The English l/r Distinction for Native Japanese
Author(s): Yoshiko Higurashi and Charles McClenon

Please see “How to cite” in the online sidebar for full citation information.

Please contact BLS regarding any further use of this work. BLS retains copyright for both print and screen forms of the publication. BLS may be contacted via http://linguistics.berkeley.edu/bls/.

The Annual Proceedings of the Berkeley Linguistics Society is published online via eLanguage, the Linguistic Society of America's digital publishing platform.
The English l/r Distinction for Native Japanese
Yoshiko Higurashi and Charles McClendon
San Diego State University
Xi'an Foreign Languages Institute

1. Introduction

Sheldon (1979) at the Winter Meeting of LSA (Linguistic Society of America) presented an interesting study of native Japanese treatment of the English l/r distinction, suggesting that production may run ahead of perception. That is, Japanese speakers of English as a second language consistently appear to produce an l/r distinction before they can consistently distinguish them in perception. This is contrary to the usual wisdom that by stressing production skills, perception will take care of itself.

We wish to question the findings from two viewpoints: firstly, we have methodological objections, but secondly, assuming that the findings are at least in part true, we wonder whether there is some principled explanation and whether this might show us something of phonological interest.

Sheldon had six native Japanese students read through a randomly ordered list of l/r minimal pairs. Four native English speakers listened to the tapes of these items, and responded in a forced choice task, identifying them as either /l/-words or /r/-words. Conversely, the native English speakers read through the minimal pair list, and the Japanese responded in the forced choice task.

When the tapes made by the Japanese were heard by the English, there were only 1% misidentifications, and for no speaker were there more than 3% misidentifications. Thus, it was judged that the Japanese were good producers. However, native Japanese speakers' errors in judging native English items, averaged across all contexts, ran slightly over 10%. Thus the conclusion that there were perception problems.

We should note, however, that the tapes were made by having each speaker read through the list twice; then the best token of each item was used.

1 Earlier version of this paper was presented at the Summer Meeting of the Linguistic Society of America, August 1980.
Although this procedure was intended to assure acoustically clear tokens, it creates an obvious bias. Production by non-native speakers is presumably more variable in quality than production by natives. The effect of the procedure, then, is to throw out production errors by the Japanese. Further, the laboratory situation in general may induce more careful speech, the words were read from lists, rather than from sentences, so the tokens presented for the Japanese may be unrepresentatively accurate. We will assume, though, that the magnitude of this effect is not great, and that without this bias there would still exist a discrepancy between production and perception scores.

We would like to suggest, however, that the simple fact that native Japanese speakers consistently produced distinguishable /l/ and /r/ tokens does not prove that they have mastered production. A comparison of the phonologies of English and Japanese may suggest other possibilities.

First, however, let us note the contexts in which perception errors were more frequent.

<table>
<thead>
<tr>
<th>Table I</th>
<th>Sheldon's Results: % errors by 5 good Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial Position</strong></td>
<td>% error</td>
</tr>
<tr>
<td>read</td>
<td>0</td>
</tr>
<tr>
<td>room</td>
<td>15</td>
</tr>
<tr>
<td>road</td>
<td>0</td>
</tr>
<tr>
<td>right</td>
<td>0</td>
</tr>
<tr>
<td><strong>Stop-Liquid</strong></td>
<td></td>
</tr>
<tr>
<td>breed</td>
<td>35</td>
</tr>
<tr>
<td>broom</td>
<td>27.5</td>
</tr>
<tr>
<td>grow</td>
<td>32.5</td>
</tr>
<tr>
<td>grass</td>
<td>10</td>
</tr>
<tr>
<td><strong>Intervocalic</strong></td>
<td></td>
</tr>
<tr>
<td>mirror</td>
<td>10</td>
</tr>
<tr>
<td>berry</td>
<td>12.5</td>
</tr>
<tr>
<td>correct</td>
<td>17.5</td>
</tr>
<tr>
<td>arrive</td>
<td>12.5</td>
</tr>
<tr>
<td><strong>Final</strong></td>
<td></td>
</tr>
<tr>
<td>dear</td>
<td>0</td>
</tr>
<tr>
<td>core</td>
<td>2.5</td>
</tr>
<tr>
<td>war</td>
<td>2.5</td>
</tr>
<tr>
<td>tire</td>
<td>2.5</td>
</tr>
</tbody>
</table>
Table I shows perception errors by five good native
Japanese speakers in distinguishing English /l/ and
/r/ on four native English speakers. We note that
/l/ is misperceived as /r/ in syllable-initial posi-
tion, but that /r/ is misperceived as /l/ in clus-
ters. Intervocally, the error rate is high and
possibly the direction depends upon stress. In final
position, errors are few. Taken together, this sug-
gests that different allophonic forms of /l/ and /r/
pose different problems.

