Knowledge and Performance: An Examination of the Role of Explicit Linguistic Knowledge in L2 Phonological Acquisition

Jen Mah
981976

LING 598
Dr. John Archibald

Contents

Knowledge and Performance: An Examination of the Role of Explicit Knowledge in L2 Phonological Acquisition

1. Introduction

The acquisition of a second language is an intriguing process, differing importantly from first language acquisition in the final state achieved by the learner: while children arrive at full native competence levels in their first language, the degree of success achieved by second language learners is highly variable. Furthermore, while children acquiring their first language (L1) do so in a seemingly automatic and effortless fashion, L2 students are often seen to struggle, and the end result very rarely ends up being native-like, most remarkably so in the domain of phonology (Scovel 1969, reported in Archibald and Libben 1995), resulting in difficulty perceiving non-native contrasts and an audible non-native accent (Archibald 1996, in O'Grady and Dobrovolsky 1996).

The universal success in L1 acquisition, given that this complex process is being performed by the immature cognitive systems of young children, has led researchers to posit Universal Grammar (UG), a system of principles and parameters that constrains the possibilities considered by a child to be a part of their L1, thus facilitating the task of acquisition: principles are defined as those linguistic universals that account for the remarkable similarities that exist cross-linguistically, while parameters are those linguistic variables, often described as "switches" or a set of options (O'Grady 1996, in O'Grady and Dobrovolsky 1996), which allow for cross-linguistic variation. UG thus establishes the patterns that are permissible in human language, the child then approaches the task of L1 acquisition as a selector, setting parameters in accordance with the language of the environment. Despite the observed universal success of L1 acquisition, there are cases of individuals who have failed to fully acquire their L1. The most notable of these is the case of Genie (Curtiss 1977, reported in Berko-Gleason 1997) a girl who was deprived of exposure to the linguistic environment until the age of 11, at which time she was recovered by social services. Although she received intensive therapy, her linguistic skills never developed to native-like competence. Cases like Genie's have been touted as evidence for the Critical Period Hypothesis, which states that UG must no longer be accessible to the learner once a certain age has passed, usually set at around 13 years of age (Lenneberg 1967; Penfield 1965; both reported in Archibald and Libben 1995); however, much debate has arisen with respect to the ages associated with the Critical Period Hypothesis and the relation to neurological development (Long 1990).

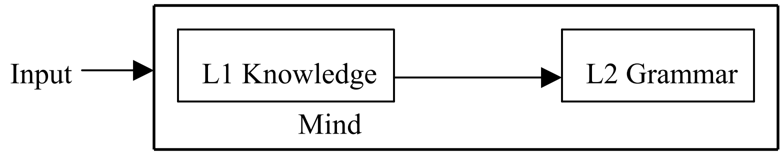
Further evidence in support of the Critical Period Hypothesis is found in L2 acquisition literature: L2 learners are found to rarely achieve native-like proficiency in the L2 (Newport 1990; Long 1990; Birdsong 1999). This body of research has prompted the proposal that UG is not accessible in L2 acquisition, formalized by Bley-Vroman (1989) as the Fundamental Difference Hypothesis: L1 and L2 acquisition differ in both the learner's definitions of a possible grammar, and in the procedure or set of procedures used to arrive at a grammar based on the available data (p. 51). Bley-Vroman (1989) argues that L2 acquisition proceeds by general problem solving skills which are not specific to the language faculty. On the other hand, White (1989) argues that UG does play a role in L2 acquisition, as some of the issues in L1 acquisition that prompted researchers to posit UG are also present in L2 acquisition. Both arguments do capture the fact that adult L2 learners' pronunciation is observably different from that of native speakers.

An interesting line of research has resurrected the notion that it is the L1, and not access to UG, that crucially affects competence in the L2. This transfer-based approach captures an important difference between L1 and L2 learners: only L2 learners come to the task of acquisition already knowing a language. Despite general agreement in the field that the L1 does play a role in L2 acquisition (Bley-Vroman 1989; White 1989), there is considerable debate as to how and to what extent the L1 affects L2 acquisition. Beginner L2 learners view the linguistic world as it corresponds to their knowledge of their first language; precisely how the previous knowledge affects the acquisition of new knowledge and the duration of this effect are still unknown.

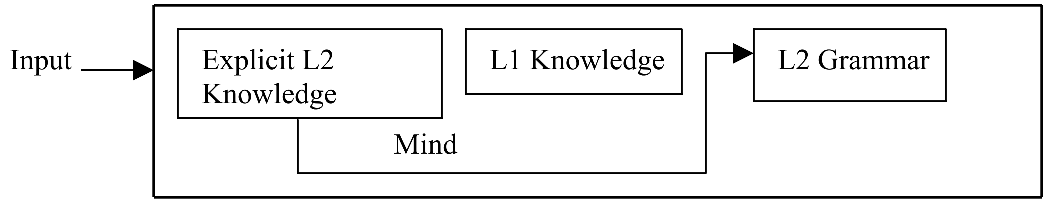
Furthermore, studies have shown that the filtering effects of the L1 are not insurmountable: there are cases of L2 learners who have been judged to have achieved a level of proficiency that is indistinguishable from the performance of native speakers (Bongaerts, Mennen and van der Slik 2000), and training in both perception and production of targeted segments has been shown to offer some improvement, albeit not to native-like proficiency (Matthews 1997). These studies, however, fail to address an important question: what is the factor that allows the learner to bypass knowledge imposed by the L1 and thus build the new structure required for L2 acquisition?

The present research aims to evaluate the possibility that it is explicit knowledge that is the key factor in determining the level of achievement possible in phonological acquisition for an adult L2 learner through examination of production data for evidence of acquisition of a novel phonological contrast. Here, the term explicit knowledge refers to the learner's conscious awareness of some linguistic property's presence and operation in the L2. The following diagrams may help to illustrate this point: (1) illustrates the input being modified by L1 knowledge in the mind of the learner before reaching the developing L2 grammar; (2) illustrates the possibility being examined here, that explicit linguistic knowledge intercepts the input before it is modified by the L1 knowledge, allowing the input to reach the developing L2 grammar as is.

(1)



(2)



The selection of L2 phonology as the domain to be examined is not accidental; a considerable body of research has been devoted to the study of the pronunciation of L2 learners. There are two reasons for this: firstly, as will be further discussed in Section 2 below, infants have been shown to demonstrate a sensitivity to their native language's phonological system that develops within the first year of life, both in terms of prosody (Mehler et al. 1998, reported in Jusczyk et al. 1995; Jusczyk, Cutler, and Redanz 1993) and segmental distinctions (Werker and Tees 1983, reported in Jusczyk 1997; Polka and Werker 1994); phonology may be the first linguistic domain to emerge through the course of L1 acquisition. Secondly, and perhaps more importantly given the context of this research, there is evidence that phonological abilities are the most sensitive to age effects. Success rates for acquisition of L2 phonology begin to decline as early as age 6 (Long 1990). These findings suggest that if L1 knowledge affects the perception and/or production of the L2 being acquired, its effects would be most observable in the L2 learner's phonology.

The paper is structured as follows: Section 2 will discuss the nature and implementation of the filter itself, through a survey of research dealing with infant perception, phonological development, and L2 learner performance; Section 3 will examine the contrasts being examined in the present research through a discussion of the relevant features of the L1 and L2 phonological systems; Section 4 will outline the methodology used to gather and analyze the data; Sections 5 and 6 will present the data collected and resulting analyses; Section 7 will sum up with a discussion of the results, their implications for both phonological theory and L2 acquisition theory, and directions for further research.

2. Linking L1 and L2 Acquisition

2.1. Early Phonological Development

Although children do not begin producing coherent, comprehensible utterances until they have reached the age of approximately two years, an extensive body of research has been devoted to the investigation of pre-verbal linguistic abilities (Burnham 1986; Friederici and Wessels 1993; Goodsitt, Morgan and Kuhl 1993; Hohne and Jusczyk 1994; Jusczyk 1993a, 1997; Jusczyk and Aslin 1995; Jusczyk, Cutler and

Redanz 1993; Jusczyk, Hohne and Mandel 1995; Polka and Werker 1994). A variety of innovative investigative techniques have been developed that allow researchers to observe pre-verbal infants' linguistic behaviour by measuring their physical reactions to a change in the linguistic environment. Techniques such as the head-turn procedure, the preferential looking task, and the high-amplitude sucking procedure allow us to reliably assess the perceptual abilities of infants who have not yet begun to speak.

These studies of infants' perceptual abilities show a definite developmental path: at earlier stages of development, perceptual abilities can be termed as being language-general, in that they are able to discriminate a variety of contrasts, both native and non-native; conversely, at later stages of development (still in the first year of life) their discriminatory abilities are limited to those contrasts found in the L1 being acquired, which can be termed as language-specific abilities. Shortly after birth, infants have been shown to exhibit a high degree of sensitivity to the prosodic structure of the L1, to which they have been amply exposed in utero: Mehler et al. (1988, reported in Jusczyk, Hohne, and Mandel 1995) demonstrated that 4-day old infants were able to distinguish utterances in their mother's L1 from utterances in other languages. Furthermore, these same infants were unable to distinguish between utterances of two languages when neither was the mother's L1. Mehler et al.'s (1988) argument was that these discriminatory abilities were based on prosody; the wall of the uterus acts as a low-pass filter, allowing only the lower frequencies to reach the infant, thus the newborn has had ample experience with the prosodic structure of the L1 and is able to make the observed discriminations. More important to the present research, however, is the development of perceptual abilities with respect to segmental contrasts in the L1. Hohne and Jusczyk (1994) found that 2-month old infants displayed sensitivity to a variety of allophonic distinctions, and argue that it is unlikely that these very young infants are analyzing the segments as allophones, but rather are able to distinguish them since the distinctions may later become important phonological features used in setting up the L1's phoneme inventory and segmenting the speech stream. Jusczyk (1997) reports on a number of studies that provide evidence that infants are initially able to distinguish both native and non-native contrasts for a variety of segments, including novel contrasts. These infants are thus equipped with the perceptual abilities to acquire the phonological system of any of the world's languages.

