The English Pronunciation of a Japanese: Theoretical Linguistic and Sociolinguistic Perspectives

Ivan B. Brown

Joetsu University of Education

Abstract

This paper analyzes particular phonetic features, particularly final obstruent devoicing, of an extract of L2 English speech spoken by an L1 Japanese speaker. This is done from two theoretical perspectives, namely, generative phonology and optimality theory. The issues raised are then discussed from the sociolinguistic perspectives of the currently evolving status of English as an international language and the consequent question of what set of English phonological features would be an appropriate benchmark for analysis (and target for learners), as well as perspectives on phonological features of speech in social interaction.

Key Words:

generative phonology	optimality theory	sociolinguistics
English	Japanese	L2

1 - Introduction

The phonetics of L2 speech is an issue which has received the attention of second language educators, second language acquisition researchers and sociolinguists. The motive for this attention has often been to analyze the "errors" of second language learners in order to posit theories of acquisition, to evaluate L2 speakers' performance by comparing them with native speakers, or to ascertain L2 learners' needs in order to plan instruction to help them acquire native-like pronunciation. This paper's main focus is not on evaluating or criticizing these motives; rather, it explores two theoretical perspectives, namely generative phonology and optimality theory, in terms of how they can be employed to analyze the L2 English speech of an L1 Japanese speaker. It then critically discusses the necessity of identifying L2 speakers' "errors" with

reference to the currently evolving status of English as an international language. Firstly, this paper outlines the main phonetic features of English and Japanese, and then analyzes a dataset of L2 English spoken by a particular L1 speaker of Japanese.

2 - The Phonetic Inventory of English

Needless to say, there are hundreds of varieties of English pronunciation in many countries over the world. For the sake of simplicity of the overview, the basic features of British Pronunciation (RP) of English are given. Figure 1 shows cardinal vowel diagram for the pure vowels of RP, based on Crystal (2003), a work which can also be consulted for cardinal vowel diagrams for a variety of regional variations of each vowel, also citing notable differences which exist between RP and Standard American English (SAE), specifically for the vowels [ɒ] and [ɜ:]. The choice of symbol is related to lip spreading or rounding, and the location is related to the tongue position. The vowel [ɒ], as in the name *Bob*, has open lip rounding in RP but no lip rounding in SAE. The vowel [ɜ:], as in the word *her*, is lower in conservative RP, whereas it is slightly higher and shorter when followed by [r] in SAE (*ibid.*).

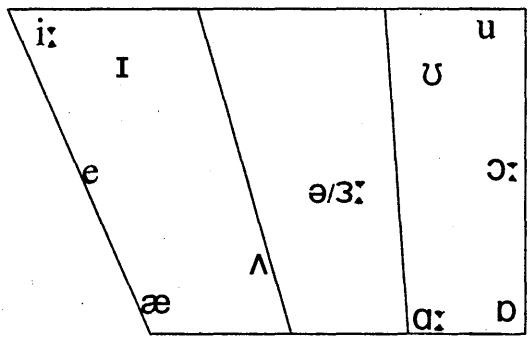

Figure 1: Cardinal vowel diagram for pure vowels in English (RP)

As well as pure vowels, English RP has a number of diphthongs, namely, [eɪ], [ɔɪ], [aʊ], [ɛə], [aɪ], [əʊ], [ɪə] and [ʊə]. These are also subject to significant regional variation, examples of which are also shown in detailed cardinal diagrams in Crystal (2003). For example, [aʊ], as in *house*, has a centralized first element in Canadian English, and [eɪ], as in *train* and *station*, has a lower first element in Cockney (East London) and broad Australian (*ibid.*). Table 1 shows the consonants of RP, again based on Crystal (2003). The main differences

between RP and SAE here are the post-vocalic rhotic [r] in SAE, as well as the flap [ɾ] variably replacing [t] in SAE (*ibid.*).