2. Phonological Characteristics

McCawley (1968) claims that Japanese is a mora-
counting syllable language. This means that rules
such as accent insertion should assign an accent to
a syllable and rules such as pitch realization should
be applied to morae.

A mora can consist of a consonant plus a vowel
(CV), a vowel alone (V), mora nasal (N), or a first
obstruent of geminate consonants. A syllable con-
sists of a consonant plus a vowel (CV), a vowel alone
(V), a long vowel ((C)V:), a vowel plus a mora nasal
((C)VN), or a vowel plus a first obstruent of gemi-
inate consonants ((C)V(C)). For instance, kyooka 'sub-
ject' is a three mora word (kyoo-o-ka) as well as a
two syllable word (kyoo-ka). McCawley defines mora
as a "unit of phonological distance" and syllable as
a "prosodic unit."

The mora is a basic metrical unit. The Japanese
syllabary is based upon morae: each kana character
represents each mora. Thus, it is hard for the Japa-
nese without linguistic training to perceive  she as a
sequence of  and a. In other words, the Japanese
writing system is mora-based and the mora is the psy-
chological minimum unit of sound.

Japanese is also a language with only one pho-
nemic liquid, /r/, but that liquid varies in produc-
tion from a retroflex flap, [r], to a retroflexed
lateral, [1], to a tap [ ]. Its allophones are in
relatively free variation and within them it covers
the space of both English /l/ and /r/. Japanese has
only five phonemic vowels, but these undergo various
contextually predictable variations in coloring.

The Japanese speaker learning English must find
a strategy to distinguish /l/ and /r/, both in per-
ception and production. They differ in at least
three dimensions: (1) the articulatory position of
the consonant, per se, (2) the duration of the con-

sonant, and (3) the effect on the surrounding vowels. The first is, of course, a phonemic distinction, but we should note that /r/ is generally longer than /l/, and that vowels before /r/ are quite colored, and that even when /r/ is initial, following vowels tend to be slightly centralized (e.g., "light", "right").

The "correct" answer for production is to make a consistent distinction between the two consonants, but in the absence of that, it is still possible to produce consistently different minimal pairs which will resemble the English forms, merely by manipulating vowel color and consonant length. We have in fact been told by some native Japanese that they cannot distinguish /l/ and /r/ well in isolation, but perceive and produce them distinctly in context through these cues.

It could be further argued that this is a natural naive approach for a Japanese speaker who, taught to read in a mora-based system, has never analyzed initial consonant and vowel as separate units, but always as a single whole. Therefore, the new distinction between /l/ and /r/ likewise is not treated in isolation. Syllable final liquids, on the other hand, are in a separate mora from the vowel, so the Japanese listener will tend to isolate them. What we interpret as r-coloring on the vowel, the Japanese listener may interpret as a short vowel and longer consonant.

Why should the weight of the distinction be shifted away from the consonant? Many studies have shown that categorical perception is less marked for vowels than for consonants. Categorical perception (usually a synthetically generated one), acoustically marginal between two phonemes, will judge it to be one or the other, not some mixture of the two. This holds especially for consonant transition curves. There is, to the ears of a speaker of English, no halfway between the /t/ and /k/ phonemes.

But categorical perception does not hold to nearly the same degree for vowels. Listeners can tell that two tokens of the same vowel are not identical, whereas they frequently cannot with stops. They can tell that two vowel stimuli, of the same phoneme, differ in a particular dimension, that one is higher, or more central or that one is nearer the boundary of some other phoneme category.

To our knowledge, categorical perception in second languages has not been much investigated. Despite the absence of experimental evidence, it
seems reasonable to assume that speakers will tend to perceive consonants in a second language according to the categorical boundaries of their native language, but will, at an earlier stage, be aware of variations in vowel quality.