Yet these language-general abilities do not remain, and, as was demonstrated in the case of prosodic structure, infants' perceptual abilities with respect to segmental structure follow a developmental path whereby the infant becomes attuned to the phonological contrasts and features of the L1 alone. Jusczyk (1997) reports on a number of investigations finding that a decrease in perceptual sensitivity to non-native consonant contrasts began setting in at 8 to 10 months of age, and was firmly in place by 10 to 12 months. In their examination of infants' perception of non-native vowel contrasts, Polka and Werker (1994) found that the decrease in perceptual abilities detected for non-native consonant contrasts set in earlier for non-native vowel contrasts: at 6 to 8 months of age, a decrease in sensitivity was observed, and this decrease was even more apparent among 10- to 12-month olds. The earlier onset of decrease in perceptual abilities for non-native vowels was argued to be due to the greater saliency of these segments in the speech stream. These studies demonstrate that within the first year of life, the infant's perceptual abilities undergo a massive developmental change, the result being that only those contrasts active in the language being acquired are recognized.

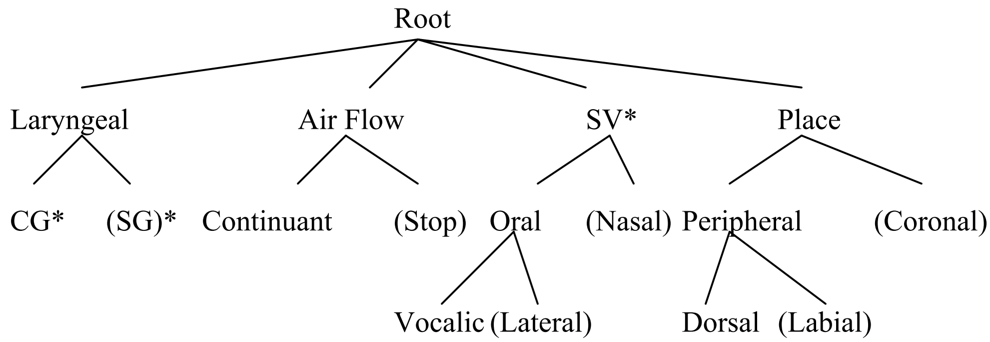
Researchers argue that this decline in perceptual abilities is beneficial to the task of L1 acquisition, as the system is appropriately constrained so as to reduce the processing load on the infant's immature cognitive system by restricting limited attention resources to the relevant contrasts in the environment (Jusczyk 1993). Yet it is precisely this developmental change in perceptual abilities that results in the difficulty L2 learners frequently have in perceiving and producing certain non-native contrasts. It appears that the filtering effects of L1 knowledge serve a beneficial purpose during L1 acquisition; yet this same knowledge serves a detrimental purpose through the course of L2 acquisition. This results in the L1 phonology's considerable influence on speech perception and the developing L2 system.

2.2. Building Knowledge– A Model of Phonological Development

Rice and Avery (1995) develop a theory of segmental elaboration based on feature geometry and present a model under which acquisition proceeds following pre-determined learning paths. Children begin the process of phonological acquisition with only phonological universals, the innate endowment from UG. The phonemic inventory of the L1 is built up through elaboration of structure: the child begins with an impoverished structure, and as phonemic contrasts are detected in the environment, the appropriate distinctive features are added to the grammar.

The notion of a learning path enters the picture when one considers how feature geometry organizes segmental structure. According to the feature geometry model, segments have internal structure, which is in turn organized into a hierarchical tree, with features grouped into nodes that reflect dependency and markedness relationships among the features. The basic tree given in (3) details the organization of the nodes and features that are used in describing segments using a model of feature geometry (Rice and Avery 1995, p. 31); unmarked features are enclosed in parentheses. These are not thought of as fully active in the phonology until they are motivated by an appropriate contrast; once a contrastive feature is added to the tree, then both features are specified.

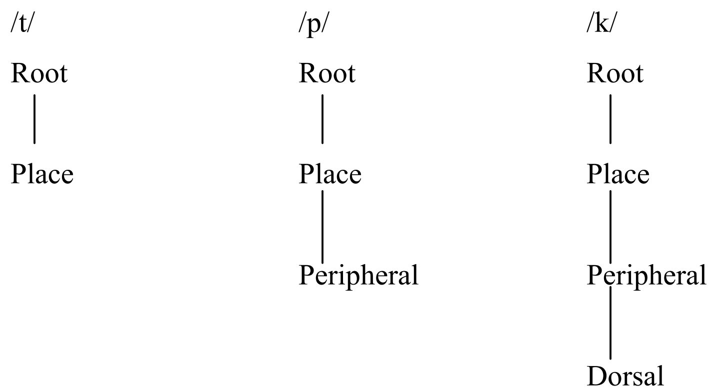
(3)



*It should be noted that the abbreviation SV is used here to denote the Sonorant Voice node, which "organizes features normally associated with sonorant segments such as nasals, laterals, and r-sounds" (Rice and Avery 1995, p. 31). This does not refer to the feature of Voice, which, while not illustrated in the above tree, is associated with the Laryngeal node. The abbreviation CG is used to denote the Constricted Glottis node; the abbreviation SG is used to denote the Spread Glottis node.

Phonemic segments are underspecified; that is, redundant information (such as the unmarked features illustrated in (1)) is absent from the underlying phonological representations, as exemplified in (2) below. Given the contrast /t/ - /p/ - /k/ in (2), the step-by-step addition of distinctive features to the child's grammar allows development of the appropriate phonemic contrast.

(2)



(Rice and Avery 1995, p. 32)

Given the structures posited here, this model makes predictions about the order of acquisition of contrasts (i.e., children will first distinguish /t/ vs. /p/, followed by the three way contrast), which Rice and Avery (1995) claim are generally borne out in the existing data. Furthermore, given the dependency relations inherent in these structures, learning can only proceed along a predetermined path within any particular organizing node. Differential expansion of phonemic inventories cross-linguistically is argued to be the result of elaboration within different nodes at of development; individual variation within a language is attributed to differential detection of (and importance attributed to) contrasts in the linguistic environment.

Thus we have seen evidence, from both observational studies and a theoretical standpoint, for filtering effects of L1 knowledge in the learner's perception of linguistic input (and potentially the learner's production output). Although they begin life with minimal phonological structure, enabling them to acquire any of the world's languages, the relevant segmental structure is elaborated so that the appropriate L1 contrasts are represented in the grammar. Children's perceptual abilities also become attuned to the L1 within the first year of life, severely reducing the ability to perceive the non-native contrasts that were perceptible at an earlier stage of development.

2.3. Filtering Effects of L1 Knowledge

Given the results from the studies reviewed above, an intriguing possibility presents itself to the researcher: there may be a causal relationship between phonological development and speech perception. We now turn to a study conducted by Brown (2000), in which she argues that the L1 filtering effects observed in both L1 and L2 acquisition are indeed a result of the segmental elaboration of feature geometry that takes place during L1 phonological acquisition. Brown (2000) claims that the acquisition of a phonological system not only determines the course of development for speech perception in L1 acquisition, but continues to constrain speech perception in L2 acquisition.

A survey of data from studies of categorical speech perception (reported in Brown 2000) suggests that the phonological system mediates between the acoustic signal and the linguistic system, allowing the speech stream to be broken down into meaningful units for processing and comprehension according to the phonemic category boundaries set out by the phonological system. This results in acoustically distinct stimuli being perceived as one segment, provided both tokens fall within the appropriate category boundaries. Thus the loss of perceptual abilities observed in studies of infant perception is not so much a loss, but a reorganization of those same perceptual abilities: the infant can still detect the variability of the acoustic signal, but has developed phonemic categories in which numerous tokens are deposited and interpreted as one distinct segment. That is to say, the system no longer attributes any linguistic importance to such variability in the speech stream.

This sort of perceptual reorganization has readily identifiable benefits, not only for first language acquisition but also for processing of language among adult native speakers: irrelevant "noise" occurring in the acoustic signal as a by-product of everyday language use, such as variable realizations, inter-speaker variability, and coarticulations, are disregarded, thus reducing the listener's processing load. Yet it quickly becomes apparent that this link between phonological development and speech perception acts as an impeding force that the listener must overcome in order to successfully acquire the L2 phonological system: it is fully possible that two phones used contrastively in the L2 may be interpreted as two separate instances of one phoneme in the L1, in which event the learner will be unable to perceive the contrast, such as Japanese learners' inability to differentiate between English /l/ and /r/, owing to the fact that Japanese has only one liquid phoneme (Matthews 1997). Effectively, Brown's (2000) proposal suggests that the phonological system of the learner's L1 will accurately predict which non-native contrasts will be perceived and thus successfully acquired, and which ones will not.

