

*Japanese Adults can Learn to Produce English /ɹ/ and /l/ Accurately**

JAMES EMIL FLEGE

University of Alabama at Birmingham

NAOYUKI TAKAGI

Shosen University, Tokyo

VIRGINIA MANN

University of California at Irvine

KEY WORDS

second language acquisition

speech production

Japanese

liquid consonants

speech style

acoustic analysis

ABSTRACT

As is well known, Japanese adults who have just begun to learn English often err in producing /ɹ/ and /l/ because their native language does not possess such liquid consonants. The aim of this study was to determine if Japanese adults eventually learn to produce /ɹ/ and /l/ accurately in words like *read* and *lead*. Liquids spoken by 12 native Japanese speakers who had lived in the United States for an average of two years were often misidentified by native English-speaking listeners. Their productions of /ɹ/ and /l/ also received much lower (and thus foreign-accented) ratings than did the native English speakers' liquids. On the other hand, liquids produced by native Japanese speakers who had lived in the United States for an average 21 years were identified correctly in forced-choice tests. This held true for liquids in words that had been read from a list as well as for words that had been spoken spontaneously. The ratings of liquids produced by 10 of the 12 experienced Japanese speakers fell within the range of ratings obtained for the 12 native English speakers. These findings challenge the widely accepted view that segmental production errors in a second language arise from the inevitable loss of ability to learn phonetic segments not found in the native language.

INTRODUCTION

Very few if any individuals who begin to learn a second language (L2) in adulthood manage to speak it without a detectable foreign accent (e.g., Oyama, 1976; Suter, 1976; Flege & Fletcher, 1992; Flege, Munro, & MacKay, 1995b). This has suggested to some that effective speech learning is limited to a critical period that ends near the beginning of

*This research was supported by NIH grant DC00257. The authors thank M. Munro, A. Schmidt, W. Strange, and Y. Tohkura for comments on an earlier version.

Address for correspondence: J.E. Flege, Department of Biocommunication, University of Alabama at Birmingham, VH 503, Birmingham, AL 35294 USA. E-mail: flegeje@biocom1.bioc.uab.edu

adolescence (e.g., Lenneberg, 1967; Scovel, 1969, 1988). However, it is uncertain at present whether foreign accents result from the deterioration, slowing, or complete loss of some basic speech learning mechanism(s), from inadequate phonetic input and/or motivation to speak the L2 like a native speaker, or from difficulty in preventing the L1 and L2 phonological systems from interacting with one another (see Flege, 1987a, 1988, 1995).

One question that is pertinent to these issues is whether adults can learn to produce *any* L2 vowel or consonant just like monolingual native speakers of the target L2. Flege (1987b) provided evidence that certain “new” vowels in an L2 can be mastered by adult learners, whereas other L2 vowels that are similar but not physically identical to vowels in the L1 can not be mastered. That study was limited to acoustic measurements, however, and so it is uncertain if the nonnative speakers’ productions matched the full range of properties that defined the L2 vowels of interest. In fact, evidence obtained in a recent study that made use of a perceptual evaluation technique suggested that even highly experienced adult learners of an L2 may persist in producing L2 vowels inaccurately, even vowels that are quite different from any L1 vowel (Munro, Flege, & MacKay, in press).

The present study examined the production of English /ɹ/ and /l/ by native speakers of Japanese. English /ɹ/ is a dorsal, sometimes retroflexed, approximant. The other English liquid, /l/, is a lateral continuant. Japanese /r/ appears to occupy a position in phonological space that is somewhere between English /l/, /ɹ/, and /d/ (and possibly /w/). Although the /r/ phoneme of Japanese is often referred to as a “liquid” and is usually represented with the phonetic symbol “r” its articulatory characteristics led Koutsoudas and Koutsoudas (1983) to represent it with the phonetic symbol “l”. Jones (1967) claimed that phonetically distinct variants of Japanese /r/ are produced “indiscriminately” as a sound resembling English /ɹ/, as a lingual flap (/ɾ/), as a “kind of” retroflex /d/, a “kind of” /l/, or something “intermediate” In a study by Sekiyama and Tohkura (1993), native English listeners identified the initial consonant in Japanese /ra/ syllables as /l/ (51% of judgments), /gr/ (17%), /dl/ (10%), /r/ (10%), “Spanish /r/” (10%), or /wl/ (2%). Japanese /r/ may be produced more consistently than these reports would suggest, however.¹ According to Vance (1987), Japanese /r/ is usually realized as an apico-alveolar tap /ɾ/, but may be palatalized when spoken in the context of /i/ and /y/. Vance reported that when the tip of the tongue (which is held lightly against the alveolar ridge for Japanese /r/) is released rapidly, native English-speaking listeners may hear Japanese /r/ as /d/.² When produced with emphasis, the Japanese /r/ may be heard as /l/ because the tongue tip constriction along the midline, when maintained, permits the lateral passage of airflow.

¹ Trubetzkoy (1939/1969) noted that “phonemic false evaluation” may arise through talker-listener mismatches, especially when realizations of a foreign phoneme straddle the boundary between two phonemes in a listener’s L1. For example, Hale (1885; cited by Buckingham & Yule, 1987) reported an experiment that involved two transcribers, himself and Alexander Melville Bell. The two individuals transcribed Mohawk words with /r/. In instances where one heard an /ɹ/, the other heard /l/. This suggested to Hale that seemingly free variation in productions of Mohawk /r/ was more likely due to the listeners than to the talkers.

² Vance (1987) reports that children learning Japanese as an L1 substitute /d/ for /r/.

The overall phonetic similarity of English /ɹ/ and /l/ to the Japanese /r/ is attested by loanword phonology data. Word-initial singleton tokens of /ɹ/ and /l/ are rendered as Japanese /r/ in English words that have been borrowed into Japanese (Lovins, 1976; see also Takagi, 1993). However, the nature of the perceived relation between the two English liquids and Japanese /r/ is controversial. Takagi (1993) found that inexperienced native Japanese (NJ) speakers of English identified word-initial English /ɹ/ and /l/ tokens in terms of Japanese /r/, although English /ɹ/ tokens were judged to be more distant from Japanese /r/ tokens than were English /l/ tokens. Best and Strange (1992) suggested that both English /ɹ/ and /l/ may be identified by NJ listeners as poor exemplars of the Japanese /w/ (or possibly /r/) category. Sekiyama and Tohkura (1993) reported that English /ɹa/ syllables were heard by NJ listeners as English /ɹ/ (67%), Japanese /r/ (13%), /w/ (13%), and /g/ (7%). They concluded that Japanese "does not have a consonant that resembles" English /ɹ/ (1993, p. 443). Similarly, Henly and Sheldon (1986, p. 517) stated that neither English /ɹ/ nor /l/ can be compared directly to Japanese /r/ in articulatory terms and that, from an acoustic standpoint, NJ speakers of English have no "perceptual representation" that can be used to identify English /ɹ/ or /l/.

Individuals who learn a second language often identify phones in an L2 with phonic elements (position-sensitive allophones, or perhaps phonemes) of the L1. This process, called interlingual identification, occurs even when the L2 and L1 phones differ acoustically and articulatorily, and when the difference between them can be detected auditorily. The perceived relation of English /ɹ/ or /l/ to consonants in the Japanese phonetic inventory may vary as a function of the age at which English is learned. Several reports suggest that Japanese speakers who were first exposed to native-produced English in early childhood identified and discriminated English /ɹ/ and /l/ more accurately than did Japanese speakers who were first exposed to English in adulthood (e.g., Shimizu & Dantsuji, 1983; Yamada & Tohkura, 1992; Nakauchi, 1993; Goto, 1971; Miyawaki, Strange, Verbrugge, Liberman, Jenkins & Fujimura, 1975; MacKain, Best, & Strange, 1981; Logan, Lively, & Pisoni, 1991; Takagi, 1993). Perhaps NJ adults are more likely than NJ children to identify English /ɹ/ and /l/ in terms of Japanese /r/.³

Interlingual identification at a *perceptual level* is said to trigger the replacement of L2 phones by L1 phones in *speech production* (Lehiste, 1988; Flege, 1988). It is uncertain at present whether an L2 phone, once identified with a phone in the L1, may later cease to be so identified. Given the apparent perceptual similarity of Japanese /r/ and the two English liquids, one might expect NJ speakers to use Japanese /r/ when producing English words that contain /ɹ/ and /l/. In fact, previous production studies have shown that adult NJ speakers often produce English /ɹ/ and /l/ inaccurately in word-initial position (Goto, 1971; Dickerson, 1974; Cochrane, 1977; Mochizuki, 1981; Sheldon & Strange, 1982). Their intended singleton /ɹ/s may be heard as /l/ by native English-speaking

³ NJ adults' errors in identifying English liquids are bi-directional: English /ɹ/ tokens are identified as /l/, and vice versa. There is, however, a great deal of variability in NJ adults' perception of English /ɹ/ and /l/. Also, their perception of singleton liquids is generally more accurate in word-final than word-initial position (e.g., Mochizuki, 1981; Sheldon & Strange, 1982), perhaps because word-final but not word-initial tokens of English /ɹ/ and /l/ are classified differently by native speakers of Japanese (see Takagi, 1993).

listeners, and vice-versa. However, these previous studies have been characterized by large inter-speaker differences. The bases for these differences are unknown, but may have included differing amounts of English-language experience, training in English, or aptitude for speech learning. Given that most previous work has examined the production of liquids in isolated English words that were read from a list, it is also uncertain if the results obtained in previous studies would generalize to more spontaneous forms of speech.

The present study examined the production of English liquids by native Japanese speakers who had studied English at school in Japan, but who were not massively exposed to English until they arrived as adults in the United States (US). Two groups of Japanese adults differing in length of residence in the US (2 vs. 21 years) produced minimally paired English words beginning in /ɪ/ and /l/ in three speaking conditions. Speakers in the first group resembled the speakers who have been examined in many previous studies. However, no previous study has examined Japanese speakers as experienced in English as those in the latter group. In two experiments, liquids spoken by the Japanese speakers and a group of native English speakers were presented to listeners for identification as /ɪ/ or /l/. In two other experiments, /ɪ/ and /l/ tokens were presented in separate blocks to listeners, who rated them for degree of foreign accent. To conclude that the Japanese speakers had mastered English /ɪ/ and /l/, it would be necessary to show that their liquids were identified correctly as often as native English speakers' liquids, and received ratings that were comparable to those obtained for liquids spoken by NE speakers. As expected, we found that some Japanese speakers' liquids received lower ratings than NE speakers' liquids, suggesting that their liquids were foreign-accented. Therefore, as a final step, acoustic analyses were undertaken to determine the basis of perceived foreign accent in liquids spoken by native Japanese speakers.

GENERAL METHODS

Speakers

Characteristics of the three groups of speakers who produced the liquid consonants examined in this study are summarized in Table 1. The four males and eight females in each group were all tested on the campus of the University of California-Irvine. The Japanese speakers in two groups began to study English in school at the age of 12 years but were first massively exposed to English when they came to the US. The experienced Japanese (EJ) speakers of English had all lived in the US for at least 12 years when they were tested ($M = 20.8$ years), whereas the relatively inexperienced Japanese (IJ) speakers had lived in the US for less than three years ($M = 1.6$). Three EJ and two IJ speakers reported having taken an intensive English course when they arrived in the US. Four of the EJ speakers, but none of the IJ speakers, were married to a native speaker of English. The IJ speakers reported using English less often on a daily basis than did the EJ speakers.

