

Realizations of Consonant Clusters in L2 Phonology of Japanese Learners of English

Yoshiho Shibuya

The present paper is a proposal part (Part 1) of research on the topic stated above. The actual procedure for the whole research includes two experiments -- (1) an acoustic experiment and (2) an articulatory experiment. The experiments and their results will be discussed and conclusion will be presented in the next *Kiyou* (#32, 2000).

1. Introduction – The Problem

More and more studies have been conducted recently on L2 phonology, with the understanding that careful investigation on L2 acquisition may help us understand more fully the cognitive processes of language learning or the nature of language itself. There is a growing awareness that not only L1 transfer, which has often been argued to be one of the main factors that affects L2 phonology (Broselow, 1984; Hancin-Bhatt, 1994b; Sato, 1983; Tarone, 1980), but also universal processes or constraints on linguistic complexity (Major, 1986; Eckman and Iverson, 1993; James, 1993; Hancin-Bhatt and Bhatt, 1997, p.342) are important factors affecting L2 phonology. Several attempts have been made to demonstrate how these universal constraints and L1 transfer interact with each other, with the focus on the strong explanatory potential of phonological theory and L2 acquisition theories (Epstein, Flynn, & Martohardjono, 1996; Sato, 1983; Tarone, 1980; Broselow and Finer, 1991; Eckman and Iverson, 1993; Hancin-Bhatt and Bhatt, 1997; among others). While these studies provide us with insights on this important issue, many of them seem to focus merely on the phonological aspect and little attention has been paid to the possibility that phonetics, either universal or language-specific, may influence the process of language acquisition.

Hancin-Bhatt and Bhatt (1997), for example, conducted experiments investigating the interactions of transfer and developmental effects (due to universal processes or constraints) in the acquisition of consonant clusters in syllable-onsets. Their study extended Broselow and Finer's Minimal Sonority Distance model in L2 syllables, providing evidence to support the idea that transfer effects play an important role in L2 syllable structure and that they do interact with developmental effects. They also argue that using an OT (Optimality Theory) approach provides a better hypothesis than previous studies on the nature of these effects. In their study, they followed Broselow and Finer's (1991) Minimal Sonority Distance Parameter-Setting model. Focusing on

the acquisition of consonant clusters in syllable onsets, Broselow and Finer (1991) argued that in the process of acquisition, both markedness (defined in terms of a set of parameters and parameter settings given by Universal Grammar (UG), p. 35) and transfer (defined as the carrying over of the L1 parameter setting into the L2, p.35) are involved, and that at a certain stage of acquisition learners arrive at a parameter setting that is midway between their L1 and L2 settings. In their study, they introduced the minimal sonority distance (MSD) parameter. Following Wexler and Manzini (1987), they assumed that the settings associated with these parameters are in subset relations (in that the more marked setting includes the constructions permitted by less marked one). They conducted an experiment to test the ability of learners to acquire L2 consonant clusters that differed in degrees of markedness. They investigated whether the minimal sonority distance parameter is set at a new value during the syllable structure acquisition of English. The sonority hierarchy can be shown as in (1) below:

- (1) least sonorous <-----> most sonorous
 stops – fricatives – nasals – liquids – glides – vowels¹
- (2) stops – 1; fricatives – 2; nasals – 3; liquids – 4; glides – 5

For example, if a language has a sonority distance setting of 5, no consonant clusters are allowed. If a language has a setting of 3, then any cluster with at least a difference of 3 is allowed. They argue that the need to keep the degree of sonority difference in the speakers' native language (L1) is also respected in L2, resulting in cross-language transfer (See Selkirk, 1982; Broselow and Finer, 1991 for detail). From the cross-language transfer point of view, MSD model predicts onset sequences with a lower MSD setting than the learner's L1 to be more difficult to acquire than the other way around. However, from the point of view of developmental effects', if the MSD of the two consonants is low, the acquisition of the cluster will be more difficult due to the universal constraints on sonority sequencing (i.e., in complex onsets, larger differences in sonority are more preferred than smaller differences). Thus, according to Broselow and Finer's MSD model, both L1 sequencing possibilities and universal sequencing constraints affect the L2 speakers' acquisition of consonant clusters in syllable onsets.