This same concept of L1 influence in L2 acquisition was the driving force behind earlier studies involving contrastive analysis, in which the phonemic inventories of the L1 and L2 were compared; the learner was expected to struggle with L2 contrasts involved phonemes that were absent from the L1 (Lado 1957, reported in Archibald and Libben 1995). This prediction, however, was not borne out: not only did L2 learners experience difficulty on certain novel contrasts, as was expected, they also experienced ease on other novel contrasts. This posed a rather hefty problem for the theory of contrastive analysis. Wardhaugh (1970, reported in Archibald and Libben 1995) attempted to compensate, by positing two versions of contrastive analysis: strong (or predictive), which attributes all areas of difficulty in the L2 to the structures of the L1 (this is the version originally set forth by Lado 1957), and weak (or diagnostic), which acknowledges that areas of L2 difficulty may stem from sources other than the L1. Brown (2000) claims that the failure of contrastive analysis lies in its identification of the components of the L1 grammar that cause interference: it is not the phonemes themselves, but rather the features contained in the learner's grammar that accurately predict perception of non-native contrasts. The featural level, rather than the segmental level, plays the crucial role in perception of L2 contrasts; therefore a feature-based approach is more appropriate in accounting for L2 phonological errors, rather than a segment-based approach. This is most clearly exemplified by Brown's (2000) research, which aimed to investigate the influence of Japanese, Korean, and Mandarin Chinese grammars on the acquisition of English contrasts not found in the L1.

Subjects participating in the study were native speakers of either Japanese, Korean, or Mandarin Chinese. The contrasts examined were /l/ - /r/, /b/ - /v/, /p/ - /f/, /f/ - /v/, and /s/ - /θ/ (another contrastive pair, /p/ - /t/, was included as a control item, as this contrast is active in all three L1s observed). These particular contrasts were chosen not only for their absence (by absence we are referring to an absence of the contrast, not of the phones themselves) in the native grammars of all subjects, but also for the variety of features upon which this group of contrasts relies. The table presented in (4) below illustrates the various contrasts under investigation, and their status in the subjects' L1.

(4)

Contrast	Distinguishing Feature	Feature present in		
		Japanese	Korean	Mandarin Chinese
/l/ vs. /r/	[coronal] (/r/)	no	no	yes
/b/ vs. /v/	[continuant] (/v/)	yes	yes	yes
/p/ vs. /f/	[continuant] (/f/)	yes	yes	yes
/f/ vs. /v/	[voice] (/v/)	yes	yes	yes
/s/ vs. /θ/	[distributed] (/θ/)	no	no	no

(adapted from Brown 2000)

Thus, if the features contained in the L1 grammar are indeed responsible for the observed L1 influence, then the speakers should differ according to native language with respect to their performance on the noted English contrasts. The predictions of the model are as follows: all subjects are expected to perceive and acquire the /p/ - /f/, /b/ - /v/, and /f/ - /v/ contrasts, as the L1 grammars all contain the features relevant to these pairs; none are expected to be able to perceive the /s/ - /θ/ pair, as none of the L1 grammars contain the relevant feature, thus this contrast is not expected to be acquired. The /l/ - /r/ pair, however, is expected to pattern differently: the Japanese and Korean speakers are not expected to be able to perceive this contrast, but the Mandarin Chinese speakers are, as their L1 grammars do contain the relevant feature. Contrastively, a segment-based approach would make the same predictions as the feature-based approach save the last one, in which case the segment-based approach predicts that all three groups will experience the same difficulty with the /l/ - /r/ pair, as none of the phonemic inventories contain both of these segments.

A 4IAX Discrimination Task was used to assess perception of the targeted contrasts: in each trial, subjects were presented with two non-word pairs involving the targeted segments, one of which would consist of a minimal pair (e.g., ra/ra vs. ra/la), they were then required to indicate which pair they believed consisted of two different non-words. Accurate performance in selecting the minimal pair as the instance of different non-words was taken to be indicative of ability to perceive the targeted phonological contrast. The results provided strong support for the proposed model. None of the groups were able to perceive the /s/ - /θ/ pair. Only the Mandarin Chinese speaking group was able to perceive the /l/ - /r/ contrast, despite its absence in the L1. Performance on those contrasts which are, according to the model, acquirable (but are not already present) was also shown to be initially poor, due to inexperience with the contrast, but improved drastically with time so that the performance of the more advanced learners on this task was at native-like levels of proficiency: since the L1 grammar contained the feature required to build the appropriate phonological representations for those segments, the learners were able to eventually detect the contrast in the speech stream. It would seem that in the earliest stages of L2 acquisition, the sounds of the L2 are mapped directly onto the phonological structures of the L1 in any way that fits (no matter how poorly, as the case may be). As acquisition proceeds, mismatches that occur in the early direct mapping prompt a re-evaluation of the acoustic properties of the segment(s). If the appropriate features are present in the L1 grammar, then new phonological representations can be built to better accommodate the new segments, the result being L2 phonological acquisition.

2.4. Additional Supporting Data – An Extension to Prosodic Phonology

Although the model presented in Section 2.3 above was supported by evidence from segmental phonology, it is reasonable to expect that it would extend to the domain of prosodic phonology: both domains exhibit the developmental perceptual attunement described earlier in Section 2.1, although no data has yet been presented demonstrating that infants are initially able to distinguish a variety of non-native prosodic parameters, as exposure to L1 prosody begins before birth, making it extremely difficult to observe and assess the initial state: even newborn infants exhibit attunement to the L1 (Mehler et al. 1988). Still, an examination of L2 learners' behaviour with respect to prosodic phonology would serve as an appropriate test for Brown's (2000) model. Since both segmental and prosodic phonology seem to exhibit the same developmental course in L1 acquisition, we would want to posit one mechanism to account for both, and for the resulting structures in the grammar and their operation in L2 acquisition. An earlier study by Han (1992) provides data with respect to such an extension of Brown's (2000) model.

Han (1992) examined native speakers of English with respect to their performance on Japanese geminate stops. The study looked at the performance of 4 native speakers of American English who had been judged (by a native Japanese speaker) to be fluent in Japanese on a production task designed to elicit the Japanese single – geminate stop contrast. Measurements were taken of the subjects' timing control of the stop closure, and these were compared to the performance of native speakers of Japanese on the same task.

According to Brown (2000), the English speakers, no matter how advanced their competence in Japanese, should not be able to produce the single – geminate stop contrast, since the feature for consonant length is not active in their L1 grammars (it should be noted here that some researchers have argued that English does make use of geminate consonants in connected speech, as in *bla<u>ck c</u>at*; these, however, are analyzed as phonetic, and not phonemic). Han's (1992) data do support this prediction, finding that the English speakers either missed the contrast altogether, or consistently produced one that was not native-like: while native Japanese speakers have been shown to produce a geminate – single stop mean timing ratio of approximately 3:1 (Han 1992; Homma 1981; Han 1962), the English speakers produced a mean timing ratio of approximately 2:1 (Han 1992). Table 1 below details the mean stop closure timing control ratios produced by native Japanese speakers; Table 2 describes the mean stop closure timing control ratios of the American subjects, broken down by subject (Han 1992). The graph in

(5) below illustrates the difference between the mean ratios produced by native Japanese speakers and the mean ratios produced by Han's non-native speakers.

Table 1: Native Japanese Speakers

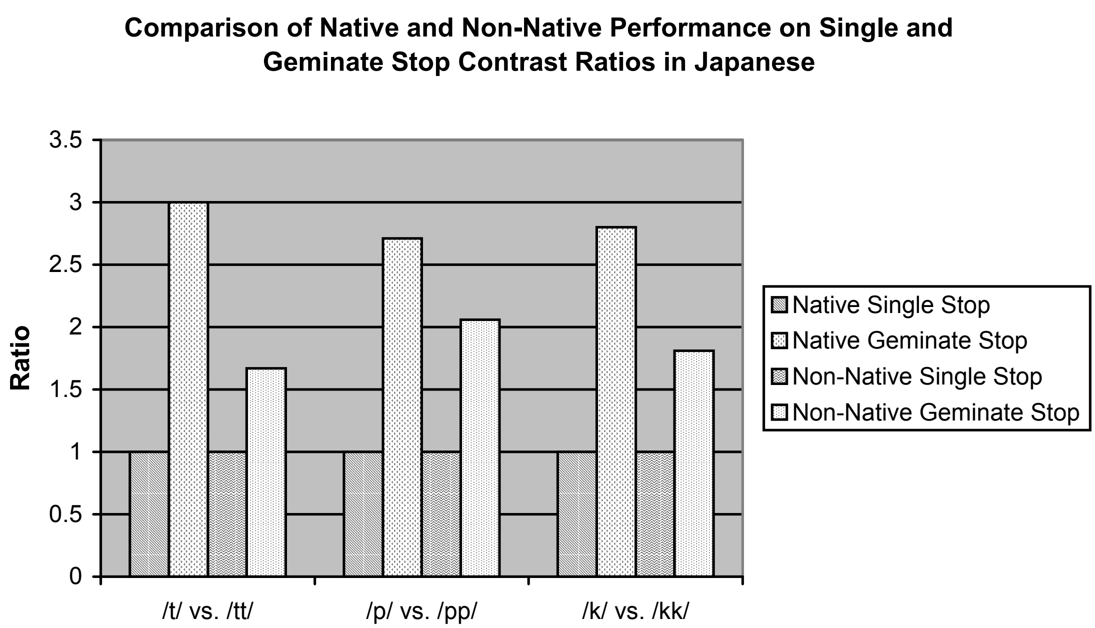
	/tt/ vs. /t/	/pp/ vs. /p/	/kk/ vs. /k/
Mean Ratio	3.00	2.71	2.80