Speech materials

The native Japanese and English speakers were tested individually in a quiet room. Each took a multiple-choice vocabulary test after responding to a language background

TABLE 1

Characteristics of the native English (NE), experienced Japanese (EJ), and inexperienced Japanese (IJ) speakers who participated in this study

	Group					
	NE		EJ		IJ	
	<i>M</i> (<i>SD</i>)	<i>Range</i>	<i>M</i> (<i>SD</i>)	<i>Range</i>	<i>M</i> (<i>SD</i>)	<i>Range</i>
Testing age ^a	36.1 (6.6)	29.0 – 47.0	43.7 (1.9)	40.0 – 47.0	35.0 (4.7)	29.0 – 44.0
Arrival age ^c	—	—	23.0 (4.8)	18.0 – 31.0	33.5 (4.6)	28.0 – 42.0
US residence ^a	—	—	20.8 (4.7)	12.4 – 28.9	1.6 (0.8)	0.7 – 3.0
Ave. use of English ^b	—	—	5.7 (1.0)	4.3 – 7.0	2.9 (0.8)	1.3 – 3.7

^a In years

^b From estimates on a 1 (never) to 7 (frequently) scale for use at home, at work, and in social settings

questionnaire in their native language. In addition to providing data of interest, the vocabulary test also introduced the English words to be produced in three subsequent speaking tasks. The English words formed the 19 minimal pairs shown in Table 2. Also included on the vocabulary test were four nonwords (*ruck, rine, leck, lun*) paired with real English words (*luck, line, wreck, run*). The native English and Japanese speakers alike were advised that not all items on the vocabulary test actually occurred in English.

A native English speaker's production of the 46 consonant-vowel-consonant (CVC) words that appeared on the vocabulary test were digitized, normalized for peak amplitude, and then tape recorded. These CVCs were later presented aurally one at a time in the same random order in which they appeared, in written form, on the vocabulary test. The experimenter (NT) paused the tape recorder after each CVC was presented. Three possible definitions were offered to the subjects for each word. Of these, one was the correct answer, and one defined a word that was minimally paired to the test word. The two other response alternatives offered for each word were "not sure" and "do not know". If one of the three definitions was selected, the test word was rated for subjective familiarity on a scale ranging from 1 ("never heard/used") to 7 ("very often heard/used"). Words for which a definition was not attempted were later assigned a familiarity rating of zero.

TABLE 2

The minimally paired English words examined in this study

<i>Vowel Context</i>	<i>Vowel Height</i>	<i>Minimal Pairs</i>		
/i/	high	reef-leaf	reek-leak,	read-lead
/i/	high	rim-limb	rip-lip	
/u/	high	root-loot	room-loom	
/ʊ/	high	rook-look		
/e'/	mid	rake-lake	rate-late	ray-lay
/o/	mid	road-load	row-low	
/ɛ/	mid	red-led		
/æ/	low	rack-lack		
/ɑ/	low	rock-lock	rot-lot	
/ɑ'/	low/mid	ride-lied	right-light	

Recordings

The native Japanese and English speakers produced the English words beginning in /ɪ/ and /l/ in what will be designated the “definition” “reading”, and “spontaneous speech” tasks. The CVC words were spoken in the same pseudo-random order in all three tasks, which always occurred in the order just given. In the definition task, the experimenter said aloud an English phrase to elicit words from the NE speakers (e.g., “What we get from the sun” to elicit the word *light*). To elicit productions by native Japanese speakers in this task, the experimenter said aloud a Japanese translation equivalent for each English word to be spoken (e.g., “hikari” for *light*). The identity of the words to be spoken was usually obvious because they had all appeared previously on the vocabulary test. The CVC words elicited in the definition task were said twice, first in isolation and then at the end of a carrier phrase (*The next word is* ___). The utterance-final tokens were analyzed because the first, isolated tokens were sometimes produced with a rising intonation contour.

In the reading task, the CVC words were read in isolation from a written list. In the spontaneous speaking task, the native English and Japanese speakers attempted to produce each word in a phrase or sentence of their own choosing. The list of words used in the reading task was used to cue production. In those few instances in which a meaningful utterance containing a particular word could not be produced, the word was said at the end of a carrier phrase instead. These productions were not examined, however.

EXPERIMENT 1

English words beginning in /ɹ/ and /l/ from the word-list reading task were presented to native English-speaking listeners. The listeners identified the initial consonant as /ɹ/ or /l/, and also indicated degree of confidence in their forced-choice judgments of English liquids spoken by native and Japanese speakers of English.

Method

The 1,368 English words from the reading task (38 words × 36 speakers) were digitized at 10.0 kHz with 16-bit resolution, then normalized for peak intensity. These words were edited to prevent word-final consonants from influencing listeners' judgments of the word-initial consonants of interest here, and to reduce the possibility that variations in the following vowel might have such an effect. An oscillographic and spectrographic representation of each word was displayed using a Kay Computerized Speech Lab (CSL). A cursor was placed at the onset of the signal, then moved toward the vowel midpoint in 20-ms steps until the quality of the vowel (e.g., /i/ in *read* tokens) could be heard. All portions of the signal that occurred after this auditorily defined segmentation point were then removed digitally, yielding consonant-vowel (CV) stimuli.

The CV stimuli were presented binaurally via headphones to adult monolingual native speakers of English (two males, eight females) with a mean age of 24 years (range: 22 to 31 years). The listeners, all graduate students at the University of Alabama at Birmingham, came from many locations in the US (Alabama, Georgia, Louisiana, Florida, Massachusetts, New York, Iowa, New Jersey). All were required to pass a pure-tone hearing screening at octave frequencies between 0.5 and 4.0 kHz before participating.

The CV stimuli were low-pass filtered (4.8 kHz), then presented binaurally at a comfortable level to the listeners in a sound booth. The listeners typed one of six digits on a keyboard to indicate their judgment of each syllable-initial liquid. The response alternatives offered to them were (1) definitely L, (2) probably L, (3) possibly L, (4) possibly R, (5) probably R, or (6) definitely R. The listeners were required to respond to each stimulus, and were told to guess if unsure. A 1.0-sec interval occurred between each response and presentation of the next stimulus. The stimuli were randomly presented in four counterbalanced sets which each contained words with high, mid, and low vowels (Set 1: *rake-lake, reef-leaf, ride-lied, rook-look, rack-lack*; Set 2: *right-light, room-loom, rim-limb*; Set 3: *ray-lay, rip-lip, rock-lock, root-loot, road-load*; Set 4: *red-led, read-lead, rot-lot, row-low*).⁴

⁴ The quality of some vowels varied as a result of the editing. For example, the vowel in tokens of *rake* and *lake* produced by the native English speakers, all originally [e']-quality vowels, sometimes sounded like [e] after the editing, and sometimes like [ɛ]. This, together with the mixing of vowels within a single set (block) was expected to desensitize listeners to possible errors in vowel quality by the native Japanese speakers.

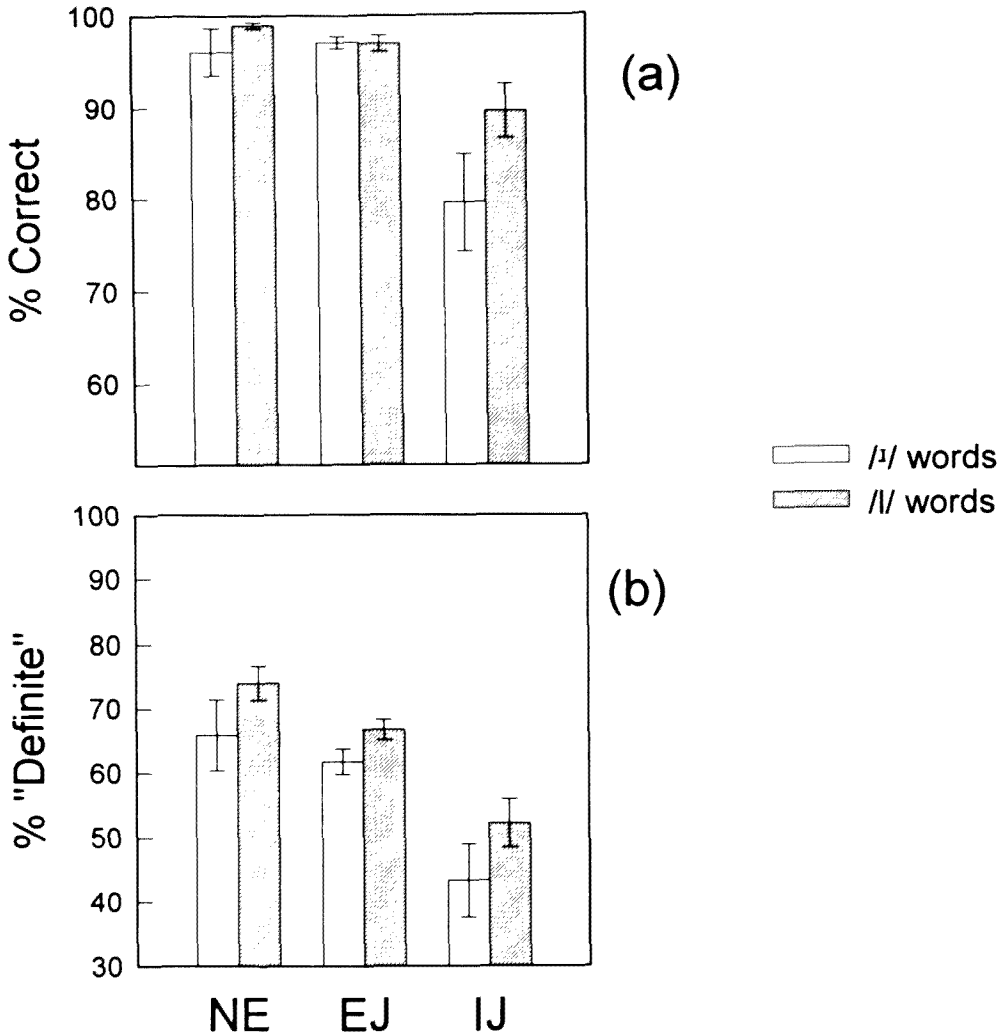


Fig. 1

(a) Mean percent correct identification scores obtained in Experiment 1 for /ɪ/ and /I/ tokens spoken by native English (NE), experienced Japanese (EJ), and inexperienced (IJ) Japanese speakers.

(b) The percentage of tokens identified as "definitely" being a realization of the intended category. The error bars in this and subsequent figures bracket ± 1.0 SE.

Results

Intelligibility. The mean percent correct identification scores shown for /ɪ/ and /I/ in Figure 1(a) have been collapsed over the three confidence levels. Liquids spoken by the NE speakers and the experienced Japanese (EJ) speakers of English were nearly always identified correctly. However, liquids spoken by the inexperienced Japanese (IJ) speakers were identified correctly less often (/ɪ/-80%, /I/-90%). Differences between the three groups could not be tested for statistical reliability because scores obtained for the NE and EJ speakers were at ceiling. However, the data suggest that the EJ speakers, but not the IJ speakers, may have produced English /ɪ/ and /I/ as accurately as did the NE speakers.

An ANOVA was carried out to examine the percentage of /ɪ/ and /l/ tokens that were judged as “definitely” being a realization of the intended category. These scores, which were not at ceiling, might be thought to reflect the subset of /ɪ/ and /l/ tokens that conformed closely to English phonetic norms. The /ɪ/ and /l/ scores obtained for the 36 speakers were each based on a maximum of 190 responses (19 words × 10 listeners). As shown in Figure 1(b), higher scores were obtained for the NE and EJ speakers than for the IJ speakers.