The consequence of this, for a native Japanese learning English, is that there may not be a clear boundary between /l/ and /r/, perceptible to the listener, but that on being taught that certain pairs are distinctive, the student will be able to find a distinction in the surrounding vowels, and in the duration of the consonants. It is not clear how fluent a speaker must become before a new boundary is located, and the perceptual burden is shifted to that dimension. Even a fluent non-native is likely to be influenced substantially by non-distinctive cues in the environment.

3. The Results

Let us note the general patterns of the errors in Table I:

1) In initial position there is a bias toward /r/.
2) In post-stop position, there is a bias toward /l/.
3) There are fewer errors on low vowels.
4) The errors are mixed but frequent in intervocalic position.
5) There are very few errors in final position.

These effects are not necessarily statistically significant, but let us ask why they would occur. What kind of perceptual strategy would cause these errors? We tentatively offer this set of perceptual principles for native Japanese listening to English:

6) /l/ is perhaps more marked, or for some unidentified reason the direction of error favors /r/ in initial position.
7) Final /r/ spreads forward onto the vowel, therefore the consonant is interpreted as longer.
8) When the liquid is intervocalic, coloring of the prior vowel is not noted as it is in a separate mora. The length difference noted in #2 is compressed.
9) Initial /r/ alters vowels somewhat, especially low vowels.
10) A short liquid is more likely to be /l/.

-- Postconsonantally, all liquids are shorter, causing a bias toward /l/.
The tendency to perceive short liquids as /l/ accounts for the results for clusters. Liquids in clusters will be shorter than those in isolation, so /r/ in a cluster may be mistaken for /l/. But a long initial /l/ may be mistaken for /r/.

There is some evidence that even native English speakers may be induced to misperceive /r/ as /l/ based on peripheral cues -- see for example Cutting and Day (1975), so it is not unreasonable for us to suppose that perception by Japanese may be swung one way or the other by such cues. Whether or not these specific principles are correct ones remains to be proven, but is not really crucial to a more general conclusion: it appears that the errors are systematic in direction, therefore that some set of cues, not phonemic in English, but usually predictable, is playing a role. The errors may occur especially when these secondary cues do not behave true to expected form.

4. Production vs. Perception

Is Sheldon justified in her conclusion that the Japanese in her study were better producers than perceivers? Not necessarily. It depends on what we mean by a "satisfactory" producer. The speakers are producing distinguishable /l/ and /r/ forms, but it is not clear that they are producing a good English /l/.

Her ultimate conclusion is that ESL teachers should perhaps pay special attention to perceptual skills, lest they discover that their apparently satisfactory students have a problem. We suspect, however, that any ESL teacher would quickly find that these students had not mastered the production of the l/r distinction even though, in a forced choice task, a sound might be recognizable as /l/ or /r/. It could very well be produced as an indeterminate liquid, surrounded by exaggerated contextual efforts. Only spectrographic or cineradiographic evidence would confirm this.

Assuming that native English speakers can make use of these contextual cues, there is no reason why they should not be able to distinguish "indeterminate" consonants. Their errors should occur only when the production of the consonant, as such, errs beyond indeterminacy toward the opposite end, that, when an /l/ is clearly /r/-like. Even if the tokens are produced by the wrong strategy, they will have to be
really bad before they are misunderstood. But the native Japanese, listening to English, will place greater weight on the surrounding context, and will misclassify if the context "errs," even if the consonant itself is perfectly well produced. The discrepancy between production and perception results does not show that the two are separate, that there is a strategy for production not available for perception.

Our moral, as phonologists, is that we simply do not know enough about the importance of context in classifying sounds. We assume that phonemes are distinctive on the basis of minimal distinctive features, and we tend to dismiss allophonic effects from the equation. There is good reason, even in a native language situation, to believe that this is an over-simplification, certainly in perception, quite likely in production as well. Sheldon's study suggests that in a second language situation, context may play an even greater role in perception and production.

Sheldon questions the traditional assumption that perception precedes production. We find the evidence for that inconclusive. However, we question an equally traditional assumption, that perception and production may be adequately described by neatly linear segmental models. We suggest that a better description of /r/ and /l/ must include features not usually considered distinctive, and not always confined to the temporal domain associated with the consonant.

References


(1983) "Nihongo no Akusento ni tsuite" (On the Accent of Japanese), Gengo (Language), Vol.12 No.1, 101-111