /n/	3 points
15. Voiceless stops	2 points
16. /ə/	2 points
17. /tʃɪ/	1 point
18. /ɪ/	1 point
19. /ʒ/	1 point
20. /dʒ/	1 point

48 NJ teachers provided 164 points for 20 sounds in total.

Average no. of points generated per participant: 3.41 points ( $SD = 1.95$ )

While the remedial approach suggested eight English-specific segmentals /æ, f, v, θ, ð, w, l, ɹ/ for NJs in conjunction with research findings from cross-linguistic analysis studies detailed in the previous section, it is shown that experienced teachers' opinions also support the validity of these sounds by assigning relatively higher points to all of these eight sounds (e.g. /æ/ - 13 points, /f/ - 12 points, /v/ - 21 points, /θ/ - 21 points, /ð/ - 5 points, /w/ - 12 points, /l/ - 8 points, /ɹ/ - 11 points). Therefore, the two approaches (the remedial and expert judgment) generally concur that eight English-specific segmentals /æ, f, v, θ, ð, w, l, ɹ/ contribute to NJs' pronunciation problems to a great degree.

It has to be acknowledged, however, that some segmental problems that the experts assigned higher scores such as /sɹ/ (20 points) and /j/ (15 points) were not included for further discussion in the present study for the following reasons. First, with respect to /sɹ/, this problem could be considered at more than allophonic levels because its difficulty is highly related to Japanese phonological rule /s/ → [ʃ] / \_ɹ (see, for other phonological assimilation in Japanese fricatives, Tsujimura, 1996). Second, /j/ was also excluded because /j/ exists both in Japanese and English phonetic systems and this subtle acoustic difference within the same category has been rarely reported as a problem by previous L2 phonology literature. In addition, it might be extremely hard not only for learners, but also for teachers, to clearly discuss the phonetic difference between Japanese /j/ and English /j/ in a classroom. One of the reasons why many experts considered /j/ as problematic could be that the private language school implements a 5-minute pronunciation lesson as a warm-up for every class and the topic of the week was /j/ when the questionnaire was conducted. This may have affected their decision.

After problematic pronunciation features were identified (i.e. eight English-specific segmentals), the experimental phase was carried out to test the extent to which these eight problematic sounds affect NE listeners' perception.

## Experimental Phase

In order to measure the role of eight *English-specific segmentals* in NJs' speech production, the present study adopted a sentence reading task by asking NJs to read two types of sentences: *loaded* sentences and *non-loaded* sentences. Afterward, all speech stimuli were rated by four NE listeners on a rubric of accentedness and comprehensibility (e.g. Derwing and Munro, 1997; Munro and Derwing, 1999).

## Participants

Participants were 20 NJs who were studying abroad at university-level schools in upstate NY at the time of the project. According to the results of the personal interview, the demographic information of these participants was as follows: their age ( $M = 27.6$  years,  $SD = 5.4$ ), length of stay in the USA ( $M = 2.3$  months,  $SD = 2.7$ ), TOEFL CBT scores ( $M = 223.4$ ,  $SD = 15.6$ ), and gender (14 females, 6 males).

## Speech Samples

In the experimental phase, first, three *non-loaded* sentences were carefully composed in such a way that they included few of the eight English-specific sounds /æ, f, v, θ, ð, w, l, ɹ/: With a



total of 26 words, three *non-loaded* sentences contained only one word that has a problematic sound /ɪ/. The average percentage of loaded words per sentence is 3 % (see Appendix).

The purpose of having NJs read non-loaded sentences is to measure any kind of pronunciation errors (suprasegmental and segmental problems) that NJs likely have, except the eight English-specific sounds, /æ, f, v, θ, ð, w, l, ɪ/. Thus, non-loaded sentences played a *baseline* role.

Second, four *loaded-sentences* were carefully composed: They consisted of 50 words out of which 41 were loaded with the targeted phones (i.e. /æ, f, v, θ, ð, w, l, ɪ/). Thus, they are hypothesized to be difficult especially for NJs to read because of NJs' L1-L2 transfer problems. The average percentage of loaded words per sentence is 82% (see Appendix).<sup>5</sup> The contents of loaded sentences are summarized in Table 2.

**Table 2.** Contents of Loaded-sentences

Targeted phones	The total number of Loaded Phones	Examples
/æ/	6	married, happy, bad
/f, v/	6	office, food, of, visa
/θ, ð/	6	think, things, they, the
/w/	4	when, woman, with, workers
/l, ɪ/*	17	read, letter, recently, limited

\*/r/ sounds in all positions (word initial, medial and final) were counted.