After arcsine transformation, the “definitely correct” scores were submitted to a (3) Group × (2) Consonant ANOVA with repeated measures on the Consonant factor. The scores were significantly higher for /l/ than /ɪ/, $F(1,33) = 9.23, p < .01$. The Group factor was also significant, $F(2,33) = 12.2, p < .01$, but not the 2-way interaction, $F(2,33) = 0.28, p > .10$. A Tukey’s HSD test examining the average scores obtained for each speaker’s productions of /ɪ/ and /l/ revealed that significantly more liquids spoken by the NE and EJ speakers than by the IJ speakers were judged to be definitely a realization of the intended category ($p < .05$). The NE and EJ speakers did not differ significantly, however. Thus, this analysis suggests that the EJ speakers produced word-initial singleton tokens of English /ɪ/ and /l/ as accurately as did the NE speakers.

In the analysis just presented, /ɪ/ and /l/ scores were computed for the 12 speakers in each group by averaging over the responses given by ten listeners. Listener-based scores were also computed. These scores were obtained for each of the ten listeners by averaging over responses obtained for the 12 speakers in each group. The ANOVA carried out to examine the 60 listener-based scores (2 liquid consonants × 3 groups × 10 listeners) yielded a significant Consonant main effect, $F(1,27) = 18.1, p < .01$, but not a significant effect of Group, $F(2,27) = 1.9, p > .10$, or a significant 2-way interaction. It appears that the listeners were not sufficiently consistent amongst themselves to support the conclusion that the NE speakers produced /ɪ/ and /l/ more accurately than even the IJ speakers. Perhaps the listeners used different criteria for what constituted “definite” instances of English /ɪ/ and /l/.

One other analysis was carried out to examine the “definitely correct” scores. This analysis examined scores computed for each of the 38 words, this time averaging over both speakers within groups and the ten listeners. The ANOVA examining the 114 word-based scores (3 groups × 38 words) yielded significant main effects of Group, $F(2,54) = 88.9, p < .01$, and Consonant, $F(1, 54) = 48.7, p < .01$, but not a significant 2-way interaction. A Tukey’s test examining average word-based scores for the two liquids revealed that the NE speakers’ liquids received significantly higher scores than the EJ speakers’ liquids, whose scores were higher than the IJ speakers’ ($p < .05$). This suggests that the EJ speakers did *not* produce English liquids as accurately as the NE speakers, or at least were not as consistent across lexical items as the NE speakers in producing liquids that conformed fully to English phonetic norms.

Inter-word variability. The IJ speakers were far more variable across lexical items in producing English liquids than were the NE and EJ speakers, whose liquids were nearly always identified correctly. The IJ speakers’ /ɪ/s were identified correctly at rates that ranged from 62% to 98% correct, depending on the word in which it occurred. The /l/s spoken by the IJ speakers were identified correctly at rates ranging from 77% to 99%. Analyses were undertaken in an attempt to account for this inter-word variability.

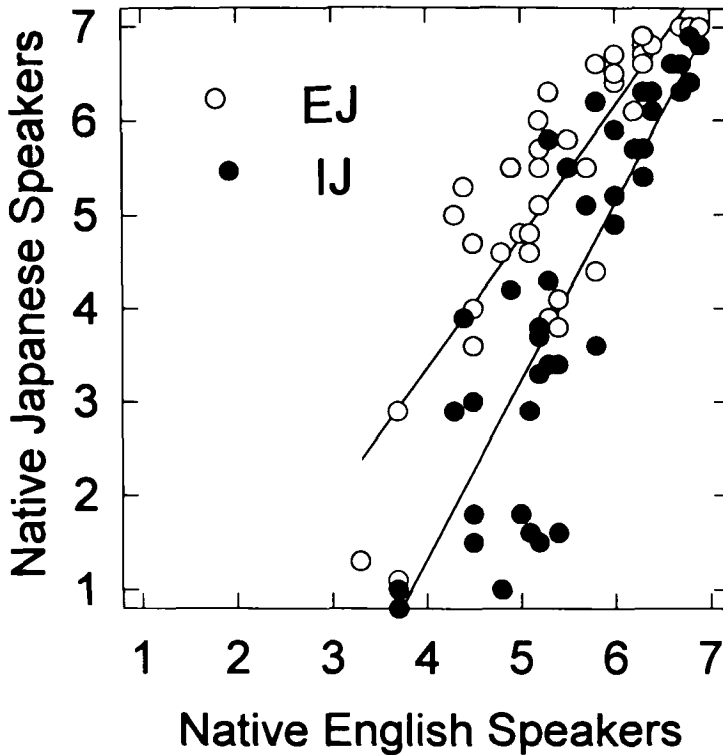


Fig. 2

Mean subjective familiarity ratings obtained for 38 English words beginning in /ɪ/ or /I/. The mean ratings obtained from 12 experienced Japanese (EJ) and 12 inexperienced Japanese (IJ) speakers have been regressed against the mean ratings obtained from 12 native English (NE) speakers.

One question of interest was whether the IJ speakers of English produced English liquids more accurately in familiar words than in words that were relatively unfamiliar to them (see Yamada, Tohkura, and Kobayashi, 1994). If such an effect were noted, it might be taken to mean that segmental phonetic representations are not freely commutable among items in the developing L2 lexicon.

As mentioned earlier, the English words examined in this study were rated for subjective familiarity before they were produced in three speaking tasks. As expected, the NE and EJ speakers judged the English words to be more familiar than did the IJ speakers. For the NE speakers, the mean familiarity ratings ranged from 3.3, for the word *rook*, to 6.9, for the word *look*. A similar range was observed for the EJ speakers. A somewhat wider range was observed for the IJ speakers because they gave average ratings of less than 2.0 to eight words (*viz.*, *reek*, *rake*, *rot*, *reef*, *wreck*, *limb*, *loot*, *loom*). Figure 2 shows that a correlation existed between the subjective familiarity ratings obtained from the NE speakers and the ratings obtained for both the IJ speakers ($r = .87$) and the EJ speakers ($r = .86$; $p < .01$).

The relation between the IJ speakers' subjective familiarity ratings of English words (38 words \times 12 speakers = 456 ratings) and the percent correct identification scores

obtained for their production of the liquids in those words was assessed. A significant correlation was not found to exist between the two variables ($r = .09$). The subjective familiarity ratings and weighted intelligibility scores to be discussed below also failed to show a significant correlation ($r = .11$). If one accepts that the NE listeners' identifications reflected how accurately liquids were produced, then one might conclude that the IJ speakers produced English liquids equally well in familiar and unfamiliar English words, at least in a reading task.

Flege and Munro (1994) found that native speakers of Spanish produced English stop consonants more accurately in English words without a cognate in Spanish than in English words that did have a Spanish cognate. Accordingly, a second analysis focused on the possible effect of cognate status. The IJ speakers were queried by mail several months after they were recorded concerning their perception of the relation that existed, if any, between the English CVC words they had produced earlier and words found in the Japanese lexicon. They were asked to rate the 38 English words on a scale ranging from "definitely related to a word in Japanese" (1) to "not related to a word in Japanese" (4). Nine or more IJ speakers thought that a relation existed between English *road, rock, rate, right, room, low, light, lip* and *lock* and a word found in Japanese. None of the IJ speakers judged such a relation to exist for 23 other English words (e.g., *red, loom*), however.

For each of the IJ speakers, percent correct identification scores were computed for subsets of English words beginning in /ɹ/ and /l/ that might plausibly be designated "cognate" words, as well as subsets of words that were surely "non-cognates" (i.e., unrelated to any Japanese word). The 2-way ANOVA examining the four scores obtained for each speaker yielded a nonsignificant main effect of Cognate Status, $F(1,11) < 1$, and a nonsignificant Cognate Status \times Consonant interaction, $F(1,11) < 1$.

Finally, we carried out an analysis to determine if vowel context might account for the inter-word variability. Mean percent correct identification scores were computed for /ɹ/s and /l/s produced in the context of high vowels (*reek, root, rook; leek, loot, look*) and low vowels (*rot, rack, rock; lot, lack, lock*). Four mean scores obtained for each of the 36 speakers were submitted to a (3) Group \times (2) Consonant \times (2) Vowel Height ANOVA, with repeated measures on the last two factors. The Consonant factor was significant, $F(1,33) = 8.14$, $p < .01$, but it did not interact significantly with the other two factors. The Vowel Height main effect was nonsignificant, $F(1,33) = 0.58$, $p > .10$, but did interact significantly with the Group factor, $F(2,33) = 3.82$, $p < .05$. Liquids were identified correctly somewhat more often when spoken by NE and EJ speakers in the context of low than high vowels (99% vs. 97%; 99% vs. 95%) whereas the opposite held true for the IJ speakers (79% vs. 83%). The trend noted here for the EJ but not the IJ speakers is similar to one reported for inexperienced Japanese speakers of English by Dickerson (1974).

Mean percent correct identification scores were also computed for /ɹ/ and /l/ tokens produced in the context of front vowels (*reek, rim, rack, rate; leek, limb, lack, late*) and back vowels (*root, rook, rock, road; loot, look, lock, load*). These scores were examined in a (3) Group \times (2) Vowel \times (2) Liquid Consonant ANOVA. Neither the Vowel main effect, $F(1,33) = 0.27$, nor the Group \times Vowel interaction, $F(2,33) = 0.05$, $p > .10$, were significant. A significant Consonant \times Vowel interaction was obtained, $F(1,33) = 13.03$, $p < .01$, apparently because the percent correct scores for /ɹ/ were slightly higher in the context of front than back vowels (92% vs. 88%) whereas the opposite held true for /l/

(93% vs. 96%). The lack of a significant 3-way interaction, $F(2,33) = 2.39, p > .10$, suggests that the effect of vowel context was much the same for all three groups of speakers.

Weighted intelligibility. Given some uncertainty as to the listeners' use of the "definite" label (see above), we also computed weighted intelligibility scores. These scores, which ranged from 0 to 10, were based on all confidence ratings, not just the "definitely correct" ratings. A score for each word beginning in /ɹ/ was computed by adding one point for each "definitely R" response given by any of the ten listeners, 0.8 for each "probably R" response, 0.6 for each "possibly R" response, 0.4 for each "possibly L" response, and 0.2 for each "probably L" response. A similar procedure was used to compute scores for each word beginning in /l/. Two scores were then derived for each speaker by averaging the scores obtained for the 19 words each beginning in /ɹ/ and /l/.

The weighted intelligibility scores obtained for the NE speakers (/ɹ/-8.9, /l/-9.3) and EJ speakers (/ɹ/-8.8, /l/-9.0) were somewhat higher than those obtained for the IJ speakers (/ɹ/-7.4, /l/-8.2). An ANOVA examining these speaker-based scores yielded significant main effects of Group, $F(2,33) = 11.8, p < .01$, and Consonant, $F(1,33) = 12.1, p < .01$, but not a significant 2-way interaction, $F(2,33) = 1.3, p > .10$. When the /ɹ/ and /l/ scores obtained for each speaker were averaged, a Tukey's HSD test indicated that the NE and EJ speakers' scores were significantly higher than the IJ speakers' scores ($p < .01$). The scores obtained for the NE and EJ speakers did not differ significantly, however ($p > .10$). The same results were obtained when 60 listener-based scores (3 Groups \times 2 Consonants \times 10 listeners) were submitted to an ANOVA.⁵

Discussion

Productions of /ɹ/ and /l/ by native English (NE) and experienced Japanese (EJ) speakers of English were nearly always identified correctly, whereas liquids spoken by relatively *inexperienced* Japanese (IJ) speakers were frequently misidentified. When intelligibility scores were weighted according to listeners' confidence in their forced-choice judgments, significantly higher scores were obtained for the NE and EJ speakers than for the IJ speakers.