Table 1: English consonants

Manner		bi-labial	labio-dental	inter-dental	alveo-lar	Post-alveo lar	pala-tal	velar	uvu-lar	glot-tal
Stops:	[+V]	b			d					
	[-V]	p			t					
Fricatives:	[+V]		f	θ	z	ʒ				
	[-V]		v	ð	s	ʃ				h
Affricates:	[+V]					ʤ				
	[-V]					tʃ				
Tap/ Flap	[+V]				(ɾ)*					
Approxi-mants: liquids	[+V]				r l					
glides	[+V]	w					j			
Nasals:	[+V]	m			n			ŋ		

2 - Phonetic Inventory of Standard Japanese

2.1 - Japanese Vowels

The Japanese vowel phonemes are shown in Figure 2 in their Japanese positions in a cardinal vowel diagram, according to the descriptions given by Kaiser (1998) and Tsujimura (1996). The sound [a] is normally shown further to the left in a generic IPA chart (IPA, 2004). Tsujimura (1996:18) makes significant comparisons with English vowels. She states that: (1) the Japanese high front vowel [i] is equivalent to the English [iː] in *please*, except that the lips are not spread; (2) the Japanese mid-front vowel [ɛ] is slightly higher than in English *pet*; (3) the low mid vowel [ɑ] is about the same height as the first vowel in *father* but more forward; and (4) the mid-back vowel [ɔ] is similar to that in English *call*, but somewhat higher and slightly more front than in English;

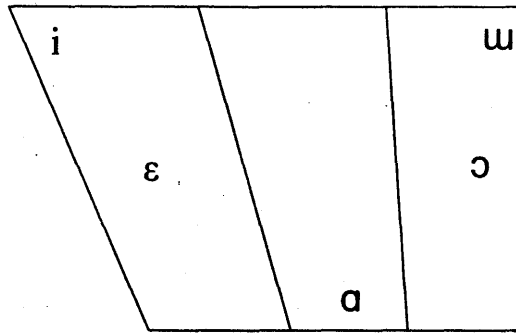

Figure 2: Japanese vowels

and (5) the Japanese high back vowel [ɯ] is unrounded (hence the symbol). She also notes, however, that this feature is more prominent in the Tokyo dialect, and less so in the dialects of western Japan, and also that Vance reported lip compression in very careful speech of Tokyo speakers (Shibatani, 1990; Vance, 1987, cf. Tsujimura, 1998:18). Each of these five vowels can occur as a short vowel or a long vowel contrastively.

The phonetic chart of Japanese consonants in Table 2 has been compiled based on that given by Tsujimura (1998:16). The post-alveolar consonants were given by Tsujimura under "alveo-palatal" articulations, using different symbols such as š and ž; that is, Fromkin and Rodman's transcription system (1973). These consonants and the alveolar affricates are given above by combining standard IPA symbols in Gimson's (1962) system. Tsujimura's table gives the alveolar liquid "(r)*" for the simplicity's sake, although she acknowledges that the Japanese phone is really a flap, as shown in Table 2. Nevertheless, the two English approximants /l/ and /r/ are well known for causing difficulties for Japanese learners because of the single Japanese flap [ɾ]. Other Japanese consonants are broadly similar to English consonants, apart from the bilabial fricative [Φ] and palatal fricative [ç], which occur where [h] does not occur in Japanese, before [ɯ] and [i] respectively. The nasal stops [n], [ň], [ɲ], [ŋ] and [ɴ] are all allophones (depending on nasal-assimilation) of the phoneme [n].