### Subjective Rating Method

Following Derwing and Munro's (1997) methodologies, the present study used a subjective rating method in which NE listeners were asked to carefully rate only *speech properties* using a 9-point scale for each of two rubrics: accentedness (1. native-like – 9. heavily accented) and comprehensibility (1. no effort to understand – 9. very hard to understand). The assumption of the present study is that approaching two types of speech stimuli (loaded and non-loaded sentences) from two different points of view (accentedness and comprehensibility) will provide a clear picture of the effects of the eight English-specific segmentals on NE listeners' speech perception.

**Listener-Raters.** Four NE listeners were recruited at a university in upstate New York. All of them had grown up in the USA and reported having normal hearing. Since they were instructors of phonetics or ESL classes in the linguistics department, they also reported their familiarity with NNE speech; it is possible that the listener factors of these participants were homogeneous.

**Overall Procedure.** First, all 20 NJ participants read four loaded sentences as well as three non-loaded sentences presented in a randomized order. All of their speech stimuli were recorded by speech analysis software, *Praat* (downloadable at [www.praat.org](http://www.praat.org)), and 140 speech stimuli were created: 7 sentences (4 loaded and 3 non-loaded sentences)  $\times$  20 participants = 140 stimuli. Second, two native English speakers were hired (one female, one male) and asked to do the same task as NJs (2 NEs  $\times$  7 sentences = 14 stimuli). Their

recordings were added as controlled stimuli, which NE listeners were expected to rate '1' for both accentedness and comprehensibility (Derwing et al., 1998). Third, after all 154 speech stimuli (140 NJs' stimuli + 14 controlled stimuli = 154 stimuli) were randomized and put in one data CD, four NE listeners were asked to listen to the CD and rate each token both for accentedness and comprehensibility respectively. Four listeners took part in a joint training session: By listening to 10 speech stimuli randomly selected from the data pool, all listeners were asked to discuss together what points the stimuli deserved both for accentedness and comprehensibility, and check with each other's rating scale.

**Data Analysis.** The data were computed in the following manner. First, all four listeners' ratings were averaged for each token. For example, if one NJ participant received '4', '5', '5' and '6' for one of his/her sentence reading performance, he/she received '5' as a total score. Second, all individual scores were averaged across the four loaded sentences and three non-loaded sentences respectively, both in accentedness and comprehensibility, creating four contexts: (a) accentedness/loaded sentence, (b) accentedness/non-loaded sentence, (c) comprehensibility/loaded sentence, and (d) comprehensibility/non-loaded sentence. For example, if one NJ participant received '4.32', '3.53', '3.84' and '4.16' for four loaded sentences in comprehensibility, he/she received '4.21' as a total score in the context of comprehensibility/loaded sentence.

## Results

First, all four listeners successfully rated 14 controlled stimuli as '1' both on accentedness and comprehensibility without any exceptions. Second, inter-class correlation (between all four raters) was calculated for accentedness ( $r = .68$ ) and comprehensibility ( $r = .55$ ) by following the manner of Ebel (1951). Therefore, all results generated by the four listeners were statistically significant at  $p < .001$  level. The descriptive results of the rating scores were summarized according to (a) the Assessment factor (accentedness/comprehensibility) and (b) the Loadedness factor (loaded/non-loaded sentences) in Table 3.

To examine whether these two factors (Assessment, Loadedness) made significant effects on the NE rating patterns, all scores were submitted to a two-way ANOVA with two repeated measures on Assessment and Loadedness. An overall Assessment  $\times$  Loadedness interaction effect was statistically significant,  $F(1, 19) = 6.066$ ,  $p = .023$ . According to Bonferroni adjustment for multiple comparisons, the score difference between loaded and non-loaded sentences was significant in accentedness ( $p < .001$ ) as well as in comprehensibility ( $p < .001$ ), confirming that loaded sentences (including the eight segmentals /æ, f, v, θ, ð, w, l, ɪ/) exert a more negative influence on