As expected from previous research examining relatively inexperienced Japanese subjects, the IJ speakers' errors in producing /ɹ/ and /l/ were bi-directional. Their /ɹ/s were sometimes heard as /l/, and their intended /l/s were sometimes misidentified as /ɹ/. The IJ speakers' success in producing /ɹ/ and /l/ varied across the lexical items examined in this study. The inter-word variability could not be readily accounted for by variation in the subjective familiarity of words for the IJ speakers, the words' perceived status vis-a-vis words found in Japanese, or the immediately following vowel context.

One possible explanation for the inter-word variability is that the IJ speakers substituted Japanese /r/ for English /ɹ/ and /l/. One of us who is a native speaker of

⁵ The ANOVA examining listener-based weighted intelligibility scores yielded significant main effects of Group, $F(2,27) = 6.8, p < .01$, and Consonant, $F(1,27) = 172.0, p < .01$, but not a significant two-way interaction, $F(2,27) = 3.0, p > .05$. Again, a Tukey's HSD test indicated that the NE and EJ speakers' scores were significantly higher than the IJ speakers' scores ($p < .01$), whereas the NE and EJ speakers did not differ significantly ($p > .10$).

Japanese (NT) judged some of the IJ speakers' productions of English /ɹ/ and /l/ to be very much like Japanese /r/ (see also Cochrane, 1977). Japanese /r/ tokens may be heard by native English listeners as a variety of English phones, including /ɹ/ and /l/ (Jones, 1967; Sekiyama and Tohkura, 1993). Alternatively, one might speculate that the IJ speakers were developing phonetic representations for English /ɹ/ and /l/. These representations may have been less well defined than those of native English speakers so that, when evoked in the process of speech production, they did not provide a perceptually adequate specification of English liquids.

According to the subjective observation of one of us who speaks English natively (JEF), the IJ speakers never said /w/ in place of /ɹ/ or /l/, something that occurs in the speech of certain children learning English as their native language (e.g., Snow, 1963). The lack of such a substitution in the speech of the IJ speakers is interesting because L1 sounds are often used in place of L2 sounds, and because Japanese has a /w/ (see Wode, 1978). Perhaps such substitutions occurred at an earlier stage of L2 learning. Alternatively, the sequence of development in children's L1 acquisition may not be recapitulated by adults learning an L2.

Two conclusions can be drawn from Experiment 1 with some certainty. One is that the EJ speakers produced English liquids more accurately than did the IJ speakers. The other is that certain Japanese adults who are experienced in English can learn to produce correctly identifiable tokens of English /ɹ/ and /l/. One caveat with respect to this last finding should be mentioned here. The 12 experienced Japanese speakers who participated in this study are not necessarily representative of the majority of Japanese speakers who learn English in adulthood. It proved difficult to recruit a highly experienced group of Japanese speakers because Japanese adults who live in the US typically return to Japan after several years. Perhaps members of our EJ group decided to remain in the US, at least in part, because they were unusually successful in acquiring English as an L2.

EXPERIMENT 2

It would probably be premature to conclude, on the basis of the Experiment 1 results, that the experienced Japanese (EJ) speakers had "mastered" English liquids. A high rate of correct identifications does not mean necessarily that their liquids conformed fully to English phonetic norms. For example, Munro et al. (in press) found that highly experienced native Italian speakers of English produced English vowels that were identified correctly nearly as often as native English speakers' vowels, yet the native Italian speakers who had begun learning English as adults were found to produce English vowels with a foreign accent. The purpose of this experiment was therefore to evaluate English liquids spoken by the native Japanese speakers for degree of perceived foreign accent.

Method

Ten listeners (three males, seven females) with mean age of 27 years (range: 22 to 32 years) rated English liquids spoken by native and Japanese speakers of English. The listeners were graduate students at the University of Alabama at Birmingham (originally from Massachusetts, New York, Kansas, Louisiana, California, or Tennessee) who all passed a pure-tone hearing screening. Six had participated previously in Experiment 1.

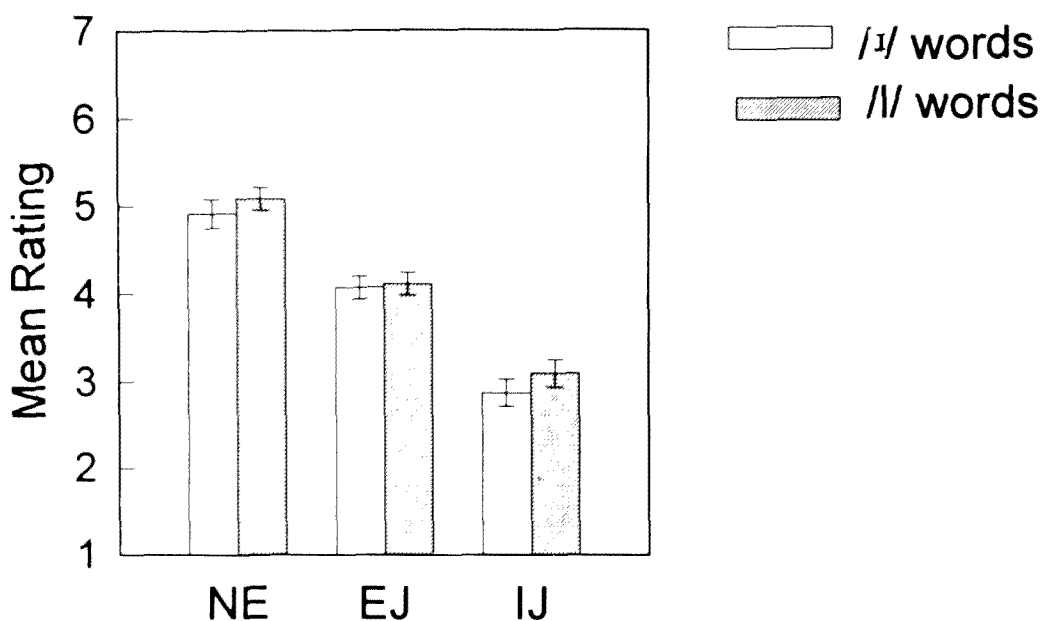


Fig. 3

Mean foreign accent ratings (1 = most foreign; 7 = most native) obtained in Experiment 2 for word-initial tokens of /ɹ/ and /l/ spoken by native English (NE), experienced Japanese (EJ), and inexperienced Japanese (IJ) speakers.

The stimuli were ten CVs from Experiment 1 that were derived from the word-list productions of *read-lead*, *root-loot*, *rate-late*, *road-load*, and *rot-lot*. The vowels in these words (viz., /i e¹ a o u/) resemble the five vowels of Japanese. This was expected to reduce the possibility that an unfamiliar vowel context might impair the Japanese speakers' production of English liquids, or that non-English vowel qualities, if heard by listeners, might influence their ratings of the liquids. The CVs with /ɹ/ and /l/ were randomly presented twice each in separate, counterbalanced blocks with a fixed 1.0-sec interval between each response and presentation of the next stimulus. The listeners were informed of the intended identity of the liquid being rated in each block of CVs. They were told to rate each liquid on a scale ranging from "strong foreign accent" (1) to "no foreign accent" (7), and were urged to use the whole scale. Thus, the more accurate were the speakers' productions of the English liquids, the higher the rating.

Results

The mean ratings shown in Figure 3 have been averaged over the five vowel contexts. Similar ratings were obtained for the CVs with /ɹ/ and /l/ (4.0, 4.1). Averaged over the two liquids, ratings obtained for the NE speakers were higher (5.0) than were the ratings obtained for the EJ or IJ speakers (4.1, 3.0).

A mean rating was obtained for each speaker's production of /ɹ/ and /l/ in the five vowel contexts. These means, which were each based on 20 ratings (10 listeners × 2 random presentations), were submitted to a (3) Group × (2) Consonant × (5) Vowel Context ANOVA. This analysis yielded a significant main effect of Group, $F(2,33) = 20.0$, $p < .01$, but a nonsignificant Consonant main effect, $F(1,33) = 1.0$, $p > .05$, and 2-way interaction, $F(2,33)$

$< 1, p > .10$. A Tukey's HSD test was carried out to test for pairwise differences between groups. It examined the average ratings obtained for /ɹ/ and /l/ spoken by each speaker. The NE speakers' liquids received significantly higher ratings than the EJ speakers', whose liquids in turn received higher ratings than the IJ speakers' ($p < .05$). This finding differs from that one obtained for weighted intelligibility scores in Experiment 1. It suggests that the EJ speakers did not produce English liquids as accurately as did the NE speakers.⁶ A second ANOVA examined ratings obtained for each of the ten listeners by averaging over the stimuli produced by the 12 speakers per group. The analysis of listener-based ratings yielded the same results as the analysis of speaker-based ratings.⁷

Discussion

Liquids spoken by the EJ speakers received significantly lower ratings than did the NE speakers' liquids. This finding suggests that although the EJ speakers produced highly intelligible liquids, the liquids produced by at least some of them diverged from the phonetic norms of English. An overall mean was computed for liquids spoken by all 36 speakers (each based on 10 listeners \times 5 vowel contexts \times 2 = 100 ratings). The mean ratings for /ɹ/ and /l/, which have been plotted in Figure 4 against the weighted intelligibility scores from Experiment 1, provide some insight as to why ratings obtained for the NE and EJ speakers differed in terms of foreign accent ratings but not intelligibility. Highly intelligible tokens received a wide range of foreign accent ratings. However, tokens that were unintelligible (or identified correctly, but with low confidence) typically received very low ratings, indicating that they were strongly foreign-accented. Munro and Derwing (1995) obtained a similar pattern of results in a study examining English utterances spoken by native and Mandarin speakers of English. They suggested that listeners may attend to different stimulus properties when rating utterances for comprehensibility as opposed to foreign accent. The listeners in Experiment 2 always knew the intended identity of the liquids being rated in each block, which may have encouraged them to attend to within-category differences. Listeners in the identification task (Experiment 1) may have discarded such information because it was not essential to defining category membership. Additional research will be needed to assess the relation between intelligibility and degree of foreign accent in phonetic segments produced by nonnative speakers.

⁶ The significant three-way interaction yielded by the ANOVA, $F(4, 132) = 4.08, p < .01$, appears to have arisen because somewhat higher ratings were obtained for /l/ than for /ɹ/ spoken in the context of mid and back vowels (/a/ 4.6 vs. 4.4; /o/ 4.4 vs. 3.9; /u/ 3.7 vs. 3.4) whereas the reverse held true for liquids spoken in the context of *front* vowels (/i/ 3.7 vs 3.9; /e/ 4.0 vs. 4.2). It is uncertain if the vowel context effect was due to variations in how liquids were produced, how they were perceived by the listeners, or both. Whatever the cause, the lack of a significant three-way interaction, $F(8, 132) = 1.96, p > .05$, suggests that the vowel effect was much the same for all three groups.

⁷ A significant Group main effect was obtained, $F(2, 27) = 20.2, p < .01$, but not a significant Consonant main effect, $F(1, 27) = 4.19$, or a two-way interaction, $F(2, 27) = 0.45; p > .05$. A Tukey's HSD test again revealed that ratings were significantly higher for liquids spoken by the NE speakers than by the EJ speakers, whose liquids revealed received significantly higher ratings than the IJ speakers' ($p < .05$).