Hancin-Bhatt and Bhatt (1997) tested the Broselow and Finer's MSD model more carefully by examining Spanish speakers whose L1 has some complex onsets (thus some transfer effects should be seen) as well as Japanese speakers, whose L1 has virtually no complex onsets (only developmental markedness effects should be observed). They also looked at codas as well as onsets (to see if the MSD model can be extended to codas). They conducted an experiment on 10 speakers of Spanish and 10 speakers of Japanese using pseudowords with both complex onsets and

¹ I followed the sonority hierarchy scale in Young-Scholten (1994) here.

codas to test the predictions of the MSD model and to examine more carefully the interaction of transfer and developmental effects in L2 learners' complex syllables. The results of their study strongly suggest that transfer interacts with developmental errors. They then argue that in order to account for the systematic error types they found, another, new, grammatical framework, namely OT, is necessary, since MSD model does not make predictions on language-specific phonotactics nor systematic errors whereas OT does. Their study seems to provide evidence for the interaction of the effects of both L1 transfer and developmental processes. What seems to be lacking, however, is strong phonetic evidence to support their findings since their study relies solely on the narrow phonetic transcriptions of the recordings. Neither acoustic nor articulatory data were carefully examined in their study. Not only this study, but also most of the other previous studies on L2 phonology seem to share these shortcomings of no acoustic or articulatory examination of phonetic data. Without such examination, we know neither whether phonetics is actually involved in the process, nor can we tell to what extent it is involved.

Zsiga (1995) discusses the important role that phonetics may play in L2 phonology. She argues that we need to consider the following three types of rule, or sound pattern: lexical (Kiparsky, 1985) alternations (rules that apply within words), postlexical processes (rules that apply to words in phrases), and phonetic implementation or regularities (Keating, 1988; Cohn, 1990) (language-specific patterns in the implementation of speech (Zsiga, 1995, p. 576)). Zsiga points out that in many cases lexical and postlexical rules behave differently (e.g., postlexical, but not lexical, rules transfer to L2 (Weinberger, 1994)) while postlexical rules and phonetic processes often seem to behave in a similar way. She argues, following Browman and Goldstein's *Articulatory Phonology* (1986, 1990, 1992), that "many if not all postlexical 'rules' can be seen to result from habits of articulatory coordination (p. 577)" and therefore should be described in terms of articulatory timing rather than feature-changing rules as has been done in previous studies. The basis for her argument comes from many studies done on gestural overlap in postlexical processes (e.g. Hardcastle, 1985; Barry, 1985; Browman and Goldstein, 1990; Byrd, 1994; Zsiga 1994.) They have found that there is significant overlap when two consonants are next to each other. That is, the gesture of the articulators of the second consonant often begins before or at the same time as the closure of the first one, resulting in concealing the acoustic cues. Zsiga (1995) gives many examples and concludes that it is gestural overlap (Browman and Goldstein, 1986, 1990, 1992) that lies behind postlexical processes. In other words, what was thought to be the result of phonology may turn out to be the result of timing patterns.

This kind of approach, I think, is missing from the studies done on L2 phonology I have discussed earlier (Sato, 1983; Tarone, 1980; Broselow and Finer, 1991; Eckman and Iverson, 1993; Hancin-Bhatt and Bhatt, 1997). I believe that acoustic and articulatory analysis of relevant data

is definitely called for before we further discuss the role of transfer. Relying solely on (narrow) phonetic transcription (as in the case of Hancin-Bhatt and Bhatt, 1997, for example) may not be enough considering the influence of gestural overlap. We need acoustic and articulatory analysis to carefully examine data because they can provide us with rich information that human ears often cannot capture. I believe that only after the role of phonetics is more clearly determined through acoustic and articulatory analysis, can we really discuss L1 transfer and the developmental effects in L2 phonology. The aim of this dissertation is exactly to do that. Here, I will explore the possibility of the phonetics/phonology interface regarding the realization of L2 syllables through both acoustic and articulatory experiments. As Zsiga points out (1995), what has been treated as the result of phonological rules or constraints may well be the result of the differences in timing patterns in L1 and L2. By conducting the acoustic and articulatory experiments and by analyzing the data carefully, I think that we can discuss the role of phonetics and phonology interacting in L2 acquisition, as well as the role of transfer. Specifically, I will investigate how English consonant clusters are produced by Japanese learners of English.

Studying the realization of consonant clusters is interesting and worth investigating because it offers insights to L2 syllable structures. The language group that I choose to examine is Japanese learners of English because Japanese is a language which does not have complex onsets or codas (Vance 1987) while English allows both complex onsets and codas. Two different kinds of strategies (deletion and vowel insertion) are expected to be observed in the realization of English syllables by the Japanese learners, though at different degrees and in different environments (Hancin-Bhatt and Bhatt 1997, Sato 1983, Saunders 1987, Weinberger 1994, 1997). In my study, I will collect data from Japanese speakers of English as a foreign language speaking sentences with target words containing consonant clusters, both as onsets and as codas. The specific questions I will be addressing are:

(1) *How do Japanese learners of English realize English syllables?*

If we assume L1 transfer, both deletion and vowel insertion (epenthesis) are possible ways for Japanese learners of English in realizing English syllables with consonant clusters. This is exactly what previous studies have demonstrated so far. In the present study, I will conduct acoustic analysis of data to confirm/disconfirm the findings of the previous researches. The reason why acoustic analysis of data is necessary is because what human ears “hear” may not be the same as what is actually produced. Also, and more importantly, through the acoustic analysis I want to investigate the nature of the vowel inserted between the consonants within consonant clusters, in order to determine whether it is inserted phonetically (an excremental vowel, inserted for the ease of articulation) or phonologically (an epenthetic vowel, inserted phonologically). Since the inserted vowel would be very short, human ears are unlikely to hear the difference; in other words, only acoustic

analysis can determine the nature of the inserted vowel. This is explained more fully in (2).