**Table 3.** Results of Rating

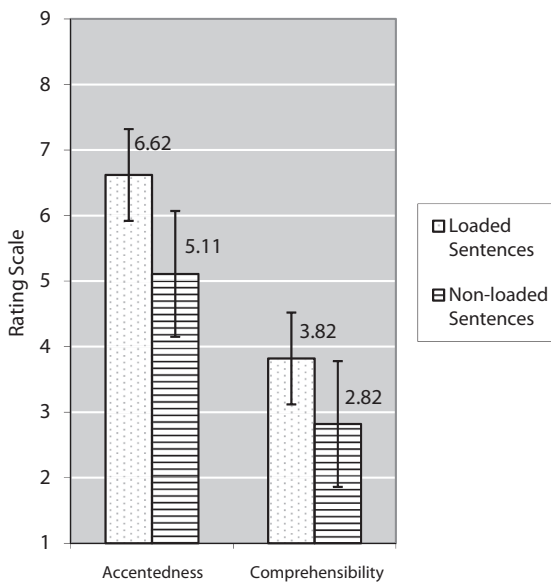
Types of domains	Types of sentences	M	n	SD
1. Accentedness	Loaded sentences	6.62	20	0.97
2. Accentedness	Non-loaded sentences	5.11	20	0.76
3. Comprehensibility	Loaded sentences	3.82	20	0.58
4. Comprehensibility	Non-loaded sentences	2.82	20	0.53

NE perception of accentedness and comprehensibility than non-loaded sentences (including few problematic segmentals) do. In addition, the results of multiple comparisons also revealed that NE listeners' accentedness ratings were higher (and thus harsher) than comprehensibility ratings both in the loaded sentences ( $p < .001$ ) and in the non-loaded sentences ( $p < .001$ ).

Given that the significant interaction effect suggests that the difference in rating scores between loaded and non-loaded sentences (i.e. the Loadedness factor) might vary for the two evaluation criteria (i.e. the Assessment factor), the effect size analyses were separately conducted on the contrasts between loaded and non-loaded sentences in accentedness as well as in comprehensibility. The results noted very similar effect sizes for both of the two evaluation domains ( $d = 1.73$  for accentedness,  $d = 1.80$  for comprehensibility). This indicates that NE listeners could be invariantly sensitive to the eight segmentals not only when they adopted the strict evaluation criteria (i.e. accentedness) but also when they employed the lenient rubric (i.e. comprehensibility). The results are graphically summarized in Figure 1.

## Discussion

In the present study, eight *English-specific segmentals*, /æ, f, v, θ, ð, w, l, ɹ/, were first identified as problematic pronunciation features for NJs drawing on expert judgements as well as cross-linguistic analyses (the identification phase). The effects of these sounds on accentedness and comprehensibility of L2 speech production was then tested through



**Figure 1.** Results of rating (9 = heavy accent/very hard to understand, 1 = no accent/no effort to understand)

subjective ratings of NE listeners (the experimental phase). Combined together, results of the present study provide both implications for future research on pronunciation teaching and practical recommendations for teachers in a classroom.

### *RQ1. How can we identify problematic pronunciation features?*

Eight English-specific segmentals identified in the identification phase not only have been described as problematic in light of cross-linguistic analysis studies (e.g. Lambacher, 1999; Riney and Anderson-Hsieh, 1993) but also received high scores from the expert judgement of experienced teachers in Japan (see Table 1). Although Levis stated, 'to large extent, pronunciation's importance has always been determined by ideology and intuition rather than research' (Levis, 2005: 369), the present study actually shed light on the credibility of experienced teachers' advice. If a certain number of expert judgements are carefully taken into account, these practitioners' opinions appear to echo SLA research findings and thus become reliable. On the contrary, the applicability of the remedial approach (i.e. finding problems by interpreting previous L2 speech studies as shown in the present study) to L2 classrooms is still ambiguous. Derwing and Munro (2005) pointed out the difficulty of employing this type of approach in L2 classroom settings as follows:

An extensive, growing literature on L2 speech has been published in journals that focus on speech production and perception...yet this work is rarely cited or interpreted in teacher-oriented publications...because it is inaccessible to those without specialized knowledge of phonetics. Moreover, some of the research may not be perceived as practical because it has been carried out under strict laboratory conditions, so that it is not immediately clear how the findings apply to the classroom (Derwing and Munro, 2005: 382).

Of course, this does not mean at all that SLA research findings can be safely ignored. Instead, the results of the current study suggest that, although it still remains unclear as to whether and to what degree one single teacher's subjective decision in regards to pronunciation teaching could be valid or intuitive (Levis, 2005), *collaborative talk* between practitioners might be a reliable index for determining teaching priorities especially when they are not familiar nor accessible to L2 speech research. Information obtained from such expert judgments tends to reflect the reality of classrooms, enabling practitioners ultimately to elaborate effective instructional L2 syllabi.