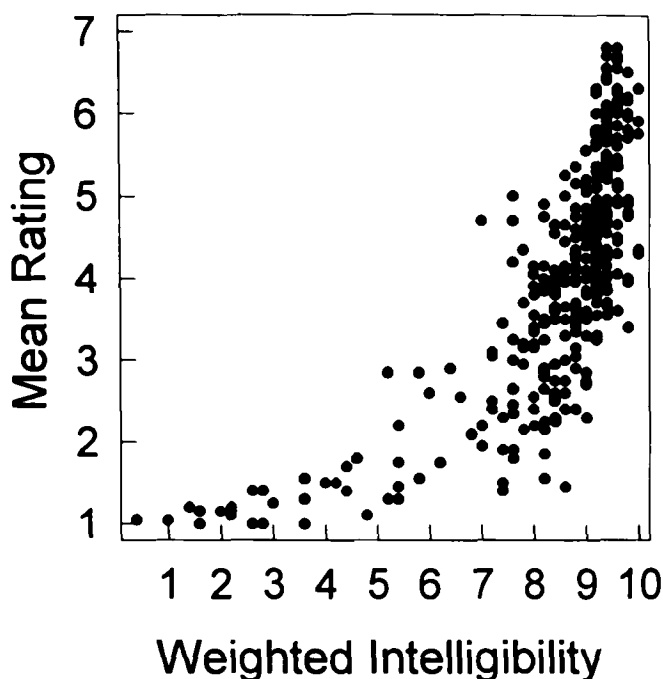


Fig. 4

The relation between weighted intelligibility scores obtained in Experiment 1 and mean foreign accent ratings obtained in Experiment 2 for English liquids spoken by 12 native English (NE) speakers, 12 experienced Japanese (EJ) speakers of English, and 12 inexperienced Japanese (IJ) speakers.

Finally, it is important to note that although the EJ speakers' liquids received significantly lower ratings on average than did the NE speakers' some EJ speakers produced English liquids quite well. The /ɹ/ tokens of seven EJ speakers (as against just one IJ speaker) received a mean rating that fell within the range of mean ratings obtained for the NE speakers.⁸ The /l/ tokens of eight EJ (but just three IJ speakers) received mean ratings that fell within the NE speakers' range.

EXPERIMENT 3

The ratings obtained in Experiment 2 suggested that roughly two-thirds of the experienced Japanese (EJ) speakers produced English liquids in a native-like fashion. One wonders, however, if such a conclusion may overestimate EJ speakers' ability to produce English liquids, for all of the liquids examined in Experiment 2 occurred in words that had been read from a written list. Students in Japan approach English primarily through the written word, and they receive explicit instruction in school concerning how to produce English liquids (Nakauchi, 1993). Perhaps the EJ speakers who seemed to produce English liquids

⁸ In this analysis we did not consider the mean ratings obtained for one NE speaker whose productions of /ɹ/ received a lower rating than did the /ɹ/s of the other 11 NE speakers.

accurately employed a special strategy in *reading* that they either could not, or would not, exploit when producing English /ɹ/ and /l/ *spontaneously*. This possibility was tested here by presenting liquids spoken in three speaking tasks to listeners for identification as /ɹ/ or /l/.

Listeners can distinguish speech that has been read from a prepared text from speech that has been produced spontaneously (Remez, Berns, Nutter, Lang, Davachi, & Rubin, 1991). It is said that nonnatives' speech varies according to the interlocutor, and according to the amount of attention they pay to their speech (Beebe, 1977; Tarone, 1979). If nonnatives pay more attention to their pronunciation of English when they read utterances than when they are speaking English spontaneously, one might expect them to produce English liquids less accurately in a spontaneous speech task than in a list reading task. Support for this hypothesis was provided by a voice onset time study carried out by Major (1992), who found that Portuguese speakers produced English /p t k/ less accurately when speaking spontaneously than when reading words from a list.

Other research suggests, however, that the results obtained in Experiments 1 and 2 did *not* overestimate the Japanese speakers' ability to produce English liquids. Munro and Derwing (1994) compared ratings of extemporaneous English narratives spoken by native Mandarin speakers to ratings obtained for read versions of the same narratives (obtained later from the same speakers, using written transcripts of the extemporaneous speech samples). The two sets of foreign accent ratings did not differ significantly. Moreover, data reported by Wenk (1979) and Dickerson and Dickerson (1977) suggested that nonnatives may produce English consonants accurately *more often* in spontaneous speech than in speech that has been read. Comparable results were obtained for longer stretches of speech by Oyama (1976) and Thompson (1991).

Methods

Stimuli. As reported earlier, neither subjective lexical familiarity, cognate status, nor vowel context seemed to exert an important influence on the intelligibility of liquids spoken by the IJ speakers. Experiment 3 therefore focused on a subset of the minimally paired English words examined in Experiment 1: *right-light*, *rock-lock*, *read-lead*, and *rate-late*. One reason for selecting these particular words was that the Japanese speakers produced more of them in spontaneous utterances than certain other words in our corpus (see General Methods section). On average, the EJ speakers produced more spontaneous utterances containing the CVC words of interest than did the IJ speakers (86%, 66%), but fewer than the NE speakers (100%). There were just nine missing tokens of the eight words selected for analysis here.⁹

Stimulus preparation. Preparation of words from the reading task was described earlier. Similar procedures were used to derive CV stimuli from words produced in the two other

⁹ Two missing tokens were for EJ speakers (both tokens of *lead*) and seven for IJ speakers (*lead-1*, *lock-2*, *light-1*, *read-1*, *rate-2*). The NE speakers' spontaneous utterances were slightly longer ($M = 5.0$ words per utterance, including function words) than the Japanese speakers' (EJ-4.4, IJ-4.3).^{*} The average number of words in spontaneous utterances containing the eight words to be examined here was 4.7 for the NE speakers, and 4.4 words per utterance for both Japanese groups.

speaking tasks. Words from the definition task occurred following a fricative that terminated the carrier phrase *The next word is* __. Words that were spoken in the spontaneous speech task occurred in a variety of phonetic contexts. Thus, deriving CV stimuli from words spoken in these latter two tasks required special care.

A Kay CSL was used to display a time-domain waveform of each word being edited along with a spectrogram (and, in many instances, an overlaid LPC formant track). As had been the case for words from the reading task, the first segmentation point for CVCs preceded by silence was the onset of signal energy. For words preceded by a phonetic segment, an attempt was made to preserve as much as possible of the “steady state” portion of the word-initial /ɪ/ or /l/ (if there was one) without including any trace of the preceding phonetic context. For example, when *read* was said in a phrase such as “*I read...*” the preceding phonetic context (/a¹/) was evident by a falling transition into the /ɪ/. In such an instance, the editing cut was made after the completion of the fall in formant frequencies. Had this not been done, the acoustic evidence of coproduction might have made the /ɪ/ token sound distorted when presented in isolation even if it had not sounded distorted in its original context. After editing, no trace of the preceding context could be heard by the first author. To make the second editing cut, the second author (NT) placed a cursor at the onset of the signal, then moved it toward the vowel midpoint in 20-ms steps until he could first hear an identifiable vowel quality (e.g., /i/ in tokens of *read*). Everything following this point was discarded.

Table 3 provides a tabulation of the phonetic contexts which preceded liquids in words spoken in the spontaneous speech task. The speakers in all three groups produced more tokens of /l/ than /ɪ/ following the fricatives /z/ and /s/ (e.g., in phrases such as *He was late*). However, there does not appear to have been systematic differences between groups in the nature of the preceding phonetic contexts. Differences in the duration of CVs derived from words spoken in the three tasks might also be expected to affect perceptual judgments of word-initial liquids. The mean durations of the edited CV stimuli are summarized in Table 4. (No attempt was made to demarcate the “liquid” and “vowel” portions because of the difficulty inherent in doing so.) The CV durations were submitted to a (3) Group × (3) Task × (2) Consonant ANOVA. Neither the Group main effect, $F(2,33) < 1$, nor any interaction involving this factor reached significance. The Task factor did reach significance, however, $F(2,66) = 12.9, p < .01$, because definition task CVs were significantly longer than reading and spontaneous task CVs according to a Tukey’s HSD test ($M = 169$ vs. 152, 148 ms; $p < .05$).

In some instances, the second editing cut was made before the (normalized) peak amplitude was reached in the original CVC words. Peak amplitude was less than full scale in CVs edited from these words. This occurred less often for CVs edited from definition-task words (20% of tokens) than for CVs edited from words spoken in the reading and spontaneous speaking tasks (47% and 46% of tokens). However, the percentage of CVs in which the editing cut was made before maximum amplitude in the word had been reached was much the same for all three groups (NE-40%, EJ-36%, IJ-36%). It is uncertain if the duration and amplitude differences across speaking tasks were due to differences in production, or to slightly different auditory criteria in segmentation.

Listeners. The 12 native English listeners (three males, nine females) who rated CV

TABLE 3

Tabulation of the phonetic contexts preceding /ɹ/ and /l/ tokens in spontaneous utterances produced by native English (NE), experienced Japanese (EJ), and inexperienced Japanese (IJ) speakers

CONTEXT	NE		EJ		IJ	
	/ɹ/	/l/	/ɹ/	/l/	/ɹ/	/l/
schwa or hesitation vowel	17	12	16	10	8	8
other vowel	9	5	9	8	6	10
silence	8	12	10	16	13	9
/s, z/	1	10	3	8	3	8
stop, affricate	4	4	5	2	6	0
nasal	4	2	4	1	6	6
liquid	5	2	1	1	2	3
other	0	1	0	0	1	0
TOTAL	48	48	48	46	45	44

TABLE 4

Mean duration (ms) of CV stimuli derived from words produced in three speaking tasks by native English (NE), experienced Japanese (EJ), and inexperienced Japanese (IJ) speakers

	NE		EJ		IJ	
	/ɹ/	/l/	/ɹ/	/l/	/ɹ/	/l/
Reading Task	167	157	151	138	154	148
Definition Task	168	180	159	162	177	166
Spontaneous	145	144	155	155	142	147

stimuli in this experiment had a mean age of 27 years (range: 23 to 34 years). They were graduate students at the University of Alabama at Birmingham who came from a variety of places in the US (New York, North Dakota, California, New Jersey, Kansas, Massachusetts, Alabama). All passed a pure-tone hearing screening.

Procedure. The stimuli were low-pass filtered (4.8 kHz) before being presented binaurally at a comfortable level to the listeners, who were tested individually in a sound booth. The listeners pushed a button marked “L” or “R” depending on what they heard. They were required to respond to each stimulus, and were told to guess if unsure. The 288 stimuli derived from words produced in each speaking task were randomly presented in separate, counterbalanced blocks. There was a 1.0-sec interval between each response and presentation of the next stimulus. Different randomizations were used for each listener. Ten practice stimuli at the beginning of each block were not analyzed.

Results and Discussion

Six percent correct identification scores were calculated for each speaker, one each for /ɹ/s and /l/s spoken in the three speaking tasks. Each mean was based on 48 forced-choice judgments (4 words × 12 listeners). As shown in Figure 5(a), liquids produced by the NE and EJ speakers were identified correctly more often (96%, 97%) than were liquids produced by the IJ speakers (79%). The significance of these between-group differences could not be tested, however, because of significant differences in variances across speaker groups.

An effect of speaking task is evident only for the IJ speakers, perhaps because only their scores were not at ceiling. The IJ speakers’ liquids were identified correctly more often in words produced in the reading and definition tasks (85%, 83%) than in the spontaneous speech task (68%). After arcsine transformation, the percent correct identification scores obtained for the IJ speakers were submitted to a (3) Task × (2) Consonant repeated measures ANOVA. Neither the Consonant factor, $F(1,11) < 1$, nor 2-way interaction reached significance, $F(2,22) < 1$. The Task factor was significant, however, $F(2,22) = 9.44, p < .01$

A Tukey’s HSD test examining scores averaged over the two liquids revealed that the IJ speakers’ liquids were identified correctly significantly less often in the spontaneous speech task than in the other two tasks ($p < .01$). This suggests that results obtained in a reading task may overestimate the performance of relatively inexperienced speakers of an L2, at least when segmental articulation is evaluated in a forced-choice test. One might speculate that individuals whose first exposure to the target L2 occurred in an academic setting benefit from the use of certain production strategies when reading L2 words from a list. Additional research will be needed, however, to understand the influence of speech elicitation techniques (and, more generally, speech style or register) on L2 production samples.