(2) *If insertion takes place, is the vowel inserted phonetically or phonologically?*

According to Levin (1989), there are two kinds of vowel insertion processes: epenthesis (phonological insertion) and excrescence (phonetic insertion). If a vowel is inserted for the ease of articulation, it is excrescence. The excrescent vowel is expected to show gradient, continuously varying duration (Levin 1989). Levin (1989) examined the Piro language, a member of the Arawakan family, and Hua, a language of the Eastern New Guinea Highlands, and found that these languages exhibit the existence of excrescent vowels, vowels that are clearly different from underlying vowels of the language in several ways. The summarized properties of excrescent vowels by Levin are: (1) they play no role in lexical or post-lexical rules; (2) they are significantly shorter than regular vowels; (3) the tongue position for them is conditioned by surrounding consonants (that is, they act as a mediator for adjacent articulations that require some degree of constriction). Because of these properties, excrescent vowels are expected to show gradient, varying duration as well as quality, depending on the surrounding consonants (Levin, 1989). If one of the properties of the excrescent vowels is mediating the articulations of adjacent consonants, they may be schwa-like and they may be viewed as what Browman and Goldstein (1992b) call “targetless” schwa, which can be described as reduced schwa. “Targetless” schwa is characterized as having no specified tongue position of its own (hence, targetless). This will be discussed later.

On the other hand, if epenthesis takes place, the epenthesized vowel is expected to show a categorial difference (i.e., either there is a vowel or no vowel). If there is an epenthetic vowel, I expect it to be longer and clearer than the excrescent vowel since it is inserted phonologically. That is, the vowel should have the characteristics of a full vowel that is in the inventory of the language. In other words, the epenthesized vowel is more likely to have clearer formants than the excrescent vowel. In Japanese speakers’ production of consonant clusters, I assume that the inserted vowel to be either /u/ or /o/, if not excrescent (in that case, it will be schwa-like).² Since /o/ appears only after /t/ and /d/ and /u/ elsewhere in Japanese loanword phonology, determining the nature of the vowels that are inserted will be crucial to my study. If the vowel after /t/ or /d/ are /o/ and the vowel after other consonants are /u/, then it is an epenthetic vowel, a vowel inserted as a result of phonological processes, not an excrescent vowel.

² An epenthetic vowel cannot be schwa since it is not included in the inventory of vowels in Japanese (Vance, 1987; Tsujimura, 1995).

Although acoustic analysis will provide more information than mere transcription of the recordings, it may not be suffice because, as some previous studies) have shown (Byrd 1992, 1996, Byrd and Zsiga 1990, Nolan 1989) the effect of gestural overlap (cf. Browman and Goldstein 1992) also needs to be considered. For example, what seems like deletion in the acoustic data may actually be the result of gestural overlap, i.e., the gesture of the second consonant overlapping (covering) with that of the first in consonant clusters.

2. Phonological Issues

2.1 Japanese and English syllable structures

The current study focuses on the issue of IL phonology, more specifically, the realization of consonant clusters by the Japanese learners of English. It is therefore our first task to look into the syllable structures of the two relevant languages, Japanese and English.

Japanese is a language with fairly simple syllable structure, one which does not have complex onsets or codas (Vance 1987). I will follow Poser (1990) on the inventory of Japanese syllables.³

(1) Japanese syllables⁴

Type	Example	Gloss
CV	to	'door'
CV:	too	'tower'
CVG	toi	'water pipe'
CVN	toN	'ton'
CVQ	totta	'took'
CV:N	tyeeN	'chain'
CVGQ	gendaikko	'modern youth'
CGV:Q ⁵	tookyookko	'Tokyoite'
CVNQ:	rondoNkko	'Londoner'

English, on the other hand, is quite free in its syllable structures (Kenstowicz 1994), allowing both complex onsets and codas. Somewhat simplified versions of syllable structure for the two languages are illustrated below:⁶

³ The inventory in (1) follows Poser (1990), but is not the same. Whether *i* in *oi* or *ai* is a real glide or another full vowel is still controversial. Poser suggests that it is a glide. I will not discuss this issue here.

⁴ C = consonant, V = vowel, : = vowel length, G = glide, N = moraic (syllabic) nasal, Q = the first half of geminate

⁵ Whether Japanese has an onset consonant cluster like this (CG) with a glide [j] following a consonant is controversial, but I will not discuss this here.

⁶ What is in the parenthesis () is optional. The small numbers indicate the possible number of consonants.