2.2 - Japanese Consonants

Table 2: Japanese consonants

Manner		bilabial	alveolar	Post-alveolar	palatal	velar	uvular	glottal
Stops:	[+V]	b	d					
	[-V]	p	t					
Fricatives:	[+V]		z					
	[-V]	Φ	s	ʃ	ç			h
Affricates:	[+V]		dz	ʤ				
	[-V]		ts	tʃ				
Tap/Flap	[+V]		ɾ*					
Approximants:								
liquids	[+V]		(r)*					
glides	[+V]				y	w		
Nasals:	[+V]	m	n	(ň)	ɲ	ŋ	N	

3 - The Japanese Speaker of English

3.1 – Speaker's background and the elicitation method

The following dataset of a Japanese native ("Japanese Speaker 3", JS3) speaking English was retrieved from the Speech Accent Archive, (Weinberger, 2004a). JS3 was a 49 year-old (at the time of recording) female born in Kofu, Yamanashi Prefecture in Japan, her native language was Japanese, German was listed as another second language, she had started learning English academically at the age of 12, and she had lived in the USA for three years. Like all the other speakers on the Speech Accent Archive, she had been recorded reading the following elicitation paragraph, which was then transcribed as shown in Figure 3.

> *Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.*

[pɹiz kɔːl stɛ̝la æsk hɑɹ tu̟ː bɹ̃ɪŋ
ðiz θĩŋ̟z wɪ̪s hɛːɹ fɹõm d̪ə stɔəɹ
siks spũnz̟ ʌbə fɹæʃ̟ə snŏ̟ piːs fəɪvə
θɪk slæ̝bz ʌvə βlɯː tʃis ænd mebi
ə snæk fo hɛːɹ bɹʌd̪ə bob wĭ ɔlso
nid ə smɑl plæstɪk sneɪk æ̟nd ə
big tɔɪ fɹɒg fŏʔ ðə kits ʃi kæn
skʊ̟p d̪is̪ ədiz̟ θĩŋ̟z ɪntŭ θɹi ɹɛd
bæ̝gs æn wi wɪl gɔː mit hɛɹ
wɛ̃nsd̪e æt d̪ə tɹɛ̃ːn steʃʊn]

Figure 3: IPA Transcription of JS3 (Weinberger, 2004b)

4 - Generalizations made by JS3

According to Weinberger (2004b), JS3 made the following phonological generalizations, as shown in Table 3.

Table 3: Generalizations for Speaker 3

segmental generalizations	syllable structure generalizations
consonants • final obstruent devoicing • interdental fricative --> stop vowels • vowel shortening • vowel raising	• vowel insertion • r-deletion

4.1 - Final obstruent devoicing

An example of final obstruent devoicing is JS3's rendition of the obstruent [z] in *cheese* ([tʃis] instead of [tʃiːz]) and the double obstruent [dz] in *kids* ([kits] instead of [kIdz]). However, JS3 voices other final obstruents such as *spoons*. Potential causes of this generalization include her native language, Japanese, and universal constraints. Some languages have only unvoiced word-final obstruents, such as German, which happens to be another second language of this speaker. Japanese has voiced and unvoiced obstruents, but has no word-final obstruents at

all (Tsujimura, 1996). The only type of word-final coda it allows is an alveolar nasal stop. It does allow word-internal unvoiced obstruent codas as part of a geminate. This error could be a result of JS3 transferring features of Japanese into her English. She could be transferring the Japanese geminate rules into word-final rules for her English, or she could be transferring the word-final obstruent neutralization properties of German into English by over-generalization. These may be partially successful attempts at producing word-final obstruents.

4.2 - Interdental fricatives replaced by dentalized alveolar stops or fricatives

JS3 is also apparently inconsistent with this error. She produces a dentalized alveolar fricative [s] instead of the intended interdental fricative in *with*, pronounces *brother* and two instances of *the* with a dentalized alveolar stop, and pronounces other instances of interdental fricatives correctly, as in *these things*. There are no interdental consonants at all in Japanese. These are therefore instances of the speaker failing to achieve the target place of articulation, and producing the nearest alternatives in Japanese. The Japanese alveolar stops and fricatives are somewhat dentalized in any case (Tsujimura, 1996).