Table 2: American Speakers of Japanese as a Second Language

Subjects	[tt] vs. [t] 1	[tt] vs. [t] 2	[tt] vs. [t] 3	[tt] vs. [t] 4	[tt] vs. [t] 5	[pp] vs. [p] 1	[pp] vs. [p] 2	[kk] vs. [k] 1	[kk] vs. [k] 2	[kk] vs. [k] 3	Mean
A	1.37	1.55	2.10	1.48	1.26	2.03	1.73	1.45	2.27	1.71	1.70
B	2.57	3.35	2.76	2.97	2.27	3.56	2.75	2.55	4.02	1.94	2.87
C	1.21	1.85	1.01	1.01	1.50	2.40	1.89	1.75	1.65	1.08	1.53
D	1.00	1.18	0.90	1.06	0.93	1.10	.098	0.99	1.22	1.14	1.05
Mean	1.54	1.98	1.69	1.63	1.49	2.27	1.84	1.68	2.29	1.47	1.79
			1.67			2.06		1.81			

(adapted from Han 1992, p. 118)

(5)



2.5. Remarks

We have seen how recent developments in language acquisition research have led to the proposal of a model that links L1 phonological acquisition to speech perception, and the related difficulties encountered in the acquisition of L2 phonology. In surveying the previous work, however, two interesting questions are raised. First, Brown's (2000) model makes predictions about the English speaker's control of length for both consonants and vowels in Japanese: we would expect the early learner to perform well with respect to vowel length, since English vowels are specified as being either mono- or bi-moraic (more on this in Section 3 below), all the while performing poorly with respect to consonant length, even at an advanced stage of acquisition, since English consonants are not specified for length. Second, Brown's (2000) model offers a rather bleak view of the final state of L2 acquisition: it can never be completely native-like if a relevant feature is missing from the L1 grammar. Yet Bongaerts, Mennen, and van der Slik (2000) report on a number of non-native speakers of Dutch from a variety of L1 backgrounds (German, English, French, Spanish, Armenian, Berber, Czech, Greek, Italian, Swedish, and Turkish) who have achieved a level of proficiency with respect to pronunciation that has been deemed, by native speakers of Dutch, to be indistinguishable from that of a native speaker. Clearly it is possible to overcome the filtering influence of the L1; the question that remains is how. Han (1992) suggests that explicit instruction may be the key to improving the learner's acquisition of Japanese length contrasts. The present research aimed to address both of these questions by evaluating a native speaker of Canadian English's performance on Japanese consonant and vowel length contrasts; the subject was provided with various types of explicit knowledge throughout the investigation. Before moving on to our discussion of the experiment and its results, a discussion of the relevant phonological structures in both English (the L1) and Japanese (the L2) is in order.

3. Comparing the L1 and L2 Phonology

The present study sought to examine performance on Japanese length contrasts by a native speaker of Canadian English. The following discussion of English and Japanese phonology will elucidate the reasons behind the choice of this particular set of contrasts.

Let us start with the phonemic inventories of the L1 and L2. Tables 3 and 4 below detail the consonant inventories for English and Japanese, respectively; Tables 5 and 6 below detail the vowel inventories for English and Japanese, respectively.

Table 3: Consonant Phonemes of English

		Bilabial	Interdental	Alveolar	Palato-Alveolar	Palatal	Velar	Glottal
Stops	Voiceless	p		t			k	
	Voiced	b		d			g	
Fricatives	Voiceless	f	θ	s	ʃ			h
	Voiced	v	ð	z	j			
Affricates	Voiceless				tʃ			
	Voiced				dʒ			
Nasals		m		n			ŋ	
Liquids	Lateral			l				
	Rhotic			r				
Glides		(w)*				j	w	

(adapted from O'Grady and Dobrovolsky 1996, p. 32)

*The placement of /w/ in parentheses here is due to its classification in English as a labio-velar, having properties of both labials and velars.

Table 4: Consonant Phonemes of Japanese

		Bilabial	Alveolar	Alveolopalatal	Palatal	Velar	Glottal
Stops	Voiceless	p pp	t tt			k kk	
	Voiced	b	d			g	
Fricatives	Voiceless		s ss	ɕ ɕɕ			h
	Voiced		z				
Affricates	Voiceless			tɕ			
	Voiced			dʑ			
Nasals		m mm	n nn				
Liquids			ɺ				
Glides					j	w	

(adapted from Han 1962)

*The use of [ɺ] to represent the alveolar lateral flap is based on Rogers (1991), p. 228.

Table 5: Vowel Phonemes of English

	Front	Mid	Back
High	ij ɪ		uw ʊ
Mid	ej ɛ	ʌ	ow ɔ
Low	æ		ɑ
Diphthongs	aj, oj		aw

(adapted from O'Grady and Dobrovolsky 1996, p. 35)

Table 6: Vowel Phonemes of Japanese

	Front	Mid	Back
High	i i:		u u:
Mid	e e:		o o:
Low		a a:	

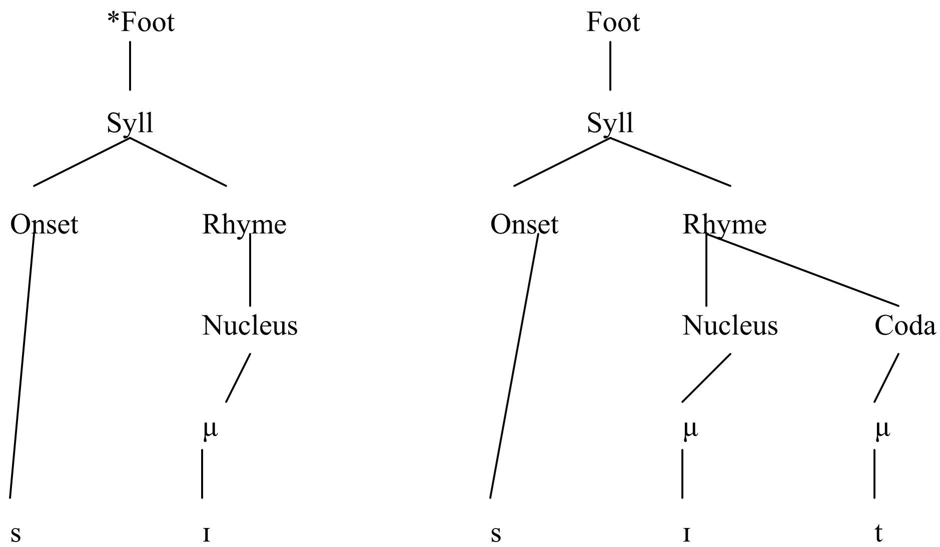
(adapted from Han 1962)

*Homma (1973) transcribes /u/ and /u:/ as /ɯ/ and /ɯ:/ respectively, with some instances being recorded phonetically as [ɨ] and [ɨ:]. It should be noted that the subject also tends to perceive these vowels as being unrounded; additionally, /o/ and /o:/ are perceived as /ɔ/ and /ɔ:/. Thus, /ɯ/, /ɯ:/, /ɔ/, and /ɔ:/ are the transcriptions that will be used throughout the remainder of the paper.

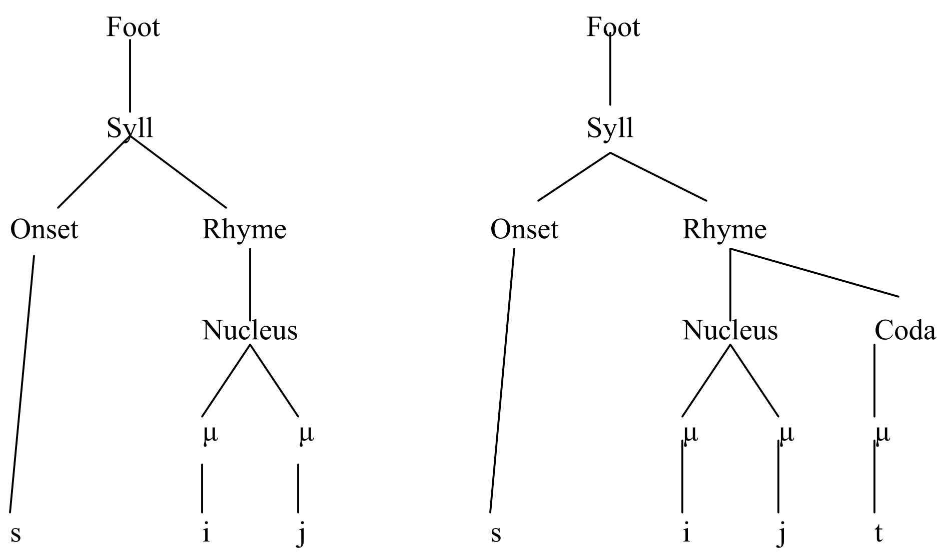
From these tables, we can see clearly that length is a contrastive feature in Japanese for both consonants and vowels, but only for vowels in English. Furthermore, English has a larger vowel inventory with respect to spectral quality of vowels, as well as three diphthongs which are absent in the Japanese vowel inventory. Thus, it is of immediate interest to examine the learner's performance not only in terms of length contrasts (for both consonants and vowels), but also in terms of vowel quality.