(2) Japanese syllable structure

$$\sigma = (C) V (V/N/Q)$$

(3) English syllable structure

$$\sigma = (C^{0-3}) V (V/C^{0-3})$$

2.2 Epenthesis and deletion

Since Japanese, characterized as a CV language, is much more restricted than English in its syllable inventory (i.e., Japanese inventory is a subset of English possibilities), it has been argued that L1 transfer is the most prominent source of syllable structure errors that Japanese learners of English make (Hancin-Bhatt and Bhatt 1997, Sato 1983, Saunders 1987, Weinberger 1994, 1997).

One of the earliest study on syllable structure errors is conducted by Oller (1974, as cited in Saunders 1987), who suggested that the preferred strategy that L2 learners employ was the use of epenthetic vowels. Oller demonstrated that L2 learners epenthesize vowels in order to cope with the difficulties in cluster production while L1 learners tend to use deletion. Since Oller, numerous attempts have been made by scholars on the issue of syllable structures of L1 and L2 and how that affect the production of consonant clusters in L2 (transfer) (Hancin-Bhatt and Bhatt 1997, Ioup 1984, Ito 1987, Sato 1987, Saunders 1987, Tarone 1976, 1980, Levin 1987, Weinberger 1990 among others), showing that both epenthesis and deletion are common strategies employed by L2 learners and L1 transfer does affect L2 phonology acquisition greatly. In the early studies where the focus was merely on the differences between L1 and L2, researchers discussed this issue within the framework of Contrastive Analysis Hypothesis and attributed all the errors to L2 transfer (see Broselow 1987 for review).

More recently, however, the main focus has been shifted to conducting cross-linguistic work and to identifying the effect of universal developmental processes or constraints (Eckman 1987, 1991, Hancin-Bhatt and Bhatt 1997, Ito 1987, Sato 1987 among others). Before I go into more detail on this issue, I will briefly introduce Japanese loanword phonology as background, in the following section.

2.3 Japanese loanword phonology

As stated earlier, one of the main purposes of the present study is to identify the nature of the vowels that are inserted in L2 production of consonant clusters. More specifically, I will examine the vowels inserted between consonants in consonant clusters to identify whether they are inserted lexically or post-lexically. It is therefore relevant to examine loanword phonology of Japanese since it exhibits the case of epenthesis. It is well-known that languages employ different strategies such as deleting one or more of the sounds or inserting sounds in their loanword phonology when the source language has a syllable structure that is not identical to its own. Languages differ from one another not only on deciding which strategy to use (deletion or epenthesis) but also on

deciding which sounds to delete or epenthesize. Japanese is a language that does not allow complex onsets or codas, which employs vowel epenthesis as its repair strategy (Ito 1989, Katayama 1998, Ohtake 1996). It is generally assumed that unrounded high back vowel /u/ is the unmarked, default vowel that is inserted in Japanese, whereas /o/ is the marked vowel that appears only after coronals /t/ or /d/ (Katayama 1998, Lovins 1975, among others). This is illustrated in the following:

(4) English loanword in Japanese

clubs /klʌbz/ → kurabuzu /kurabuzu/; strike /straɪk/ → sutoraiiku /sutoraiiku/

In (4), /o/ is inserted after /t/ while the default vowel /u/ is inserted elsewhere (after /s/ and /k/ in this case). I, therefore, argue in the present study that if the vowel found to be inserted in the L2 production of consonant clusters is identified as either /u/ or /o/ and their distribution is identical to that of loanword phonology, phonology is clearly involved in the process (i.e., epenthesis).

3. L2 Phonology

3.1 L2 Phonology

L2 acquisition had long been neglected or underestimated except in the area of language pedagogy because traditionally what it could offer was thought to be (at least by some scholars) limited to classroom teaching alone (Eckman and Iverson, 1994; Newmeyer & Weinberger, 1988 as cited in Epstein, Flynn, & Martohardjono 1996; Yava, 1994). Recently, however, more and more scholars have become aware that L2 acquisition processes can and do offer deep insights to our understanding of the nature of language as well as language learning in general. A growing number of research has been conducted, mainly in the field of syntax, but in the field of phonology as well.

In the past L2 phonology had been paid only little attention compared to many other areas in linguistics. As mentioned briefly above, however, recently a growing number of studies have been conducted on the issue of L2 phonology because the field is thought to have the explanatory potential of phonological theory and linguistic learning theories in the area of phonological acquisition (James, 1994; p. 189). It had been revealed that, although transfer is still considered the main factor affecting L2 phonology (Broselow, 1984; Hancin-Bhatt, 1994b; Sato, 1983; Tarone, 1980; Young-Scholten, 1994; among others), the view that something more than transfer is also being involved in the process of L2 acquisition by adult learners is also considered extensively.