4.3 - Vowel raising

As shown earlier, referring to Kaiser (1998) and Tsujimura (1996), Japanese has only five vowels, two of which are high, namely, [i] and [u]. The mid-font vowel [e] is slightly higher than in English *pet*. English has a similar front high vowel [i], a high back vowel [u], but also a slightly lower central front high [ɪ]. The speaker raises the vowels in *Stella*, *fresh*, *red*, *Wednesday*, *things*, *six*, *big* and *kids*. This could be an instance of JS3 transferring the Japanese vowel heights into her spoken English. Alternatively, one could argue in some cases that the heights of the vowels are being assimilated with the heights of adjacent consonants.

4.4 - Vowel insertion

JS3 inserts schwas at the ends of the words *of*, *fresh* and *five*. This can be explained in terms of Japanese syllable structure. The only word-final coda that occurs in Japanese is the alveolar nasal stop [n] (Tsujimura, 1996). It is possible that JS3 inserts these vowels in order to comply with Japanese syllable structure and make the English syllables easier to pronounce.

5 - Analyses of Final Obstruent Devoicing in the Dataset

5.1 - Generative Phonology

In generative phonology, the feature matrices of the segments in question are examined. The feature value for obstruents is [-sonorant] (Clark and Yallop, 1995). The environment for word-final position could be denoted as: __##. Therefore a rule can be written to stipulate that obstruents are voiceless when in word-final position:

[-sonorant] → [-voiced]/__##

However, not all the word-final obstruents in JS3's dataset are voiceless. Those in *cheese*, *kids*, *peas*, *these* in *She can scoop these* and *bags* are voiceless by over-generalization but she successfully voices the word-final obstruents in "*please*", the earlier occurrences of *these*, *things*, *spoons*, *slabs*, *Bob*, *need*, *and*, *big*, *frog* and *red*. This seems to indicate that either the phonological system of this speaker's interlanguage is in an unstable fluctuating state, or that it is a complex phonotactical system, where the voicing of word-final obstruents depends on the environment. A generative theory would search for abstract underlying forms (Lass, 1984), from which one or more transformational phonological rules operating in a particular order would produce the surface phonological forms such as those above.

5.2 - Optimality Theory

In optimality theory, one would search for an explanation to this problem by referring to a number of constraints, and placing them in a ranking order so as to produce the desired set of outputs and point to an explanation (Archangeli, 1997). Two constraints that could be relevant here are: (1) Obs/Voi: An obstruent must be voiceless, and (2) Faith[Voice]: Voicing in the output must be faithful to the voicing in the input (Pulleyblank, 1997). The ranking of these determines whether a phonological system can produce voiced obstruents or not. If (1) is ranked higher than (2), i.e., Obs/Voi >> Faith[Voice], then it is more important that obstruents be voiceless than that the voicing is faithful. However, this problem concerns word-final obstruents. Earlier it was pointed out that Japanese has voiced obstruents but not in word-final or syllable-final positions. Hammond (1997:41) drew attention to the argument for the syllable as a unit of cognitive organization, and claimed: "All polysyllabic words can be analyzed as sequences of well-formed syllables." On this basis, one can develop a constraint hierarchy for the Japanese syllable, which might point toward a possible hierarchy for

JS3's interlanguage. NoCoda: Syllables end with a vowel (Archangeli, 1997, p7) and Faith[Nasal] (Pulleyblank, 1997, p73) could be combined to account for Japanese syllables with respect to codas in syllable structure:

Faith[Nasal] >> NoCoda

The occurrences of neutralization of final obstruent voicing in JS3's interlanguage with respect to codas appear to be similar to that of Russian, as examined by Pulleyblank (1997). He ranks the constraints ContrastiveCoda (a coda does not bare contrastive features), Faith[Voice] and Obs/Voi to account for this:

ContrastiveCoda >> Faith[Voice] >> Obs/Voi

This can be applied to the word *cheese* from the dataset, as shown in Table 4.