Japanese is a language in which moras are the units of prosodic significance (Han 1994, Port et al. 1987). Moras measure out syllable weight: a light syllable consists of one mora, a heavy syllable consists of two (or more) moras (Spencer 1996). Thus, Japanese short vowels are analyzed as being mono-moraic, while long vowels are bi-moraic. The analysis of English vowels is similar: in English, two classes of vowels can be identified: tense vowels (/ij, ej, uw, ow, ɑ/) and lax vowels (/ɪ, ɛ, æ, ʊ, ɔ, ə/), of which tense vowels are all inherently longer than the corresponding lax vowels. This has been analyzed as a phonemic distinction, resulting in an analysis of English lax vowels as being mono-moraic, while English tense vowels are bi-moraic. Evidence for this analysis comes from examination of English mono-syllabic feet, which must be bi-moraic, as illustrated by (6) below: a mono-moraic lax vowel alone may not form the rhyme of a mono-syllabic foot (**si*); however, once a consonant fills the coda position (and is consequently assigned one mora), legal foot structure is obtained (*sit*). Conversely, a bi-moraic tense vowel may form the rhyme of a mono-syllabic foot, both with (*seat*) and without a Coda consonant (*see*), as illustrated in (7)[1]. (Spencer 1996)

(6)



(7)



Han (1992) argues that it is not the actual durational values obtained through measurement that will provide the accurate picture of the durational contrast, as these are easily confounded by differences in rate of speech and speaker-to-speaker variation; rather, it is the ratio obtained from the mean measurements that will provide a meaningful illustration of the nature of the contrast. The ratio of geminate to single stop consonants, illustrated in Section 2.4 above, is approximately 3:1 (Han 1992,

[1] This analysis of vowel length in English is not without controversy; Dobrovolsky (in discussion) notes that English vowel length has been argued to be a phonetic feature resulting from the Advanced Tongue Root distinctive feature in the grammar: vowels that are specified as [-ATR] cannot be long.

Homma 1981, Han 1962), while the ratio of long to short vowels is between 2:1 and 3:1; Han (1962) claims that standalone vowels are produced using a ratio of 2:1, vowels preceded by a voiced consonant are produced using a ratio of 2.5:1, and vowels preceded by a voiceless consonant are produced using a ratio of 3:1. (It should be noted here that due to vocabulary limitations, the consonantal environment of the vowel was not taken into account in this study.)

Given the phonological descriptions presented above, Brown's (2000) model would make the following predictions about L2 performance:

1. A native speaker of English will perform poorly on Japanese consonant length contrasts, even at advanced stages of acquisition.
2. A native speaker of English will perform well on Japanese vowel length contrasts in more advanced stages of acquisition, but may perform poorly in early stages.
3. Substitution of English vowels may occur in early stages as the learner's developing interlanguage grammar attempts to cope with the new vowel system, but this should disappear at later stages of development.

Remembering Han's (1992) proposal that explicit instruction may assist the learner in acquiring Japanese consonant length contrasts, we can make the following additional prediction:

4) A native speaker of English who has received explicit instruction on Japanese consonant and vowel length contrasts will perform well on these, even at early stages of acquisition.

Indeed, any data supporting Han's (1992) proposal would suggest that linguistic knowledge gained through explicit instruction is available to the developing L2 grammar, a topic still under considerable debate.

4. Research Methodology

4.1. Subject

Only one subject (the researcher) provided data for this investigation: a 22-year old female student at the University of Calgary whose native language is Canadian English. The speaker has been a resident of Calgary, Alberta, for the past ten years, prior to that she resided in Edmonton, Alberta. In addition to English, this speaker is also fluent in French and speaks some Spanish: she began acquiring French at the age of five through a French immersion education program, and she began acquiring Spanish at the age of 18 when she began her studies at the university. Although she is not monolingual, her competence with these other languages is not expected to affect her acquisition of Japanese due to the large typological differences that have been observed; particularly, neither French nor Spanish makes use of length contrasts for either consonants or vowels.

The subject began acquiring Japanese in a university classroom setting at the age of 22. Classes were held four times a week, for one hour each class, with an additional hour per week spent on drill exercises that emphasized grammatical use of the linguistic structures of the Japanese language. Very little attention was given to pronunciation; the instructor made minimal comments about long vowels (referring to them as "stretched"), and even fewer comments about geminate consonants (describing the production of these as "swallowing the first part of the sound"). Prior to beginning classes, the subject's linguistic knowledge of Japanese was extremely limited: due to exposure to examples from various Linguistics classes, she knew that Japanese had a five vowel system (in terms of vowel quality), long and short vowel contrasts, and single and geminate consonant contrasts; however, any other knowledge of the language or its operation was nonexistent.

4.2. Data Collection

Data was collected in two sessions, two months apart: once after approximately four months' classroom exposure, and a second time after approximately six months' classroom exposure. Fifteen Japanese sentences were designed to elicit tokens of the targeted contrasts, albeit not in minimal pairs due to vocabulary restrictions. These sentences were each written on an index card using the *hiragana* Japanese script, which the subject had been required to master after one month of classes as part of the course requirements, so that attention and processing abilities would be devoted to decoding meaning from the *hiragana* script, rather than focussing on pronunciation. The cards were randomized, then presented to the subject for reading. To avoid unnatural pauses (and any lengthening that may have resulted) in the reading, only words that were present in the subject's working vocabulary were used; hence the absence of minimal pairs in the data. Another unfortunate result of this vocabulary familiarity condition is that a small number of tokens containing geminate consonants and/or long vowels were deemed appropriate for elicitation. It should be noted here that an additional sentence was presented in the second round of recording to elicit tokens of a geminate consonant (/ss/) that were not deemed appropriate to the first round of recording due to the aforementioned vocabulary restraint. The test sentences are listed in Appendix A.

The subject read each sentence three times, each of which was recorded using a Sony TCD – D100 Digital Audio Tape (DAT) recorder and a Sony ECM – MS908C electret condenser microphone. The recorded data were then re-digitized at a sampling rate of 22.2 kHz using the SoundScope 8 One Channel Analyzer in the Phonetics Laboratory at the University of Calgary. Wide-band spectrograms were made of the recorded data using this same device.

4.3. Measurement Procedures

The following measurements were all obtained through examination of the wide-band spectrograms obtained from the recorded data. The segments under scrutiny were isolated visually using the spectrogram, and then played back to ensure the absence of neighbouring sounds.

4.3.1. Consonant Closure Duration

Stop closure duration was measured by the absence of noise on the spectrogram; fricative closure duration was measured from onset to endpoint of the characteristic "noise" on the spectrogram; nasal consonant duration was measured from onset to endpoint of the characteristic "nasal murmur" on the spectrogram. Due to the difficulty in determining the closure duration of stops in initial position, only intervocalic stops were included in the data analysis. Similarly, since fricatives in final position have a tendency to decrease in amplitude slowly, making it difficult to determine the actual endpoint for these sounds, these were also excluded from analysis.

4.3.2. Vowel Duration

Vowel duration was measured from the onset of glottal vibration, indicated by regular vertical striations on the spectrogram, to the following closure. In the case of vowels in final position, the endpoint of the vowel was taken to be where glottal vibration ceased.

4.3.3. Vowel Quality

Vowel quality was measured in terms of the values of the first and second formants (F1 and F2, respectively) of the vowel, which contain the most information with respect to that vowel's identity in acoustic space (Borden, Harris, and Raphael 1994). F1 and F2 values are easily obtained from a wide-band spectrogram.

4.3.4. Remarks

In addition to being examined for significant differences in duration between single and geminate consonants and short and long vowels, all durational measurements were compiled into mean ratios, and compared against those obtained from studies of native speakers' performance on production of the targeted contrasts, as well as Han's (1992) results involving non-native speakers. Vowel quality measurements were compared against averaged measurements of a female native speaker's productions (Homma 1973) as well as measurements of the subject's own production of English vowels. Additionally, the results from Time I were compared against those from Time II in order to identify any changes in the subject's performance.

5. Results – Time I

5.1. Consonant Closure Duration

The graph in (8) below illustrates the subject's mean timing control of single and geminate consonants; Table 7 details the timing control ratios produced, as illustrated in (8).

(8)