Wexler and Manzini (1987) (and Berwick, 1985), assuming that only 'positive evidence' is available to the learners, defines Subset Principle as "a method for specifying a markedness hierarchy when alternative values yield languages which are in a subset relation." The less marked setting is included in the more marked ones. All that is generated in the smaller circle as well as additional forms can be generated in the bigger circle (i.e., it is less restricted). They assume

that, in L1 acquisition, initially, a parameter is in the most restrictive setting (the small circle, i.e., the subset), and that only exposure to positive evidence triggers a parameter resetting from a less marked to a more marked one. Wexler and Manzini focused their study on parameters and learnability in syntax (binding theory), but Broselow and Finer (1991) extended this to phonological acquisition as well as syntax in L2 (as introduced in the previous section). Assuming Sonority Sequencing Generalization (SSG) (segments within a syllable tend to be arranged in order of decreasing sonority approaching the syllable margins), they argued that clusters in which consonants are closer in sonority are more marked than clusters in which consonants are farther apart on the sonority scale. The (universal) SSG⁷ is combined with the (language particular) setting of the MSD and they together constrain the structure of consonants. For example, Japanese, a language which only allows clusters that consist of two consonants that are very far apart (obstruent and glide) is very restrictive and unmarked, while English, which has more marked setting, allows more complex onsets than Japanese. In their study, Broselow and Finer investigated the relevance of the MSD parameter to L2 acquisition. They conducted an experiment on 24 Korean and 8 Japanese speakers producing words with different complex onsets, to examine whether the more marked clusters (clusters which are closer in sonority) are more difficult to learn than the less marked ones. The results of their study suggest that learners seem to converge on a setting of the parameters that is somewhere midway between the L1 and L2 settings. In other words, they seem to support the assumption that learners begin with L1 setting and move toward L2 setting. They conclude that principles of UG constrain the L2 acquisition and L2 learners do have access to UG, but that adult learners appear to transfer their L1 parameter settings regardless of whether their L1 setting is more or less marked than the L2 setting (p. 55). They suggest that, although more marked features may be harder to learn, less marked features are not necessarily easier and that we need to consider both markedness and transfer in L2 acquisition.

Young-Scholten (1994) discusses the issue of ultimate attainment in L2 phonology in terms of whether or not post puberty learners retain access to UG and concludes that adult learners do have access to UG based on Subset Principle (Wexler and Manzini 1987) and Asymmetry Hypothesis (Young-Scholten 1994).

By first reviewing previous studies on L2 phonology, she discusses the existence of interlanguage (IL) phonology, which have often been shown to be neither L2-like nor L1-like, but to be systematic and to contain the kinds of rules found in natural languages, suggesting that they fall within the range of phonological systems that are allowed by UG. This leads to the view

⁷ As Broselow and Finer points out, relative sonority is generally assumed to show variation across languages (cf. Selkirk, 1984; Zec, 1988; among others).

on second language acquisition that UG is available to adults. She, however, points out that such a conclusion cannot provide us with an explanation as to why so few adults completely master an L2 phonology, as many previous studies have revealed (Patkowski, 1990, as cited in Young-Scholten, p. 194; See also Long, 1990 for review) so far (i.e., lack of postpuberty ultimate attainment). Young-Scholten, focusing on the Subset Principle (Wexler and Manzini, 1987), and adopting the view that UG is indeed accessible in L2 phonological acquisition and the resetting of parameters can be done, argues that there are situations where positive evidence will not suffice. If the learners' L1 grammar is less restricted than that of L2, the learners can be informed of the fact that the assumptions based on their L1 may produce ungrammatical forms in L2, only through negative evidence. It seems that there are situations in which the Subset Principle shows that the negative evidence is required while it assumes only positive evidence to be available. Instead of abandoning the theory of Subset Principle as an explanation for L2, however, Young-Scholten formalizes the idea which she terms the Asymmetry Hypothesis:

- (5) a. If the L1-L2 situation is such that positive evidence can effect acquisition, then acquisition will ultimately take place.
- b. If the L1-L2 situation is such that negative evidence is required, acquisition will not take place. (Young-Scholten, 1994, p. 201)

Her investigation on the acquisition of flapping and resyllabification by the German speakers learning (American) English and the English speakers learning German showed that prediction made by her Asymmetry Hypothesis was borne out and that the Subset Principle does seem to effect learnability. However, as she addresses at the end of her paper, further research is needed to fully test the Asymmetry Hypothesis. If the learners who start out with the subset (restricted grammar) do actually acquire the superset (less restricted grammar) and if the reverse case does not take place (i.e., learners who begin with the superset do not acquire the subset), then, we can conclude that the Asymmetry Hypothesis is at work.

Regarding the learnability of consonant clusters, Subset Principle makes a prediction that Japanese learners of English (with the more restricted consonant cluster types, i.e., the subset) would receive positive evidence in the input that certain syllable types are possible (since more syllable types are possible in English than in Japanese). Their learning, therefore, may be facilitated, while English learners of Japanese should have difficulty acquiring syllables with fewer consonant cluster types because of lack of such positive evidence. Empirically, however, this prediction seems to be incorrect. As a teacher of English in Japan, I have found that there are many Japanese learners of English who have great difficulty learning English consonant clusters. I presume that not only the principles of UG but also the negative interference of L1 (Japanese) is affecting their L2 (English).