### *RQ2. How much do these features affect NE listeners' perception?*

The results of NE listeners' ratings revealed that NJs have more difficulty in pronouncing loaded sentences than non-loaded sentences not only in accentedness but also in comprehensibility. That is, L2 learners such as NJs likely have L1-L2 transfer problems at segmental levels, resulting in a negative influence on NE listeners' perception regardless of the evaluation criteria. These findings support the relative difficulty of the eight English-specific segmentals /æ, f, v, θ, ð, w, l, ɪ/ which were originally hypothesized to be problematic for NJs not only via the previous research findings but also by the

teacher questionnaire. Thus, one of the implications from the current study could be that the combination of remedial and expert judgment approaches is empirically reliable to identify problematic pronunciation features for NJs.

Furthermore, the present study illustrated that all NE listeners rated NJs' speech more strictly in the domain of accentedness than comprehensibility both in loaded and non-loaded sentences, showing that accentedness and comprehensibility are essentially two different evaluation variables. This result coincides with many other L2 speech studies (e.g. Derwing and Munro, 1997; Munro and Derwing, 1999). Although some L2 learners are highly motivated to achieve accent-free speech and the nativeness principle still affects teaching practice in L2 classrooms (Levis, 2005), as previous SLA studies have convincingly shown, attaining native-like pronunciation after early childhood is extremely difficult (e.g. Piske et al., 2001). Derwing and Munro (2005: 384) claimed that 'it may do more harm than good for teachers to lead learners to believe that they will eventually achieve native pronunciation or to encourage them to expend time and energy working toward a goal that they are unlikely to achieve'.

In summary, the present study suggests that both teachers and learners in classrooms should consistently weigh two crucial factors to attain successful L2 speech perception and production: the evaluation criteria and segmental difficulty factors. First, they should adopt the more realistic and appropriate evaluation method for adult L2 learners (i.e. comprehensibility) to assess students' improvements in L2 pronunciation depending on their learning goals. Next, teachers should also help L2 learners overcome segmental difficulties (i.e. L2 specific sounds) by effectively implementing segmental-based instruction.

## Conclusion and Future Direction

Although identifying and prioritizing problematic features for comprehensible pronunciation is such a complex phenomenon, the present study provides one possible framework (i.e. the combination of the remedial and expert judgment approaches) and tests its validity by conducting experimental phonetics experiments to measure the extent to which these features are problematic for NE listeners' perception. Truly, the results of the present study need to be interpreted with caution due to the obvious limitations (e.g. the questionnaire consisted of only one open-ended question; only segmentals were focused and the problematicity of these sounds was tested only through the sentence reading tasks). However, it has to be pointed out again that the purpose of the present study is to take a first step toward establishing trustworthy as well as feasible methodologies for both researchers and practitioners to identify and prioritize problematic pronunciation areas for particular EFL learners to acquire comprehensible pronunciation; it is impossible to ignore the promising results of the present study.

As suggestions for future studies, three more topics are further posed to refine and expand the possibility of the present study's framework. First, because the questionnaire with only one open-ended question did not allow teachers who likely received little training for pronunciation teaching to report many pronunciation rules, the contents of the questionnaire as well as the amount of teachers' awareness for pronunciation teaching need to be reconsidered. For example, it would be intriguing to ask teachers who have sufficient experience for pronunciation teaching to rate the problematicity of a number

of individual pronunciation problems already identified by previous SLA research (e.g. English-specific segments, syllable structures, intonation, word and sentence stress, speech rate, fluency) to determine the relative weights of these problems toward comprehensible pronunciation. However, it needs to be emphasized again that it could be difficult to find a good number of teachers with sufficient experience in pronunciation teaching due to the general lack of awareness towards pronunciation teaching in many L2 classrooms (Derwing and Munro, 2005).

Second, although the eight segmentals were identified as crucial factors for comprehensibility from the teachers' perspectives, such information about what is more important also needs to be elicited from the participants who are the insiders themselves because phonological awareness raising tasks tend to be effective especially when they follow English models which the learners themselves aspire to approximate (Zhang, 2005).