EXPERIMENT 4

This experiment used a rating task to evaluate the accuracy with which the NE and native Japanese speakers produced English liquids in three speaking tasks.

Method

The 12 listeners from Experiment 3 rated the same stimuli for degree of foreign accent in a session held several days later. The stimuli were presented in eight blocks, one each for

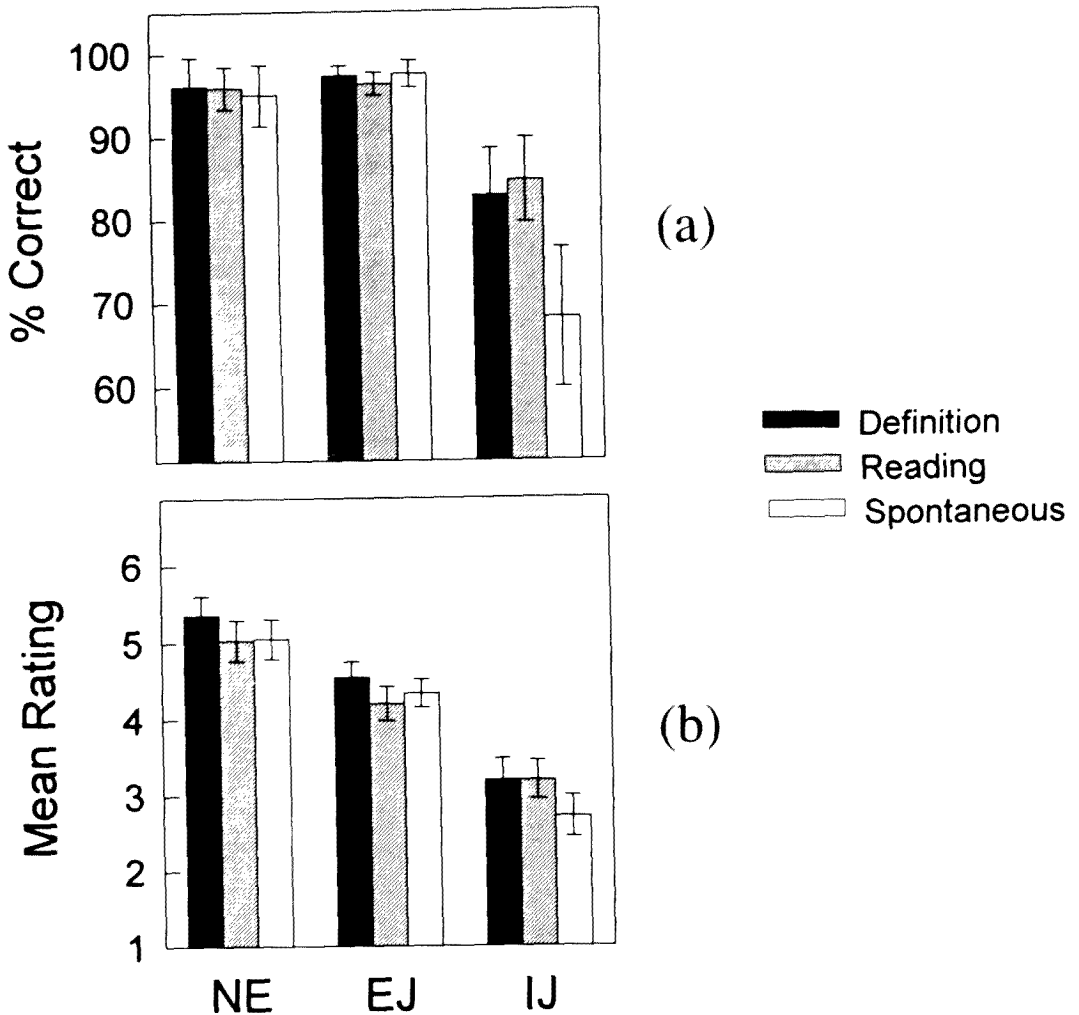


Fig. 5.

(a) Mean percent correct identification scores obtained in Experiment 3 for /ɹ/ and /l/ tokens spoken by native English (NE), experienced Japanese (EJ), and inexperienced Japanese (IJ) speakers.

(b) Mean foreign accent ratings obtained in Experiment 4 for the same /ɹ/ and /l/ tokens.

CVs derived from words beginning in /ɹ/ (*right, rock, read, rate*) and /l/ (*light, lock, lead, late*). Six listeners heard the /ɹ/ words first (in counterbalanced order), and six heard the /l/ words first. The listeners, who always knew the identity of the one liquid being presented in a block, rated each liquid using a scale that ranged from "strong foreign accent" (1) to "no foreign accent" (7). They were told to use the entire scale, and were required to respond to each stimulus. Most blocks contained 108 randomly presented CV stimuli. The CVs in each block were preceded by 36 practice stimuli that were not analyzed. Separate randomizations of the stimuli were used for each listener. The interval between each response and presentation of the next stimulus was 1.0 sec.

Results and Discussion

Six mean ratings were calculated for each of the 36 speakers, one each for the /ɪ/s and /I/s spoken in each speaking task. Each of these means was based on 48 judgments (4 words × 12 listeners). As shown in Figure 5(b), the NE speakers' liquids received somewhat higher ratings (5.2) than the EJ and IJ speakers' (4.5, 3.0). Ratings were slightly higher for liquids produced in the definition task than in the reading or spontaneous speech tasks.¹⁰

Mean speaker-based ratings were submitted to a (3) Group × (3) Task × (2) Consonant ANOVA. The Group × Task interaction, $F(4,66) = 4.2, p < .01$, it yielded was explored by tests of simple main effects. The simple effect of Task was significant for all three groups, F -values ranging from 4.8 to 12.1, $p < .05$. Tukey's tests revealed that liquids spoken in the definition task by the NE and EJ speakers received significantly higher ratings than liquids spoken in the reading task ($p < .05$). For the NE speakers, definition-task liquids also received higher ratings than did liquids from the spontaneous speech task ($p < .05$). For the IJ speakers, liquids spoken in the spontaneous speech task received significantly lower ratings than did liquids spoken in the other two tasks ($p < .05$). (No other differences reached significance.) The rating data obtained for the IJ speakers corresponded to the identification data obtained in Experiment 3, where liquids were identified significantly less often in spontaneously spoken words than in words that were had been read from a list.

The simple main effect of the Group factor proved to be significant in all three speaking tasks, F -values ranging from 17.0 to 30.0, $p < .01$. According to Tukey's HSD tests, liquids spoken by NE speakers in the reading and definition tasks received significantly higher ratings than did the EJ speakers' liquids, whose liquids received higher ratings than the IJ speakers' ($p < .05$). The difference between the NE and EJ speakers failed to reach significance for liquids spoken in the spontaneous speech task, but liquids spoken by these two groups received significantly higher ratings than the IJ speakers' liquids ($p < .05$).

Mean ratings were calculated for each of the 12 listeners by averaging over the ratings obtained from the 12 speakers per group. The Group × Task × Consonant ANOVA carried out to examine the listener-based ratings yielded a significant 3-way interaction, $F(4,66) = 5.54, p < .01$. The simple main effect of Group proved to be significant for all six Group × Task combinations, F -values ranging from 29.6 to 100.9, $p < .01$. Tukey's tests revealed that, in each instance, the NE speakers' liquids received significantly higher ratings than did liquids spoken by the EJ speakers, whose liquids in turn received significantly higher ratings than did the IJ speakers' ($p < .05$).

Figure 6 shows mean ratings obtained for liquids spoken by each of the 36 speakers in the three speaking tasks. The individual speaker means shown here were based on a maximum of 96 ratings (12 listeners × 4 words each beginning in /ɪ/ and /I/). Differences between the three speaking tasks were surprisingly small in most instances. The largest differences were between the reading and spontaneous speech tasks. In no instance did

¹⁰ As noted in the Method section, this may have been due to differences in the duration and/or intensity of CVs in the definition task as compared to CVs from the other two tasks.

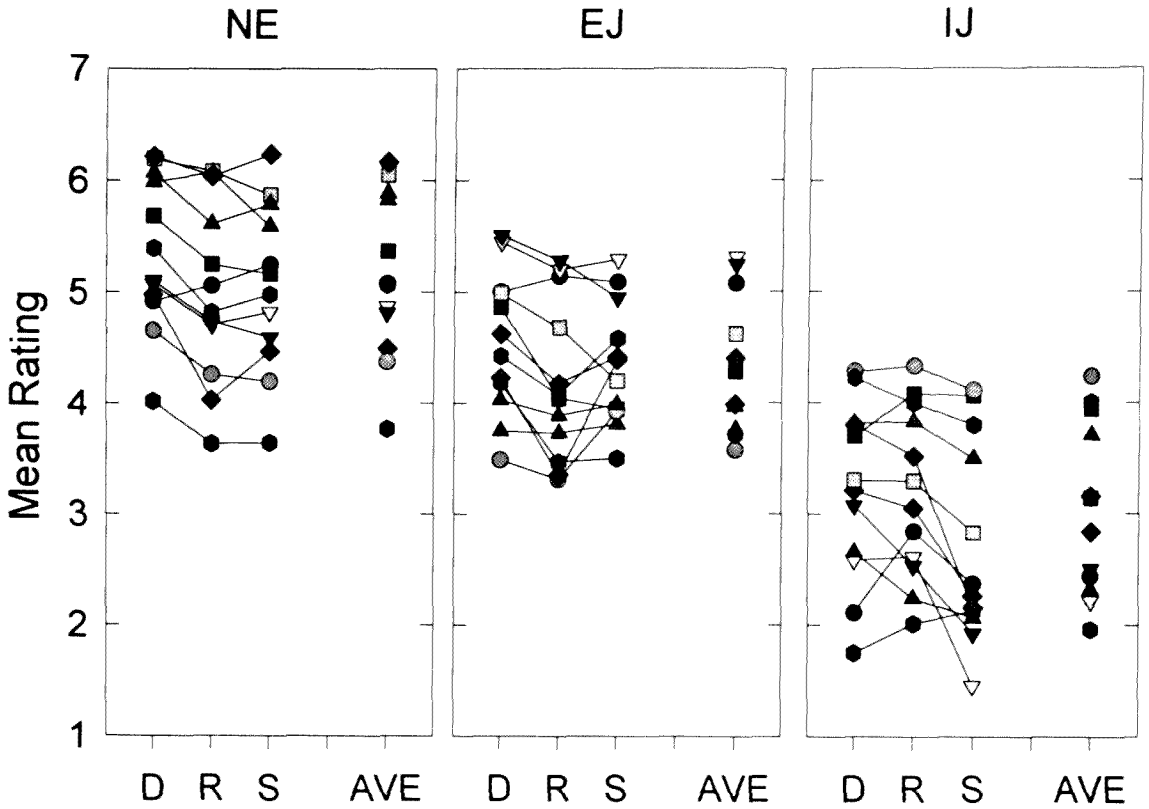


Fig. 6

Mean ratings obtained for /ɹ/ and /l/ tokens produced by 36 speakers in three speaking tasks. "D" indicates the definition task, "R" the word-list reading task, "S" the spontaneous speaking task, and "AVE" the average for the three tasks.

ratings obtained in these tasks differ by more than 0.5 for the NE speakers (on a 7-point scale). However, scores obtained for three EJ speakers in the spontaneous task exceeded those obtained in the reading task by more than 0.5; and the scores obtained for four IJ speakers exceeded the criterion, but in the opposite direction. This finding suggests that data obtained in the reading task may underestimate some EJ speakers' /ɹ/-/l/ production ability but it may *overestimate* some IJ speakers' ability.