Several studies in the field of L2 phonology have focused on the issues that are closely related to my main theme, realization of consonant clusters (Tarone, 1980; Sato, 1983; Saunders, 1987; Hancin-Bhatt & Bhatt, 1997; among others). Saunders (1987), for example, investigated the morphophonemic variation in clusters in Japanese English. He examined and analyzed data that are related to difficulties encountered by Japanese learners of English in their production of two English bimorphemic stop-fricative clusters *-s* (third person singular present tense morpheme to verbs and the plural morpheme to nouns). The aim of his study was to determine whether the phonemic environment affects the level of error and the duration of the fricative, and whether different types of convergence are likely to affect the same phenomena. He conducted two studies, the morpheme study (L2 adult longitudinal study of morpheme development in spontaneous speech) and the acoustic study (a reading designed specifically to elicit occurrences of the clusters for acoustic analysis). His findings include the following: reduction is a favored production strategy, noun attachments are subject to less error than are verb attachments, the initial consonant of the cluster affects the rate of omission (morpheme study) and considerable evidence for post-cluster epenthesis and mid-cluster epenthesis (acoustic study). While his study reveals that there is massive transfer of syllable structure constraints from L1 to L2 confirming many other studies mentioned above, it does not go beyond that. The possibility that some developmental effects such as discussed earlier (universal constraints such as MSD) may be involved in the process is lacking from his study.

The study conducted by Hancin-Bhatt & Bhatt (1997) articulates and deals with this issue using OT (Optimality Theory). They argue that transfer effects interact with developmental effects in the construction of L2 syllable structures, providing evidence from a study of ESL learners. Based on the Minimal Sonority Distance Parameter-Setting (MSD) model of Broselow and Finer, their study extends MSD by adding further evidence to support the model. They then argue that OT provides a better account of the interactions between transfer and developmental effects because OT allows us to interpret not only why L2 learners have difficulty with specific structures but also how they resolve it within one grammatical framework.

While these studies provide us with some insights on this important issue, many of them seem to have their focus merely on the phonological aspect and little attention has been paid on the possibility that phonetics may be involved in the process.

Hancin-Bhatt and Bhatt (1997), for example, conducted experiments investigating the interactions of transfer and developmental effects (universal processes or constraints). Their study⁸ extended Broselow and Finer's (Minimal Sonority Distance) model in L2 syllables, providing

⁸ They also argue that using an OT (Optimality Theory) approach provides a better hypothesis on the nature of these effects than previous studies.

evidence to support the idea that transfer effects play an important role in L2 syllable structure and that they interact with developmental effects. I will follow them in that for second language acquisition both cross-language transfer and universal constraints interact with one another in a complex way.

Their research is more thorough and insightful than Saunders' (1987), but it seems to lack phonetic support. Their data were said to have been transcribed in narrow phonetic transcriptions from the recordings, but neither acoustic nor articulatory data were carefully examined. Without such examination, however, we cannot assess the role of phonetics involved in the process. What we need is to explore the possibility of phonology/phonetics interface, regarding the realization of L2 syllables. In other words, we need to more carefully examine the data through acoustic and articulatory analysis and determine how phonetics may be involved in the realization of consonant clusters. As introduced in Section II above, Zsiga (1995) discusses the important role that phonetics may play in L2 phonology and argues that many postlexical 'rules' may have resulted from habits of articulatory coordination and thus, should be described in terms of articulatory timing, rather than feature-changing (phonological) rules. In order to determine the possible role of phonetics involved in the processes of L2 sound acquisition, therefore, conducting both acoustic and articulatory experiments and analyzing the data carefully is necessary. I think that this study will have important implications in L2 sound acquisition because many of the previous studies seem to have focused merely on the phonological side and failed to consider the role of phonetics in the process. In the next section, the importance of phonetics as well as phonology/phonetics interface will be reviewed.

4. The importance of phonetics and phonology/phonetics interface

There are two specific points that I would like to address regarding the importance of the phonology/phonetics interface issue in the present study. One is associated with what Levin (1987) proposed – the existence of epenthetic and excrescent vowels, and the other, the issue of gesture and temporal organization. Then, I will briefly review the paper discussing the role of the phonotactic properties in the perception of consonant clusters.

4.1 Epenthesis vs. Excrescence

Levin (1987) claims that there are two different types of vowel insertion processes. They are referred to as epenthesis, a phonological vowel insertion and excrescence, a phonetic vowel insertion. According to her study, epenthetic and excrescent vowels have distinct intrinsic properties which result from the point at which their values are assigned. Levin gives an example of excrescence from the Piro language. The excrescent vowels are, as opposed to the underlying vowels, do not play any role in phonological rules and are significantly shorter. Also, features of excrescent

vowels can be viewed as either the result of language-specific phonetic coarticulation effects, the result of universal phonetic transitional articulatory effects, or the result of universal phonetic default rules, and can be viewed as mediating adjacent articulations requiring some degree of constriction in the oral tract (Levin 1987, p.4).