Finally, although the findings of the present study were limited to EFL settings, it would be intriguing to investigate how not only NE listeners but also NNE listeners react to the same speech samples of NJs (i.e. EIL contexts). Whereas some scholars claimed that NNE and NE perceptions are different from each other (e.g. Jenkins, 2000, 2002), recent L2 speech studies have empirically shown that both NE and NNE listeners likely agreed on which aspects of NNE talkers' pronunciation are the most difficult and easiest to understand (e.g. Bent and Bradlow, 2003; Field, 2005; Munro et al., 2006). Therefore, it seems reasonable to speculate that NNE listeners would judge accentedness and comprehensibility of NJs' segmental difficulties in ways similar to the NE listeners in the present study; empirical studies are called for to investigate these issues.

## Acknowledgments

Parts of this study were presented at The 30th Annual NY TESOL Applied Linguistics Conference in New York City, on March 7th. I am grateful to Roy Lyster, Kim van Poeteren, Pavel Trofimovich, Candace Farris, and two anonymous RELC Journal reviewers for their helpful comments on earlier versions of this article.

## Notes

- 1 According to Munro's (2008) comprehensive review, mutual intelligibility is a bidirectional phenomenon between a speaker and a receiver, (both of whom make conscious or unconscious efforts to attain successful L2 communication). While sufficient research attention has been given to *listeners' factors* (i.e. their familiarity to foreign-accented speech; e.g. Kennedy and Trofimovich, 2008), the current study exclusively focuses on investigating which *speech properties* (i.e. segmental and supra-segmental aspects of L2 sounds) characterize intelligible pronunciation in order to inform the practical relevance in the field of pronunciation teaching research.
- 2 The selection of GA as a model in the current study also reflects a tendency in Japanese EFL classrooms: In her quantitative and qualitative analyses of EFL textbooks approved by the Ministry of Education in Japan, Matsuda (2002) found their exclusive focus on GA in all linguistic domains: 'Main texts and sentences in the morphological and syntactic rules of the "standard" American English, the selection of vocabulary and its spelling were based on the American conversation, and pronunciation guides and tapes that accompanied the textbooks also represented the phonology of American English' (2002: 437).

- 3 Given that the major difference between GA and other dialects such as Received Pronunciation and Australian English mainly lie in vowel formation (Ladefoged, 2003; Jenkins, 2000), it is notable that a low front vowel /æ/ is common in various dialects of English.
- 4 One could argue that these wordings did not allow the participants to express the degree of problematicity. Yet, the purpose of the current study is to grasp a *consensus* of teachers' opinions about NJs' pronunciation problems rather than a *precise picture* of the ordered rankings of the problematic pronunciation features.
- 5 It has to be acknowledged that word positions were not taken into account in the present study. Although Bent, Bradlow, and Smith (2007) in their small-scale study found the relatively important role of segmentals at word initial positions in intelligibility perception, future studies still need to be conducted to pursue this topic with larger speech samples.

## References

- Bent T, Bradlow A (2003) The interlanguage intelligibility benefit. *Journal of Acoustical Society of America* 114: 1600-610.
- Bent T, Bradlow A, and Smith B (2007) Segmental error in different word positions and their effects on intelligibility of non-native speech. In: Bohn O, Munro M (eds) *Language Experience in Second Language Speech Learning: In Honour of James Emil Flege*. Amsterdam: John Benjamins, 331-47.
- Celce-Murcia M, Brinton D, and Goodwin J (1996) *Teaching Pronunciation: A Reference for Learners of English to Speakers of Other Languages*. Cambridge: Cambridge University Press.
- Derwing T, Munro M (1997) Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition* 2: 1-16.
- Derwing T, Munro M, and Wiebe G (1998) Evidence in favor of a broad framework for pronunciation instruction. *Language Learning* 48(3): 393-410.
- Derwing T, Munro M (2005) Second language accent and pronunciation teaching: a research-based approach. *TESOL Quarterly* 39(3): 379-97.
- Ebel R (1951) Estimating the reliability of ratings. *Psychometrika* 16: 407-24.
- Ellis R (2006) Researching the effects of form-focused instruction on L2 acquisition. *Applied Linguistics* 19: 18-41.
- Field J (2005) Intelligibility and the listener: the role of lexical stress. *TESOL Quarterly* 39(3): 399-423.
- Hahn L (2004) Primary stress and intelligibility: research to motivate the teaching of suprasegmentals. *TESOL Quarterly* 38(2): 201-23.
- Jenkins J (2000) *The Phonology of English as an International Language*. Oxford: Oxford University Press.
- Jenkins J (2002) A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics* 23: 83-103.
- Kang O (2010) Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System* 38: 301-15.
- Kennedy S, Trofimovich P (2008) Intelligibility, comprehensibility, and accentedness of L2 speech: the role of listener experience and semantic context. *The Canadian Modern Language Review* 64: 459-89.
- Ladefoged P (2003) *Phonetic Data Analysis: An Introduction to Instrumental Phonetic Fieldwork*. Oxford: Blackwell.
- Lambacher S (1999) A CALL tool for improving second language acquisition of English consonant sounds by Japanese learners. *Computer-Assisted Language Learning* 12(2): 137-56.