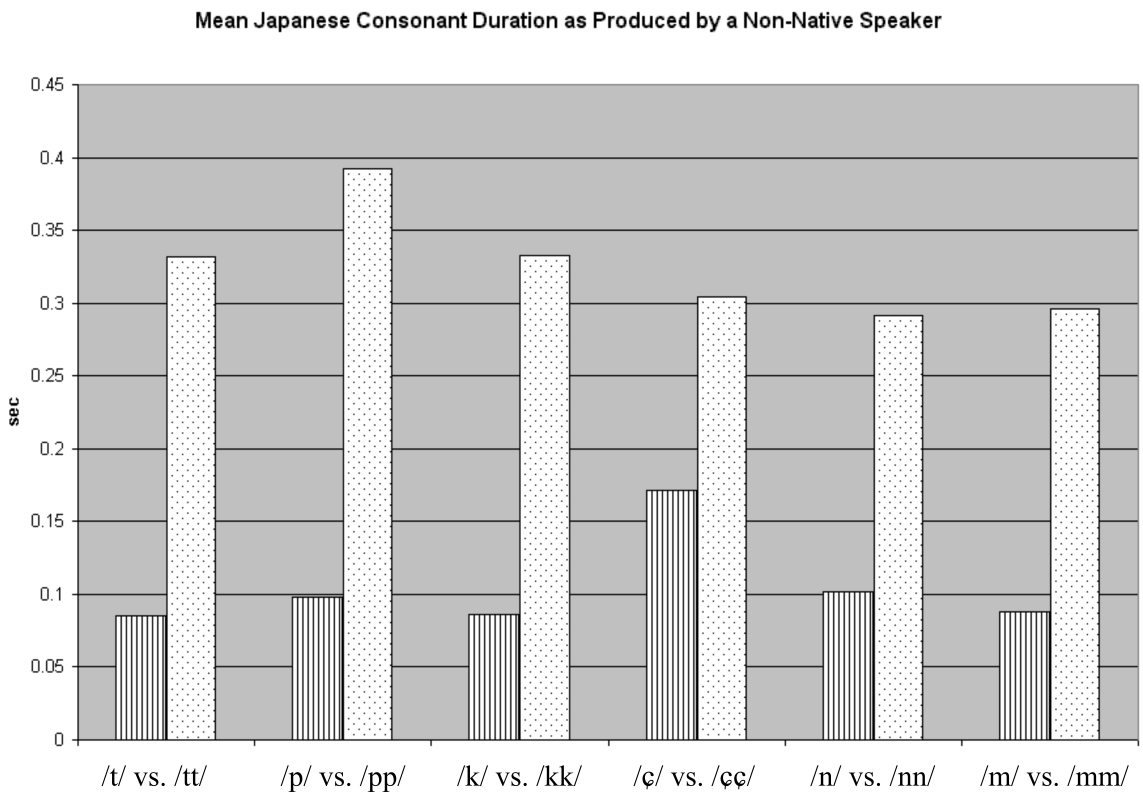


Table 7: Mean Consonant Closure Duration Ratios

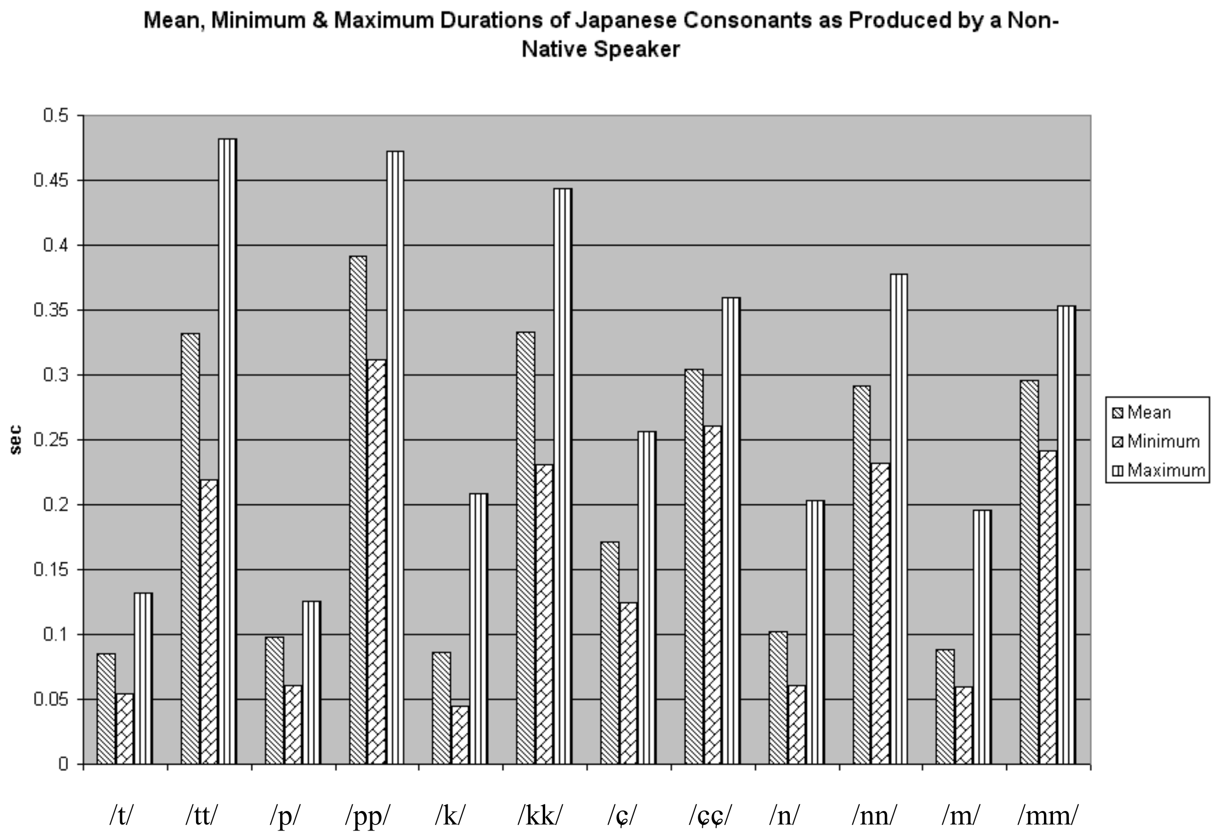
	Single	Geminate	Ratio
/t/ vs. /tt/	0.085	0.332	3.91
/p/ vs. /pp/	0.098	0.392	4.00
/k/ vs. /kk/	0.086	0.333	3.87
/ç/ vs. /çç/	0.171	0.291	1.78
/n/ vs. /nn/	0.102	0.291	2.85
/m/ vs. /mm/	0.088	0.296	3.36

Overall, the subject produced a mean consonant closure duration ratio of 3.26:1.0. This appears promising for our hypothesis, given the mean overall closure duration of 2.8:1.0 (Han 1992) to 3:1 (Han 1962, Homma 1981) reported for native speakers. For all contrasts investigated, a two-tailed *t* test revealed that geminate consonants were significantly longer than their corresponding single consonants [/t/ vs. /tt/: t = -12.820, p < 0.001; /p/ vs. /pp/: t = -10.649, p < 0.001; /k/ vs. /kk/: t = -7.292, p = 0.001; /ç/

vs. /çç/: $t = -8.044$, $p < 0.001$; /n/ vs. /nn/: $t = -7.640$, $p < 0.001$; /m/ vs. /mm/: $t = -12.657$, $p < 0.001$]. These results conflict with the predictions made by Brown's model of L1 interference, based on which we would expect poor performance on these contrasts. The subject's actual performance suggests that she has been able to acquire new phonological structure with respect to Japanese single and geminate consonants and overcome the L1 grammar's inadequacies in dealing with the novel contrast. These results also support our hypothesis that explicit linguistic knowledge is required to overcome the L1 influence: the subject was aware of consonant length contrasts in Japanese and was also able to consistently produce a significant consonant length contrast; at this point, however, it should not be expected that the subject's performance be fully native-like since she was not aware of the specific timing ratios involved.

The contrast she is producing, however, is not native-like in other important but troubling ways. Performance varies among the segment classes: among the obstruents, the ratios obtained for performance on stop contrasts are greater than those of native speakers, while the ratio obtained for the fricative contrast (/ç/ vs /çç/) is smaller than that produced by native speakers. Furthermore, closer inspection of the raw data reveals some variation in the tokens produced for each contrast: for some contrasts, the longest token of a single consonant was just as long or longer than the shortest token of the corresponding geminate consonant. The graph in (9) serves to illustrate this point. This variation both across and within categories suggests that although the subject reliably and consistently distinguishes between single and geminate consonants in her productions, she does not control the timing of these in the same way that a native speaker of Japanese would. Again, this is not surprising at this point, given that she was unaware of how her performance differed from that of a native Japanese speaker when the data were collected. If Han's (1992) proposal that explicit linguistic knowledge will enable the L2 learner to achieve native-like proficiency is correct, then we would expect to see this situation remedied at the second recording.

(9)



5.2. Vowel Duration

The graph in (10) illustrates the subject's mean timing control of vowel duration; Table 8 details the timing control ratios produced illustrated in (10).

(10)