Based on this notion of epenthesis vs. excrescence, I predict that when a vowel insertion takes place, the inserted vowel is expected to be short, gradient and unclear, if the vowel is an excrescent vowel.⁹ This prediction is in accordance with Browman and Goldstein (1986, 1989, 1992)'s analysis on "Targetless" schwa. Based on the theory of Articulatory Phonology, Browman and Goldstein investigated English reduced (schwa) vowels. Their investigations showed that these vowels take the acoustic and articulatory properties of neighboring vowels rather than having its own target. They argue for the possibility of a kind of schwa which is completely unspecified for tongue position (hence, targetless). These excrescent vowels may merely be inserted for the ease of articulation, in which case phonology cannot refer to them (Levin 1987), and are indeed targetless.

4.2 Gesture and timing

Browman and Goldstein (1986, 1990, 1992) proposed an approach to phonological representation based on the idea of gestural overlap, called Articulatory Phonology. Since then, many researchers have investigated the gestures and timing in a sequence of two consecutive consonants in English and have found that there is significant overlap of gestures in such structures. That is, when two consonants are adjacent to each other, it is often the case that the movement of the closure for the second consonant starts before the first consonant's articulation is completed. According to Articulatory Phonology, in fast or casual speech, gestural overlap of the articulators of the consonants can become so great that it can cause the perception that a consonant has been deleted or assimilated. What this suggests (among other things) is that mere phonetic transcription, which relies on 'human ears' may not accurately capture the actual processes of articulation. What seems to be deletion may not be deletion but the result of gestural overlap, the articulation of the first consonant being completely masked (overlapped) by the following consonant. If this is the case, clearly, we do not need to posit a phonological rule of deletion.

Byrd (1992, 1996) conducted series of experiments investigating the topics such as perceptual effects of gestural overlap or coproduction on the first consonant in a cluster and influences on articulatory timing in consonant sequences. The results of her latter study suggest that inter-gestural coordination is variable and affected by linguistic factors. She considered both degree of temporal overlap and variability in timing and found that place, manner, and sequence

⁹ The other case, an epenthetic vowel, was discussed earlier with respect to loanword phonology.

syllabification were all demonstrated to have an effect on consonant sequence timing.

The results of her studies suggest that in considering inter-gestural timing, Articulatory Phonology offers an interesting framework because Articulatory Phonology as the tool which fills the gap between the gradient nature of phonetics and categorical nature of phonology. Byrd's findings also seem to suggest the importance of phonetic research as the basis for salient phonological research.

Some of the other research based on Articulatory Phonology includes the study by Barry (1991, 1992), Nolan (1992), Zsiga (1993, 1994, 1995). Barry (1991), for example, used electropalatography (EPG) and found that there is some tongue-tip gesture even when an alveolar has apparently been deleted. This kind of gestural overlap can only be captured by a device such as electropalatography (EPG). In the next section, I will refer to EPG and discuss the important role it plays in the present study.

4.3 Perception of consonant clusters

Dupoux, Kakehi, Hirose, and Pallier and J. Mehler (1999) conducted four cross-linguistic experiments on French and Japanese hearers (10 of each) in order to investigate whether phonotactic properties of Japanese, which has a very reduced set of syllable types, induce them to perceive vowels that are not present ("illusory" vowels). The results of their experiments strongly suggest that Japanese, but not French, listeners 'heard' the illusory vowels between consonants. In their study, Dupoux, Kakehi, Hirose, and Pallier and J. Mehler conducted four experiments to explore the possible account for vowel epenthesis in Japanese in loan words. Specifically, they investigated the possibility that vowel epenthesis can take place at the perceptual level. Experiment 1 was conducted to assess the extent of the epenthesis effect, using nonword stimuli with a continuum ranging from trisyllabic tokens (such as *ebuzo*) to disyllabic tokens (such as *ebzo*). Japanese subjects were expected to report the presence of the vowel [u] if the epenthesis effect had a perceptual basis. French subjects, on the other hand, should report the presence of epenthetic vowel less than the Japanese. The result of their experiment confirmed their hypothesis. Other experiments also confirmed their proposal: the obtained effects are due to phonotactics. Their study has revealed that Japanese listeners show tendency to perceive an illusory epenthetic [u] vowel within consonant clusters. Based on their findings, they conclude that in order to account for phonotactically-based assimilation, we need to revise the models of speech perception. In my study, the production of consonant clusters by Japanese speakers will be examined, and whether their findings can be extended or not will be investigated.