- Lambacher SG, Martens WL, Kakehi K, Marasinghe CA, and Molholt G (2005) The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics* 26: 227-47.
- Levis J (2005) Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly* 39(3): 367-77.
- Matsuda A (2002) International understanding through teaching World Englishes. *World Englishes* 21: 436-40.
- Munro M (2008) Foreign accent and speech intelligibility. In: Hansen Edwards J, Zampini M (eds) *Phonology and Second Language Acquisition*. Amsterdam: John Benjamins, 193-218.
- Munro M, Derwing T (1999) Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning* 49, (Supplement 1): 285-310.
- Munro M, Derwing T, Morton S (2006) The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition* 28: 111-31.
- Nishi K, Strange W, Akane-Yamada R, Kubo R, and Trent-Brown S (2008) Acoustic and perceptual similarity of Japanese and American English vowels. *Journal of the Acoustical Society of America* 124: 576-88.
- Piske T, MacKay I, and Flege J (2001) Factors affecting degree of foreign accents in an L2: a review. *Journal of Phonetics* 29: 191-215.
- Riney T, Anderson-Hsieh J (1993) Descriptions of Japanese pronunciation of English. *JALT Journal* 15: 21-36.
- Riney T, Takagi Y (1999) Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning* 49: 2, 275-302.
- Riney T, Takada M, Ota M (2000) Segmentals and global foreign accent: the Japanese flap in EFL. *TESOL Quarterly* 34(4): 711-37.
- Robinson P (1996) Learning simple and complex rules under implicit, incidental rule-search conditions, and instructed conditions. *Studies in Second Language Acquisition* 18: 27-67.
- Tsujimura N (1996) *An Introduction to Japanese Linguistics*. Oxford: Blackwell.
- Vance T (1987) *An Introduction to Japanese Phonology*. Albany, NY: State University of New York Press.
- Zhang LJ (2005) Awareness-raising in the TEFL phonology classroom: student voices and socio-cultural and psychological considerations. *ITL International Journal of Applied Linguistics* 145(1): 219-68.

## Appendix

### Non-loaded Sentences

1. Most good men come to dinner on time.  
/moust gud men kam tu dinər ɔn taɪm/ (1 out of 8 words are loaded)
2. My dog eats two times each day.  
/maɪ dɒg its tu taɪmz ɪtʃ deɪ/ (0 out of 7 words are loaded)
3. He takes a bus to come to my house each day.  
/hi teɪks ə bʌs tu kam tu maɪ haʊs ɪtʃ deɪ/ (0 out of 11 words are loaded)

### Loaded Sentences

1. When do you think they are going to read letters?  
/wɛn du ju θɪŋk ðei əɪ ˈgouɪŋ tə ɪd ˈletəɪz/? (6 out of 10 words are loaded)



2. I guess a married woman is usually happy with her office life.  
/aɪ ɡɛs ə 'mæɹɪd 'wʊmən ɪz 'ju:ʒuəli 'hæpi wɪð həɪ 'ɒfɪs laɪf/ (8 out of 12 words are loaded)
3. He has at least nine things to complete on campus because of his visa.  
/hi hæz ət li:st naɪn θɪŋs tə kəm'plɪt ən 'kæmpəs bɪ'kɔ:z əv hɪz 'vɪzə/ (7 out of 14 words are loaded)
4. Recently the amount of food is very limited and that is bad for workers.  
/'ri:ntʃli ðə ə'maʊnt ə(v) fʊd ɪz 'vɛ:ɹɪ 'lɪmɪtɪd ənd ðæt ɪz bæd fɔ: wɜ:kəɪz/ (10 out of 14 words are loaded)