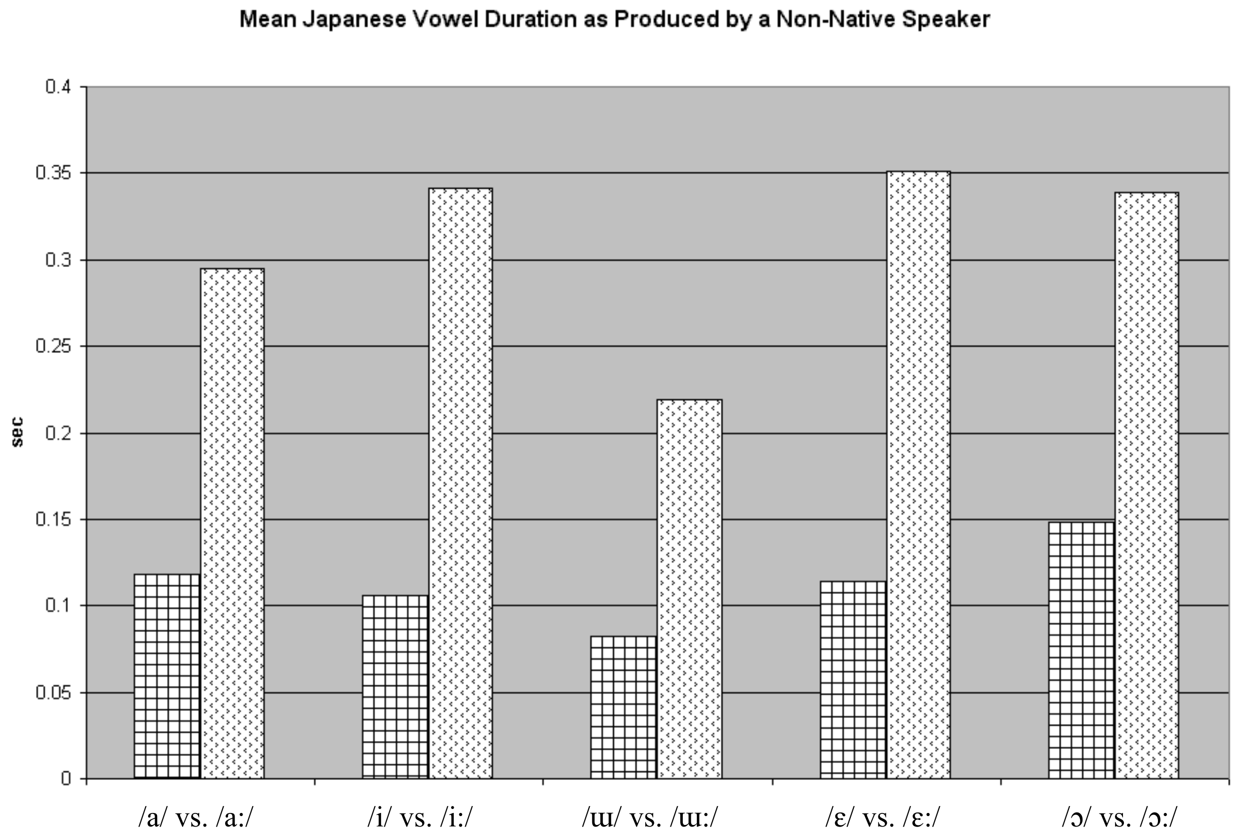


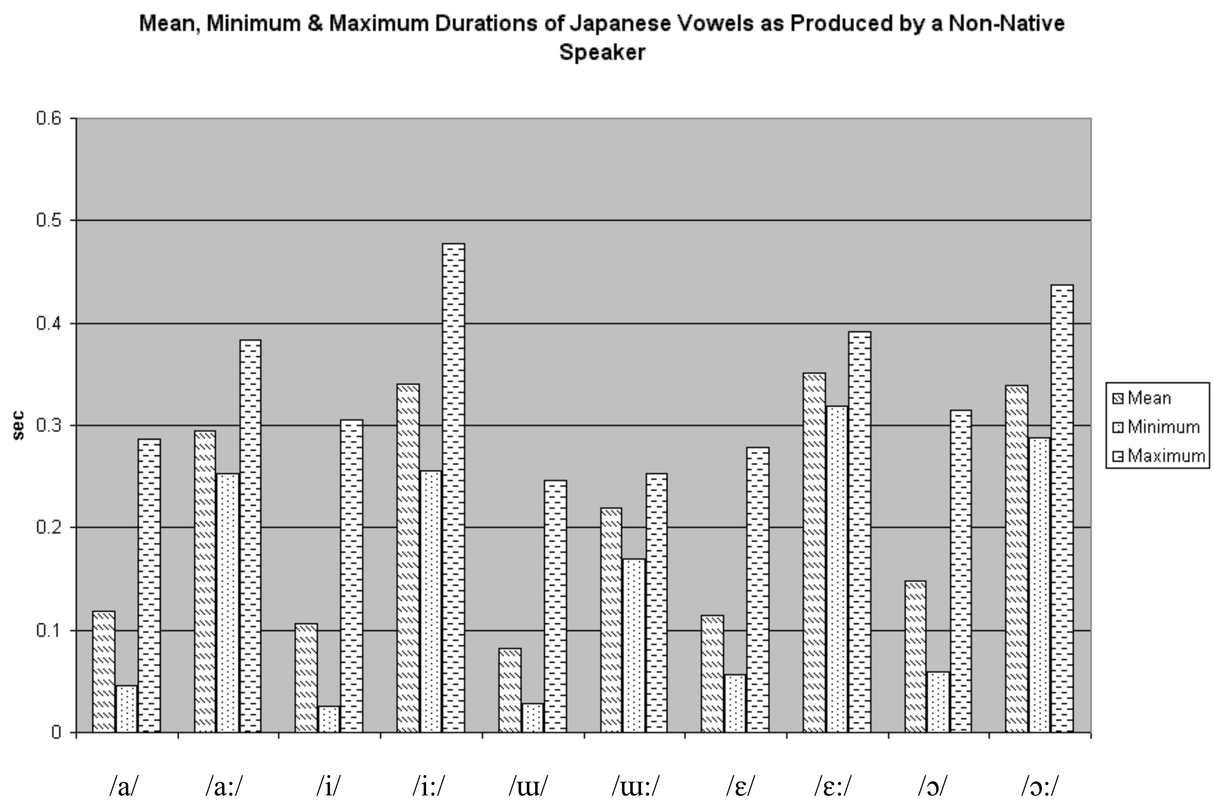
Table 8: Mean Vowel Duration Ratios

	Short	Long	Ratio
/a/ vs. /a:/	0.118	0.295	2.50
/i/ vs. /i:/	0.106	0.341	3.22
/ɯ/ vs. /ɯ:/	0.082	0.219	2.67
/ɛ/ vs. /ɛ:/	0.114	0.351	3.08
/ɔ/ vs. /ɔ:/	0.148	0.339	2.29

Overall, the subject is producing a mean vowel duration contrast ratio of 2.65:1.0, which falls well within the reported native speaker range of 2:1 to 3:1 (Han 1962). For all contrasts investigated, a two-tailed *t* test revealed that long vowels were produced significantly longer than short vowels [/a/ vs. /a:/: $t = -9.716$, $p < 0.001$; /i/ vs. /i:/: $t = -5.818$, $p = 0.002$; /ɯ/ vs. /ɯ:/: $t = -9.276$, $p < 0.001$; /ɛ/ vs. /ɛ:/: $t = -12.668$, $p < 0.001$; /ɔ/ vs. /ɔ:/: $t = -8.523$, $p < 0.001$]. This is consistent with predictions made by Brown's model, under which the subject was expected to have little difficulty in implementing the Japanese vowel length contrast since vowels in English are specified for length through the mono-moraic – bi-moraic distinction. These results are also consistent with our explicit linguistic knowledge hypothesis, as the subject was aware of the vowel length contrast in Japanese.

Again, while an examination of the mean ratios appears promising, an examination of the raw data reveals variation in the tokens produced, to a greater extent than that observed among the consonants: vowel duration is highly variable, in that for all vowels except /ɛ/, the longest token of the short vowel is as long or longer than the shortest token of the corresponding long vowel. The graph in (11) serves to illustrate this point.

(11)

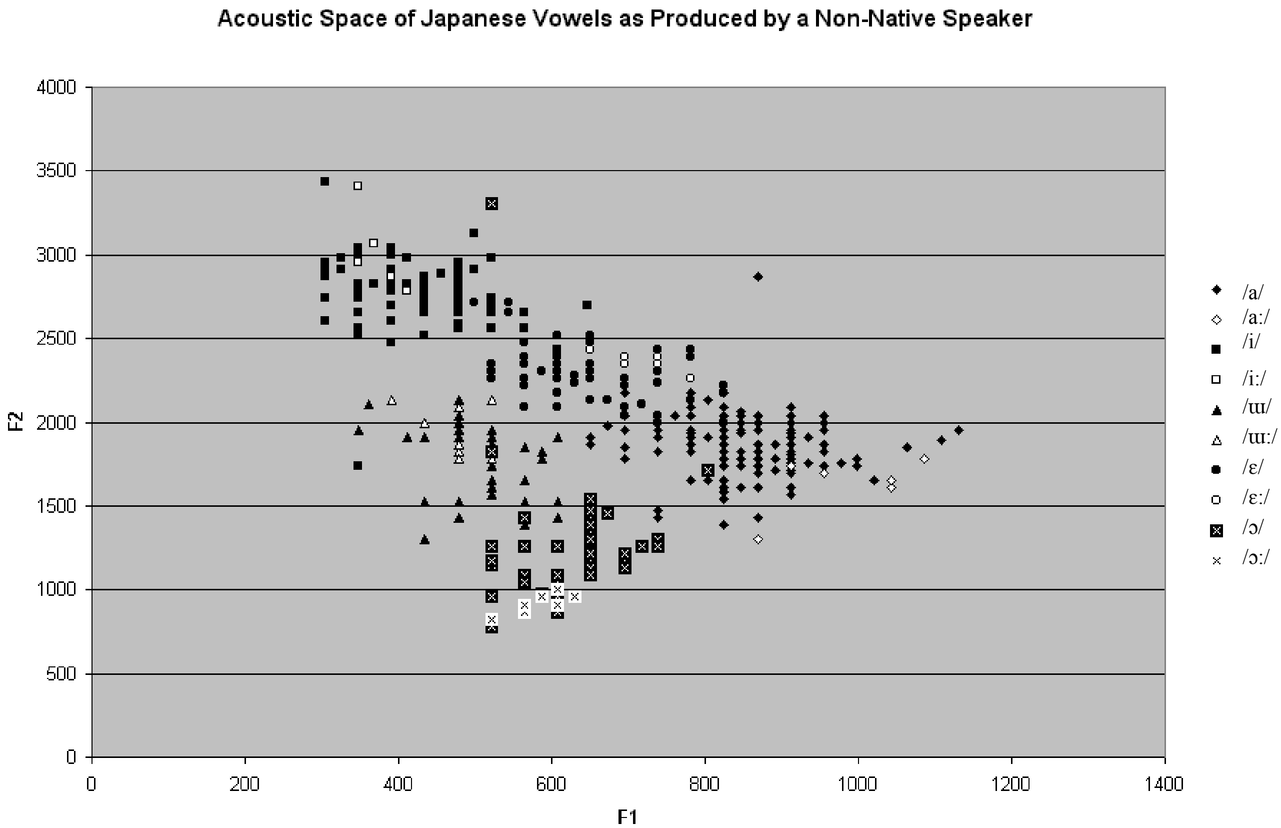


Although examination of the mean ratios yielded native-like results, this variability suggests that the subject is experiencing difficulty in controlling vowel length. Once again, this suggests a lack of native-like timing control of the L2 length contrast, a situation that our explicit linguistic knowledge hypothesis predicts would be remedied at the second recording.