SELECTED BIBLIOGRAPHY

- Anderson-Hsieh, J., R. Johnson, & K. Koehler. (1992). The relationship between native speakers judgements of nonnative pronunciation and deviance in segmentals, prosody, and syllable structure. *Language Learning* 42: 529-555.
- Archibald, J. (1994). A formal model of learning L2 prosodic phonology. *Second Language Research* 10, 3:pp. 215-240.
- Barry, M. C. (1992). Palatalization, assimilation and gestural weakening in connected speech. *Speech Communication* 11, 393-400.
- Broselow, E. (1987). An investigation of transfer in second language phonology. In Ioup, G. & S. H. Weinberger ed. *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. Cambridge: Newbury House Publishers.
- Byrd, D. (1994). Palatogram reading as a phonetic skills: a short tutorial. *Journal of the International Phonetic Association* 24:1. 21-34.
- Eckman, F. R. (1987). The reduction of word-final consonant clusters in interlanguage. In James/Leather, eds. *Sound Patterns in Second Language Acquisition*. Foris Publications.
- Eckman & Iverson. (1994). Pronunciation difficulties in ESL: coda consonants in English interlanguage. In Yavas ed. *First and Second Language Phonology*. San Diego, CA: Singular Publishing Group, INC. pp. 251-266.
- Hancin-Bhatt, B. (1994). Segment transfer: a consequence of a dynamic system. *Second Language Research* 10, 3: pp. 241-269.
- Hancin-Bhatt & Bhatt. (1997). Optimal L2 syllables. *SSLA*, 19, 331-378.
- Hashi, Westbury, & K. Honda. (1998). Vowel posture normalization. *JASA*. 104 (4), October 1998:2426-2437.
- Ingram, D. (1995). The acquisition of negative constraints, the OCP, and underspecified representations. In Archibald ed. *Phonological Acquisition and Phonological Theory*. Hillsdale, NJ; Lawrence Erlbaum Associates.
- James, A. (1994). Second language phonology anno 1994. *Second Language Research* 10, 3:pp. 189-192.
- Katayama, M. (1998). Optimality Theory and Japanese Loanword Phonology. Ph.D. Dissertation. UC Santa Cruz.
- Keating, P. (1988d). The phonology-phonetics interface. In f. J. Newmeyer ed. *Linguistics: The Cambridge Survey*. Volume 1: Linguistic theory: Foundations. Cambridge: Cambridge University Press. pp. 281-302.
- Kubozono, H. (1985). Speech errors and syllable structure. *Linguistics and Philology* 6:220-243.
- Ladefoged, P. (1993). *A Course in Phonetics*. 3rd ed. Fort Worth: Harcourt Brace Jovanovich College Publishers.
- Levin, J. (1987). Between epenthetic and excrescent vowels. WCCFL: Proceedings of the West Coast Conference on Formal Linguistics 6: 187-201.
- Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition* 12, 251-286.
- Poser, W. (1990). Evidence for foot structure in Japanese. *Language*, 66, number 1.
- Sato, C. J. (1987). Phonological processes in second language acquisition: another look at interlanguage syllable structure. In Ioup, G. & S. H. Weinberger ed. *Interlanguage Phonology: The Acquisition of a Second Language Sound System*. Cambridge: Newbury House Publishers.
- Saunders, N. J. (1987). Morphophonemic variation in clusters in Japanese English. *Language Learning* 37, 2: 247-272.
- Smith, L. C. (1992). The timing of vowel and consonant gestures. Ph.D. Dissertation, Yale University, Connecticut.
- Sugito, M. (1982). *Nihongo Akusento no Kenkyu*. Tokyo, Japan: Sanseido Shuppan.
- _____. (1996). *Nihonjin no Eigo*. Tokyo, Japan: Izumi Shoin.

澁 谷 良 穂

- Tsujimura, N. (1995). *An Introduction to Japanese Linguistics*. Cambridge, MA: Blackwell.
- Weinberger, S. H. (1994). Functional and phonetic constraints on second language phonology. In Yavas ed. *First and Second Language Phonology*. San Diego, CA: Singular Publishing Group, INC.
- _____. (1997). Making distinctions: adult second language and the Obligatory Contour Principle. Paper presented at the University of Maryland Linguistics Colloquium.
- Young-Scholten, M. (1994). On positive evidence and ultimate attainment in L2 phonology. *Second Language Research*, 10, 3: 193-214.
- Zsiga, E. C. (1993). Features, gestures, and the temporal aspects of phonological organizations. Ph. D. Dissertation, Yale University, Connecticut.
- _____. (1995). An acoustic and electropalatographic study of lexical and post-lexical palatalization in American English. In Connell, B. & A. Arvaniti eds. *Papers in Laboratory Phonetics IV*. Cambridge: Cambridge University Press.
- _____. (1995). Phonology and phonetics in the education of second-language teachers. *Georgetown University Round table on Languages and Linguistics 1995*: 575-587.