Figure 6 also shows the overall average ratings obtained for liquids spoken by each of the 36 speakers. Just three of 12 IJ speakers obtained a mean rating that fell within the range of ratings obtained for the 12 NE speakers, whereas 10 of the 12 EJ speakers obtained such a mean rating. The two EJ speakers whose ratings fell slightly outside the NE range were both 44-year-old women. One of these women had arrived in the US at the age of 25 years and indicated using English frequently; the other arrived at the age of 31 years and indicated using English infrequently.

ACOUSTIC ANALYSIS

Acoustic analyses were undertaken to identify characteristics of the Japanese speakers' productions of English liquids that might lead to the perception of foreign accent. We measured 180 /ɪ/ tokens and 180 /I/ tokens that had been rated for degree of foreign accent in Experiment 2. The acoustic measurements were submitted to multiple regression analyses with foreign accent ratings as the dependent variable.

Methods

Measurements were made from spectrograms using a Kay Computerized Speech Lab. Somewhat different measurements had to be made of the /ɪ/ and /I/ tokens because not all measures could be applied in an evenhanded fashion to both liquids. However, in view of the report that Japanese speakers may not produce a F3 difference between English /ɪ/ and /I/ (Yamada & Tohkura, 1992), F3 frequency was measured in all tokens.

The duration of a "steady state" portion in /I/ tokens was measured from the onset of energy in the region of F2 or F3 (whichever occurred first) to the end of a constriction phase. The release of lingual constriction for /I/ was often marked by a short, wide spectrum release burst (see Dalston, 1975). Secondary characteristics marking the release of /I/ were an increase in energy in the region above F2 or a sudden increase in F2 frequency. The frequencies of the first three formants (F1, F2, F3) were measured at the point of release. Frequencies were not measured during the steady state portion because F2 and/or F3 was not always evident.

The /ɪ/s typically did not show acoustic evidence of lingual release (see also Yamada & Tohkura, 1990). For /ɪ/ tokens, F1, F2, and F3 frequencies were measured at the first location in the waveform at which energy was evident for all three formants. A second set of frequency measurements was made at the point where F3 frequency began rising rapidly (or where F2 frequency increased, in the absence of F3 change). The duration of the steady state portion was the interval between the two locations where frequency measurements were made.

Not every token could be measured using the procedures just described. As indicated in Tables 3 and 4, the number of measures per liquid obtained for each group ranged from 48 to 57 (out of a maximum possible of 60 tokens per group). For example, there were a total of 20 missing values for F2 frequency in the /I/ tokens, and a total of 30 missing F3 onset frequency values for the /ɪ/ tokens.

Results and Discussion

Table 5 presents the mean acoustic values obtained for /I/ tokens produced by the three groups. There was little difference between groups in the mean values obtained for F1 and F3. However, F2 frequencies were somewhat higher in /I/ tokens produced by the two Japanese groups than by the NE speakers. Also, the steady state portion was somewhat shorter in tokens produced by the IJ speakers than by the NE or EJ speakers.

Values of the four acoustic measures in /I/ were submitted to a multiple regression analysis. It accounted for 44% of the variance (adjusted *R*-squared) in the mean foreign accent ratings obtained for the 180 /I/ tokens (36 speakers × 5 vowel contexts). Of the four

TABLE 5

Acoustic measurements of /l/ tokens spoken by native English (NE), experienced Japanese (EJ), and inexperienced Japanese (IJ) speakers (*n* = number of measurable tokens)

		NE	EJ	IJ
F1 ^a	<i>M</i>	603	617	621
	<i>SD</i>	(127)	(95)	(89)
	<i>n</i>	54	50	56
F2 ^a	<i>M</i>	1429	1724	1810
	<i>SD</i>	(252)	(404)	(409)
	<i>n</i>	54	50	56
F3 ^a	<i>M</i>	2854	2890	2942
	<i>SD</i>	(627)	(282)	(392)
	<i>n</i>	53	50	56
DUR ^b	<i>M</i>	84	85	77
	<i>SD</i>	(36)	(29)	(35)
	<i>n</i>	51	50	56

^aMeasured at the point of rapid spectral change, in Hz

^bDuration of the steady state portion, in ms

variables examined, three were identified as significant predictors of foreign accent in /l/. The beta weight for F2 frequency (-.647) was greater than the beta weight for F3 frequency (.412) or steady state duration (.247). When just the data for the NE and EJ speakers were examined, the same three acoustic variables accounted for 52% of the variance. Again, F2 frequency appeared to be the most important acoustic variable.¹¹ The simple correlation between F2 frequency and foreign accent in the 160 available /l/ tokens was $r = -0.536$ ($p < .001$). The lower (more English-like) were the F2 frequencies, the less foreign-accented the /l/ tokens were judged to be. It may be that /l/ tokens produced by the NE speakers, but not by the native Japanese speakers, were slightly velarized. This would be expected to lower F2 frequency (Ladefoged, 1975).

Table 6 presents the mean acoustic values obtained for the /l/ tokens. Both groups of Japanese speakers seem to have produced larger F3 differences between English /l/ and /l/ than did the nearly monolingual Japanese speakers examined by Yamada and

¹¹ Much the same results were obtained in analyses of /l/ and /l/ tokens when values in Hertz for these liquids were converted to Barks or Bark difference scores.

TABLE 6

Acoustic measurements of /ɪ/ tokens produced by the native English (NE), experienced Japanese (EJ), and inexperienced Japanese (IJ) speakers

		NE	EJ	IJ
F1 ^a	<i>M</i>	456	425	470
	<i>SD</i>	(72)	(69)	(94)
	<i>n</i>	48	52	50
F2 ^a	<i>M</i>	1121	1144	1311
	<i>SD</i>	(248)	(174)	(232)
	<i>n</i>	48	52	50
F3 ^a	<i>M</i>	1750	2106	2261
	<i>SD</i>	(414)	(306)	(315)
	<i>n</i>	48	52	50
F1 ^b	<i>M</i>	503	488	506
	<i>SD</i>	(70)	(74)	(103)
	<i>n</i>	57	55	55
F2 ^b	<i>M</i>	1300	1274	1416
	<i>SD</i>	(246)	(170)	(243)
	<i>n</i>	57	55	55
F3 ^b	<i>M</i>	1922	2160	2272
	<i>SD</i>	(497)	(279)	(344)
	<i>n</i>	57	55	55
DUR ^c	<i>M</i>	62	58	59
	<i>SD</i>	(29)	(27)	(32)
	<i>n</i>	48	52	49

^aMeasured at the point of onset of one or more formants, in Hz^bMeasured at the point of rapid movement of F3, in Hz^cDuration of the steady state portion, in ms

Tohkura (1992). The F2 and F3 frequencies in /ɪ/ tokens spoken by the NE speakers were closer than those in /ɪ/ tokens spoken by the two Japanese groups, both for measurements made at formant onset and at the point of rapid change. The difference was mostly due to lower F3 values on the part of the NE speakers.

The acoustic measurements accounted for 65% of the variance (adjusted *R*-squared) in ratings for /ɪ/ tokens. Of the seven acoustic variables examined, four were identified as significant predictors of foreign accent in the multiple regression analysis: F2 frequency at onset, F3 frequency at onset, F2 frequency at the point of rapid spectral change, and F3

frequency at the point of rapid spectral change. The beta weight for F3 onset was higher ($-.965$) than were beta weights for the other three variables ($.237$ to $-.467$). When just data for the NE and EJ speakers were included, 70% of variance was accounted for. Once again, the F3 onset frequency variable had the highest beta weight. F3 onset frequency showed a significant simple correlation with the foreign accent ratings ($r = -.759$; $p < .01$). The lower (more English-like) were the F3 values in /ɪ/ tokens, the less foreign-accented they were judged to be.

GENERAL DISCUSSION

As expected, productions of English /ɪ/ and /l/ by Japanese adults who were relatively inexperienced in English (two years of residence in the United States) were often misidentified by native English-speaking listeners. Their productions were also judged to be strongly foreign-accented. However, productions of English /ɪ/ and /l/ by Japanese adults who were experienced in English (21 years of residence in the US) were identified correctly, and received only slightly lower ratings than did liquids spoken by native English (NE) speakers. When ratings obtained for liquids produced in three speaking tasks were averaged, those produced by 10 of the 12 experienced Japanese (EJ) speakers received a mean rating that fell within the range of ratings obtained for 12 NE speakers. Ratings obtained for the remaining two EJ speakers fell barely outside the NE speakers' range. These results suggest that many Japanese adults eventually learn to produce English /ɪ/ and /l/ accurately.

Hurford (1991) outlined three broad hypotheses that might be advanced to explain age constraints on the acquisition of a second language (L2). According to the *exercise hypothesis*, a capacity for learning to process environmentally relevant stimuli will disappear or decline if it ceases to be used (see also Bever, 1981). According to the *maturational state hypothesis*, abilities may decline as a consequence of maturation, without regard to use (see also Johnson and Newport, 1989). Finally, according to the *interference hypothesis*, L2 learning is adversely affected by the prior acquisition of a native language (see also Grosjean, 1989). The results obtained in the present study seem to disconfirm at least a strong version of all three hypotheses. The experienced Japanese speakers examined here did not learn languages continuously between the time they acquired Japanese and the time of their arrival in the US. They were not exposed to English prior to the age (12 years) that is traditionally considered to mark the end of a critical period for L2 learning (Scovel, 1988). Finally, the EJ speakers indicated that they continued using their native language on a regular basis. Despite all of this, many EJ speakers managed to learn to produce English /ɪ/ and /l/ accurately.

As mentioned earlier, however, we cannot be certain that the EJ speakers we examined were representative of adult L2 learners, for they were not chosen randomly from a larger population. It is also important to emphasize that foreign accents in an L2 increase in strength as the age of learning an L2 increases, and that few if any adults manage to avoid speaking their L2 without a detectable foreign accent, no matter how experienced they might be in the L2 (Flege et al., 1995b). Although certain L2 vowels and consonants may be produced accurately by adult learners of an L2, this appears to be the exception rather than the rule. Highly experienced adult learners of an L2 have been shown to err persistently in producing a wide range of L2 consonants and vowels (Flege

et al., 1995a; Munro et al., in press). Additional research is clearly needed to determine *which* L2 vowels and consonants can be mastered, as well as to determine what kind of phonetic input is needed for successful learning to occur.

It would be of theoretical and practical interest to account for the phonetic errors made by the inexperienced Japanese (IJ) speakers of English. Based on the performance of their more experienced compatriots, one might conclude that the IJ speakers' errors were not the result of their having passed a critical period. As discussed in the Introduction, their errors might be attributed to cross-language phonetic interference. The IJ speakers probably identified English /ɹ/ and /l/ with Japanese /r/. If so, they may have substituted Japanese /r/ for one or both English liquids, either without any modification or else after altering their typical production to accommodate acoustic differences they had discerned to exist between the English liquids and Japanese /r/ (see Flege, 1991). Or, perhaps their errors were simply due to insufficient phonetic input. After all, many native English 5- and 6-year-olds make errors in producing English liquids. Mastery of English liquids requires years of practice.

It would also be of interest to account for the accuracy with which the EJ speakers produced English liquids. Their success may be an instance of what MacWhinney (1992) called "direct learning" in an L2, that is, learning not encumbered by previously established native language structures. If this were the case, then one might surmise that English liquids which at one time were identified with Japanese consonants later ceased to be so identified.