Table 4: Constraints for *cheese*

[tʃIz] / "cheese"	ContrastiveCoda	Faith[Voice]	Obs/Voi
[tʃIz]	*!		*
☞ [tʃIs]		*	

Models for general final obstruent devoicing in this dataset have been offered using classical generative phonology and optimality theory. In this instance, the classical phonological rule specifies word-final obstruents rather than generalizing for the codas of all syllables. The ranking of constraints given through optimality theory generalizes for all codas. It has not been possible, on this occasion, to devise models using either theory that account for the whole dataset. There are two possible explanations for this. Either the phonological system of the speaker's interlanguage is highly complex, with many interacting rules or constraints that make the apparent irregularities regular, or the inconsistencies are due to paralinguistic or non-linguistic factors in the given performance. At this juncture, it may be apposite to view this dataset and the issues it raises from a sociolinguistic perspective.

5.3 - Sociolinguistic Perspectives

It is not necessarily in the intrinsic natures of the theoretical perspectives employed above to analyze L2 speech by measuring their deficits as compared to native speakers' English. In any case, as the overview of the inventory of English phonemes indicates, it is unclear which native speaker variety should be used as a benchmark. This is an ever increasingly pertinent issue with respect to English as a target language, given its evolving role as an international language, and is attracting the attention of a growing number of researchers of sociolinguistics. Jenkins (2002), for example, highlighted the recent consensus that "non-native speakers (NNSs) using English for international communication now outnumber its native speakers," and set out to address the "need for empirically established phonological norms and classroom pronunciation models for English as an International Language (EIL), in which intelligibility for NNS rather than for native speaker (NS) receivers is the primary motivation" (p.83). Through analysis of three sets of data drawn from NNS-NNS interaction, she arrived at a proposal for a revised pronunciation syllabus for EIL, the *Lingua Franca Core*, containing phonemes that all speakers should aim to achieve in the interests of mutual intelligibility. Other non-essential features, in which variation would not seriously hinder intelligibility, were listed under "Non-core features".

Of relevance to the dataset in this paper are Jenkins' categorization of the interdental fricatives [θ] and [ð], minor differences in vowel quality as non-core features, in which case the analysis of JS3's transformation of these sounds into a dentalized alveolar fricative or a dentalized alveolar stop, and her vowel raising may have value for linguistic research but would not lead to any negative evaluation of JS3's English speaking ability. However, her syllabus, at least this version of it, does not deal with final obstruent devoicing or vowel insertion. Nevertheless, inspection of her data seems to show interactions between NNSs with different L1s in which they frequently devoice final obstruents with no apparent effect on intelligibility. Therefore, it could be speculated that word-final voiced obstruents would be categorized as non-core according to Jenkins' criteria.

In addition to the EIL paradigm, other types of sociolinguistics can also shed light on the analysis of L2 speech. The datasets on the Speech Accent Archive may be useful for the heuristic search for features of spoken data from various regions, but the fact that the recordings are obtained through the reading of an elicitation paragraph means that we cannot investigate these L2 phonological features in the context of authentic communication. Local (2004) pointed out that "The natural home of spoken language is social interaction and linguistic (phonetic) resources are systematically deployed in its management" and went

on to argue that "particular phonetic details and phonetic variability are associated with particular interactional, grammatical and lexical systems and that this 'context-embeddedness' is used in speech understanding" (2004:abstract). Applying this perspective to JS3's speech, it might be more fruitful to analyze her varied devoicing of final obstruents, for example, by observing them in authentic contextualized interactions, making it possible to investigate whether this variation was linked to social aspects of her interaction. Indeed, some L2 speaker's phonological features which have previously been labelled as "errors" are recently being seen in a more positive light as proactive communication strategies. For example, Carroll (2004) used conversation analysis to show that Japanese EFL learners' insertions of vowels (especially those with rising intonation) after word-final consonants were not merely phonological errors that impeded communication, but a time-buying strategy to indicate to their interactional partners that further words would follow after a brief pause to think. From an even more socially-oriented perspective, phonological variation has been shown to be related to communication accommodation (Giles and Coupland, 1991). For example, if JS3 were communicating with an interlocutor who tended to devoice final obstruents, and if JS3 felt inclined to identify with this interlocutor, then she might *converge* towards the interlocutor's devoicing of final obstruents. Conversely, if JS3 felt disinclined to identify with the interlocutor, she might *diverge* by tending to voice such word-final obstruents.