The Speech Learning Model (SLM) proposed by Flege (1995) would lead one to attribute the IJ speakers' liquid production errors to the absence of phonetic categories for English /ɹ/ and /l/. It would lead one to expect that EJ speakers who produced English /ɹ/ and /l/ accurately did so not by modifying previously established articulatory patterns but by establishing new categories. According to the SLM, adults remain capable of forming phonetic categories for certain L2 vowels and consonants. The model proposes that when a phonetic category is eventually formed for an L2 consonant, it will be produced accurately provided that the new consonant category is based on the features used by monolingual speakers of the target L2 and is not "deflected away" from the nearest L1 consonant in order to preserve contrast in a common L1-L2 phonological space.

This hypothesis cannot be confirmed or disconfirmed on the basis of the data presented here. To do so will require perceptual testing coupled with a careful assessment of speech production. Support for the hypothesis would be provided if it were shown that Japanese speakers who produced English liquids accurately, but not those who erred in producing English liquids, perceived English liquids like native English speakers. The categorical discrimination task described by Flege, Munro and Fox (1994) would provide an appropriate perceptual test. In this triadic task, subjects must identify the serial position of an odd item out. Three aspects of the task ensure that it tests the ability to group phone classes into *phonetically relevant* categories. (The interval between stimuli is relatively long; the three stimuli in each triad are always spoken by different speakers; and "catch" triads are included that do not contain an odd item out.) The SLM predicts that Japanese speakers who produce English liquids accurately will be able to discriminate English /ɹ/ and /l/ from one another and also from the nearest Japanese consonants (/r/, /w/) whereas speakers who err in producing English liquids will often fail to choose the odd item out.

In summary, the results obtained here suggest that Japanese adults may learn to produce English liquids accurately if they speak English for many years. This argues against the view that the errors made by inexperienced Japanese speakers in producing English /ɹ/ and /l/ result from their having passed a critical period for speech learning. However, our understanding of many crucial aspects of second language speech learning remains limited. Specific questions regarding the acquisition of L2 consonants raised in the present study must be addressed in future research.

(Part 1 received: December 13, 1993; part 2 received: August 17, 1994; accepted February 22, 1995)

REFERENCES

- BEEBE, L. (1977). The influence of the listener on code-switching. *Language Learning*, 27, 331 – 339.
- BEST, C., & STRANGE, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, 20, 305 – 330.
- BEVER, T. (1981). Normal acquisition processes explain the critical period for language learning. In K. Diller (Ed.) *Individual differences in language learning aptitude*. Rowley MA: Newbury House.
- BUCKINGHAM, H., & YULE, G. (1987). Phonemic false evaluation: Theoretical and clinical aspects. *Clinical Linguistics and Phonetics*, 1, 113 – 125.
- COCHRANE, R. (1977). *The acquisition of /r/ and /l/ by Japanese children and adults learning English as a second language*. Unpublished Ph.D. dissertation, University of Connecticut.
- DALSTON, R. (1975). Acoustic characteristics of English /w,r,l/ spoken correctly by young children and adults. *Journal of the Acoustical Society of America*, 57, 462 – 469.
- DICKERSON, L. (1974). *Internal and external patterning of phonological variability in the speech of Japanese learners of English: Toward a theory of second-language acquisition*. Unpublished Ph.D. dissertation, University of Illinois, Urbana-Champaign.
- DICKERSON, L., & DICKERSON, W. (1977). Interlanguage phonology: Current research and future directions. *Actes du 5ème Colloque de Linguistique Appliquée de Neuchatel*. (pp. 18 – 30). Geneva: Droz.
- FLEGE, J. (1987a). A critical period for learning to pronounce foreign languages? *Applied Linguistics*, 8, 162 – 177.
- FLEGE, J. (1987b). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47 – 65.
- FLEGE, J. (1988). The production and perception of speech sounds in a foreign language. In H. Winitz (Ed.) *Human communication and its disorders: A review – 1988* (pp. 224 – 401). Norwood, NJ: Ablex.
- FLEGE, J. (1991). Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, 89, 395 – 411.
- FLEGE, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 229 – 273). Timonium, MD: York Press.
- FLEGE, J., & FLETCHER, K. (1992). Talker and listener effects on the perception of degree of foreign accent. *Journal of the Acoustical Society of America*, 91, 370 – 389.

- FLEGE, J., & MUNRO, M. (1994) The word unit in L2 speech production and perception. *Studies in Second Language Acquisition*, **16**, 381 – 411.
- FLEGE, J., MUNRO, M., & FOX, R. (1994). Auditory and categorical affects on cross-language vowel perception. *Journal of the Acoustical Society of America*, **95**, 3623 – 3641.
- FLEGE, J., MUNRO, M., & MACKAY, I. (1995a). The effect of age of second language learning on the production of English consonants. *Speech Communication*, **16**, 1 – 26.
- FLEGE, J., MUNRO, M., & MACKAY, I. (1995b). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, **97**, 3125 – 3134.
- GOTO, H. (1971). Auditory perception by normal Japanese adults of the sounds “ɪ” and “I” *Neuropsychologia*, **9**, 317 – 323.
- GROSJEAN, F. (1989). Neurolinguists, Beware! The bilingual is not two monolinguals in one person. *Brain and Language*, **36**, 3 – 15.
- HALE, H. (1885). On some doubtful or intermediate articulations. *Journal of the Anthropological Institute*, **14**, 233 – 243.
- HENLY, E., & SHELDON, A. (1986). Duration and context effects on the perception of English /ɪ/ and /I/: A comparison of Cantonese and Japanese speakers. *Language Learning*, **36**, 505 – 521
- HURFORD, J. (1991). The evolution of the critical period for language acquisition. *Cognition*, **40**, 159 – 201.
- JOHNSON, J., & NEWPORT, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, **21**, 60 – 99.
- JONES, D. (1967). *The phoneme: Its Nature and Use*. Cambridge, UK: Cambridge University Press.
- KOUTSOUDAS, A., & KOUTSOUDAS, O. (1983). A contrastive analysis of the segmental phonemes of Greek and English. In B. Robinett & J. Schachter (Eds.), *Second language learning, contrastive analysis, error analysis, and related aspects*. Ann Arbor, MI: University of Michigan Press.
- LADEFOGED, P. (1975). *A course in phonetics*. New York: Harcourt, Brace, and Jovanovitch.
- LEHISTE, I. (1988). *Lectures on language contact*. Cambridge, MA: MIT Press.
- LENNEBERG, E. (1967). *Biological foundations of language*. New York: Wiley.
- LOGAN, J., LIVELY, S., & PISONI, D. (1991). Training Japanese listeners to identify English /ɪ/ and /I/: A first report. *Journal of the Acoustical Society of America*, **89**, 874 – 886.
- LOVINS, J. (1976). *Loanwords and the phonological structure of Japanese*. Unpublished Ph.D. dissertation, University of Chicago.
- MACKAIN, K., BEST, C., & STRANGE, W. (1981). Categorical perception of English /ɪ/ and /I/ by Japanese bilinguals. *Applied Linguistics*, **2**, 369 – 390.
- MACWHINNEY, B. (1992). Transfer and competition in second language learning. In R. Harris (Ed.), *Cognitive processing in bilinguals*. (pp. 371 – 390). Amsterdam: Elsevier.
- MAJOR, R. (1992). Losing English as a first language. *The Modern Language Journal*, **76**, 190 – 208.
- MIYAWAKI, K., STRANGE W., VERBRUGGE, R., LIBERMAN, A., JENKINS, J., & FUJIMURA, O. (1975). An effect of linguistic experience: The discrimination of /ɪ/ and /I/ by native speakers of Japanese and English. *Perception & Psychophysics*, **18**, 331 – 340.
- MOCHIZUKI, M. (1981). The identification of /ɪ/ and /I/ in natural and synthesized speech. *Journal of Phonetics*, **9**, 283 – 303.
- MUNRO, M., & DERWING, T. (1994). Evaluation of foreign accent in extemporaneous and read material. *Language Testing*, **11**, 253 – 266.
- MUNRO, M., & DERWING, T. (1995). Processing costs and the perception of foreign-accented speech. Manuscript submitted for publication.
- MUNRO, M., FLEGE, J., & MACKAY, I. (in press). The effect of age of second-language learning on the production of English vowels. *Applied Psycholinguistics*..

- NAKAUCHI, J. (1993). The identification of /l/ and /r/ by Japanese early, intermediate and late bilinguals. *Speech, Hearing and Language, Work in Progress*, 7, 155 – 179. London: University College, Department of Phonetics and Linguistics.
- OYAMA, S. (1976). The sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5, 261 – 285.
- REMEZ, R., BERNIS, S., NUTTER, J., LANG, J., DAVACHI, L. & RUBIN, P. (1991). On the perceptual differentiation of spontaneous and prepared speech. Paper presented at the 121st meeting of the Acoustical Society of America.
- SCOVEL, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, 19, 245 – 253.
- SCOVEL, T. (1988). *A Time to speak: A psycholinguistic inquiry into the critical period for human speech*. New York: Newbury House/Harper & Row.
- SEKIYAMA, K., & TOHKURA, Y. (1993). Inter-language differences in the influence of visual cues in speech perception. *Journal of Phonetics*, 21, 427 – 444.
- SHELDON, A., & STRANGE, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, 3, 243 – 261.
- SHIMIZU, K., & DANTSUJI, M. (1983). A study of the perception of /r/ and /l/ in natural and synthetic speech sounds. *Studia Phonologica*, 17, 1 – 14.
- SNOW, K. (1963). A detailed analysis of articulation responses of 'normal' first grade children. *Journal of Speech and Hearing Research*, 6, 277 – 290.
- SUTER, R. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning*, 26, 233 – 253.
- TAKAGI, N. (1993). *Perception of American English /r/ and /l/ by adult Japanese learners of English: A unified view*. Unpublished Ph.D. thesis, Department of Cognitive Sciences, University of California, Irvine.
- TARONE, E. (1979). Interlanguage as chameleon. *Language Learning*, 29, 181 – 191.
- THOMPSON, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Lang. Learning*, 4, 177 – 204.
- TRUBETZKOY, N. (1969). *Principles of phonology* (C.A.M. Baltaxe, Trans.). Berkeley, CA: University of California Press. (Original work published 1939).
- VANCE, T. (1987). *An introduction to Japanese Phonology*. Albany, NY: State University of New York Press.
- WENK, B. (1979). Articulatory setting and de-fossilization. *Interlanguage Studies Bulletin*, 4, 202 – 220.
- WODE, H. (1978). The beginning of non-school room L2 phonological acquisition. *International Review of Applied Linguistics*, 16, 109 – 125.
- YAMADA, R., & TOHKURA, Y. (1990). Perception and production of syllable-initial English /r/ and /l/ by native speakers of Japanese. *International Conference on Spoken Language Processing, ICSLP-91*, (pp. 757 – 760). November 1990, Kobe, Japan.
- YAMADA, R., & TOHKURA, Y. (1992). Perception of American English /r/ and /l/ by native speakers of Japanese. In E. Tohkura, E. Vatikiotis-Bateson, & Y. Sagisaka (Eds.) *Speech Perception, Production, and Linguistic Structure* (pp. 155 – 174). Tokyo: Ohmsha.
- YAMADA, R., TOHKURA, Y., & KOBAYASHI, N. (1994). Effect of word familiarity on non-native phoneme perception: Identification of English /r/, /l/, and /w/ by native speakers of Japanese. In A. James & J. Leather (Eds.), *Second Language Speech*. The Hague: Mouton.
-

Copyright of Language & Speech is the property of Kingston Press Ltd. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.