6 - Conclusion

This paper has examined the phonological dataset of a Japanese speaker of English from the perspectives of comparisons of the phonemic inventories of English and Japanese, generative phonology and optimality theory and the macro-sociolinguistic perspective on English as an international language and the micro-sociolinguistic perspective of social interaction. Each perspective arguably has its own advantages to contribute to a comprehensive picture, and their differences suggest that none of them should be exclusively relied upon.

References

Archangeli, D. (1997) Optimality theory: An introduction to linguistics in the 1990s. In D. Archangeli & D.T. Langendoen (eds.), *Optimality Theory: An Overview.* Oxford: Blackwell Publishers, pp.1-32.

Carroll, D. (2004) Vowel-marking as an interactional resource in Japanese novice ESL conversation. In K. Richards and P. Seedhouse (eds.), *Applying Conversation Analysis*. Basingstoke, England: Palgrave Macmillan.

Clark, J. & Yallop, C. (1995) *An Introduction to Phonetics and Phonology. 2nd Ed.* Oxford: Blackwell Publishing.

Crystal, D. (2003) *The Cambridge Encyclopedia of the English Language. 2nd Ed.* Cambridge: Cambridge University Press.

Fromkin, V. & Rodman, R. (1972) *An Introduction to Language.* 5th ed. 1994. New York: Holt, Rinehart & Winston.

Giles, H. and Coupland, N. (1991) *Language: Contexts and Consequences.* Pacific Grove, CA: Brooks/Cole.

Gimson, A.C. (1962) *An Introduction to the Pronunciation of English.* 6th ed. 2001. London: Edward Arnold.

Hammond, M. (1997) Optimality theory and prosody. In D. Archangeli & D.T. Langendoen (eds.), *Optimality Theory: An Overview.* Malden, MA: Blackwell Publishers, pp.33-58.

International Phonetic Association (2004) *Reproduction of The International Phonetic Alphabet (Revised to 1993, Updated 1996)* c/o Department of Linguistics, University of Victoria, Victoria, British Columbia, Canada. Retrieved from URL <http://www2.arts.gla.ac.uk/IPA/fullchart.html> on November 1st, 2004.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language, *Applied Linguistics* 23(1): 83-103.

Kaiser, S. (1998) Japanese Language II. In *Japanese Correspondence Course for JET Participants 1998-1999, Book 3.* Tokyo: Council of Local Authorities for International Relations.

Lass, R. (1984) *Phonology: An Introduction to Basic Concepts.* Cambridge: Cambridge University Press.

Local, J. (2004). Cognition, intersubjectivity and phonetics in talk-in-interaction. Paper presented at *The Architecture of the Language Faculty*, the inaugural conference of the Centre for Human Communication, University College London, on June 4th, 2004.

Pulleyblank, D. (1997) Optimality theory and features. In D. Archangeli & D.T. Langendoen (Eds.), *Optimality Theory: An Overview.* Malden, MA: Blackwell Publishers, pp.59-101.

Shibatani, M. (1990) *The Languages of Japan.* Cambridge: Cambridge University Press.

Tsujimura, N. (1996) *An Introduction to Japanese Linguistics.* Oxford: Blackwell Publishing.

Vance, T.J. (1987) *An Introduction to Japanese Phonology.* New York: SUNY Press.

Weinberger, S. (2008a). *Speech Accent Archive.* George Mason University. Retrieved from URL: <http://accent.gmu.edu/> on October 1st, 2008.

Weinberger, S. (2008b). *Japanese Speaker 3, Speech Accent Archive.* George Mason University. Retrieved from URL: <http://accent.gmu.edu/browse_language.php?function=detail&speakerid=223> on October 1st, 2008.