5.3. Vowel Quality

The scatter plot presented in (12) details the subject's vowel productions in acoustic space.

(12)



The mean F1 and F2 frequencies are given in Table 9 below, as well as the mean F1 and F2 frequencies reported for a female native speaker of Japanese (Homma 1973).

Table 9: Mean F1 and F2 of Japanese Vowels

	F1 - Native	F1 – Non-Native	F2 - Native	F2 – Non-Native
/a/	1046	855	2075	1870
/a:/	1046	985	2075	1630
/i/	354	413	2886	2783
/i:/	354	376	2886	2991
/ɯ/	367	507	2060	1772
/ɯ:/	367	473	2060	1941
/ɛ/	655	653	2209	2309
/ɛ:/	655	717	2209	2361
/ɔ/	659	625	977	1208
/ɔ:/	659	579	977	912

As is made clear by Table 9, the subject's Japanese vowels differed from those produced by a native speaker. Another important difference between native speaker performance and the subject's performance was noted: the subject produced significantly different mean vowel qualities for short vowels and their corresponding long vowels. Table 10 serves to summarize the results of statistic analysis. A two-tailed *t* test revealed that with the exception of the /ɛ/ vs. /ɛ:/ contrast, all long vowels were pronounced with a significantly different quality than their corresponding short vowels[2]. This is clearly L1 interference, as English vowels are differentiated using both spectral and temporal cues, and these results serve as counterevidence to our explicit linguistic knowledge hypothesis, as it does not predict this spectral differentiation between long and short vowels. These results are in keeping, however, with Brown's (2000) model, which predicts that the subject will perform well with respect to the vowel contrasts at advanced stages of acquisition, but may exhibit difficulty at earlier stages.

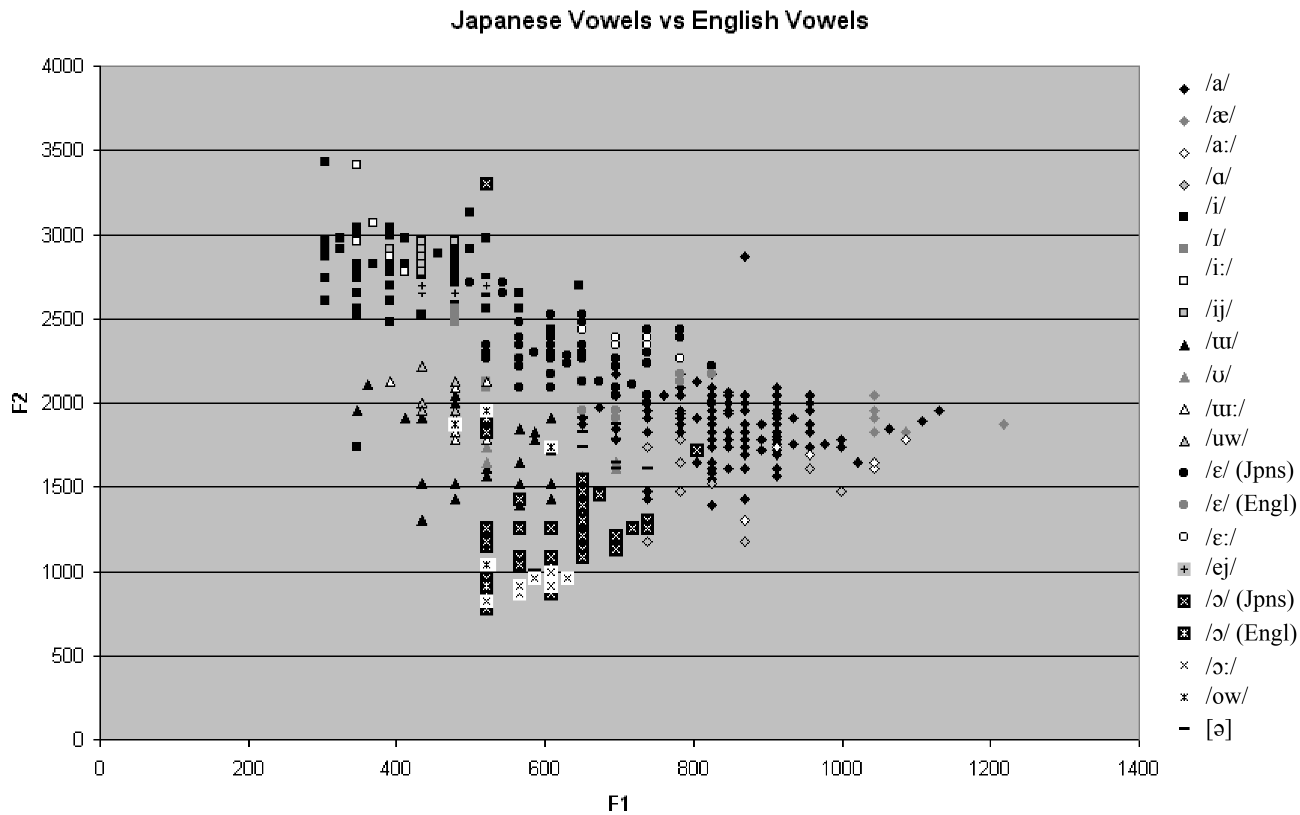
Table 10: Results of Two-Tailed t Test for Japanese Vowel Quality – Time I

Vowel Contrast		t value	p value
/a/ vs. /a:/	F1	-3.825	p < 0.001
	F2	3.464	0.001
/i/ vs. /i:/	F1	2.860	0.016
	F2	-2.663	0.009
/ɯ/ vs. /ɯ:/	F1	1.569	0.123
	F2	-2.285	0.027
/ɛ/ vs. /ɛ:/	F1	-1.795	0.077
	F2	-0.845	0.401
/ɔ/ vs. /ɔ:/	F1	3.256	0.004
	F2	2.504	0.015

The vowel quality data gathered were also compared against the F1 and F2 of Canadian English vowels to check for substitution of these in the subject's Japanese productions: Japanese short vowels were compared against English lax vowels; Japanese long vowels were compared against English tense vowels; this is illustrated in (13) below.

[2] Significance for vowel quality was determined on the following basis: if either or both of the F1 and F2 values reached significance, then the long vowel was deemed to be of a significantly different quality than its short counterpart. Since both the F1 and F2 are used in identifying the vowel in acoustic space, a significant change to either or both would result in the production of a different vowel.

(13)



Results of this analysis are mixed, and do not wholly support or reject either of our hypotheses. While most Japanese vowels differed significantly from their English counterparts, the contrasts /ɯ/ vs. /ʊ/ and /ɯ:/ vs. /uw/ did not, suggesting that the subject is substituting rounded back vowels from English for the Japanese unrounded targets. This finding is consistent with Brown's model, which predicts that /ɯ/ and /ɯ:/ should not be acquirable to our subject, as English high back vowels are not specified for roundedness (and are thus rounded by default); however, it also poses a serious problem for our explicit linguistic knowledge hypothesis, which claims that since the subject is aware of unrounded back vowels in Japanese, she should be able to produce them. Table 11 details the results of statistical analysis.

Table 11: Results of Two-Tailed t Test for Japanese and English Vowels – Time I

Vowel Contrast		t value	p value
/a/ vs. /æ/	F1	-6.647	$p < 0.001$
	F2	-0.509	0.611
/aː/ vs. /ɑ/	F1	3.245	0.006
	F2	1.111	0.287
/i/ vs. /ɪ/	F1	-5.309	$p < 0.001$
	F2	7.905	$p < 0.001$
/iː/ vs. /ij/	F1	-5.437	$p < 0.001$
	F2	0.997	0.362
/ɯ/ vs. /ʊ/	F1	-2.114	0.083
	F2	1.032	0.307
/ɯː/ vs. /uw/	F1	0.923	0.373
	F2	-1.209	0.248
/ɛ/ vs. /ɛ/	F1	-2.386	0.020
	F2	4.161	$p < 0.001$
/ɛː/ vs. /ej/	F1	10.716	$p < 0.001$
	F2	-11.939	$p < 0.001$
/ɔ/ vs. /ɔ/	F1	2.673	0.010
	F2	0.890	0.377
/ɔː/ vs. /ow/	F1	2.603	0.022
	F2	-2.908	0.033

5.4. Remarks

Given that at the time of recording, the subject was aware only of the existence of geminate consonants and long vowels in Japanese, and not of her performance with respect to that of native speakers, the results of the analysis of duration are not surprising under our hypothesis that it is explicit linguistic knowledge that allows the learner to bypass L1 knowledge in building new structure for the L2. The subject appropriately produced geminate consonants and long vowels where required; however, some troubling variation in duration was found, particularly among the vowels.

The results of the analysis of vowel quality are intriguing, as the vowels produced neither wholly resembled those produced by native speakers of Japanese, nor did they wholly resemble the subject's own production of English vowels.

The next section aims to determine if the subject was able to make use of any linguistic knowledge gained from examination of the first recording session.

6. Results – Time II

6.1. Consonant Closure Duration

The graph in (14) below illustrates the subject's mean timing control of single and geminate consonants; Table 12 details the timing control ratios produced, as illustrated in (14).

(14)

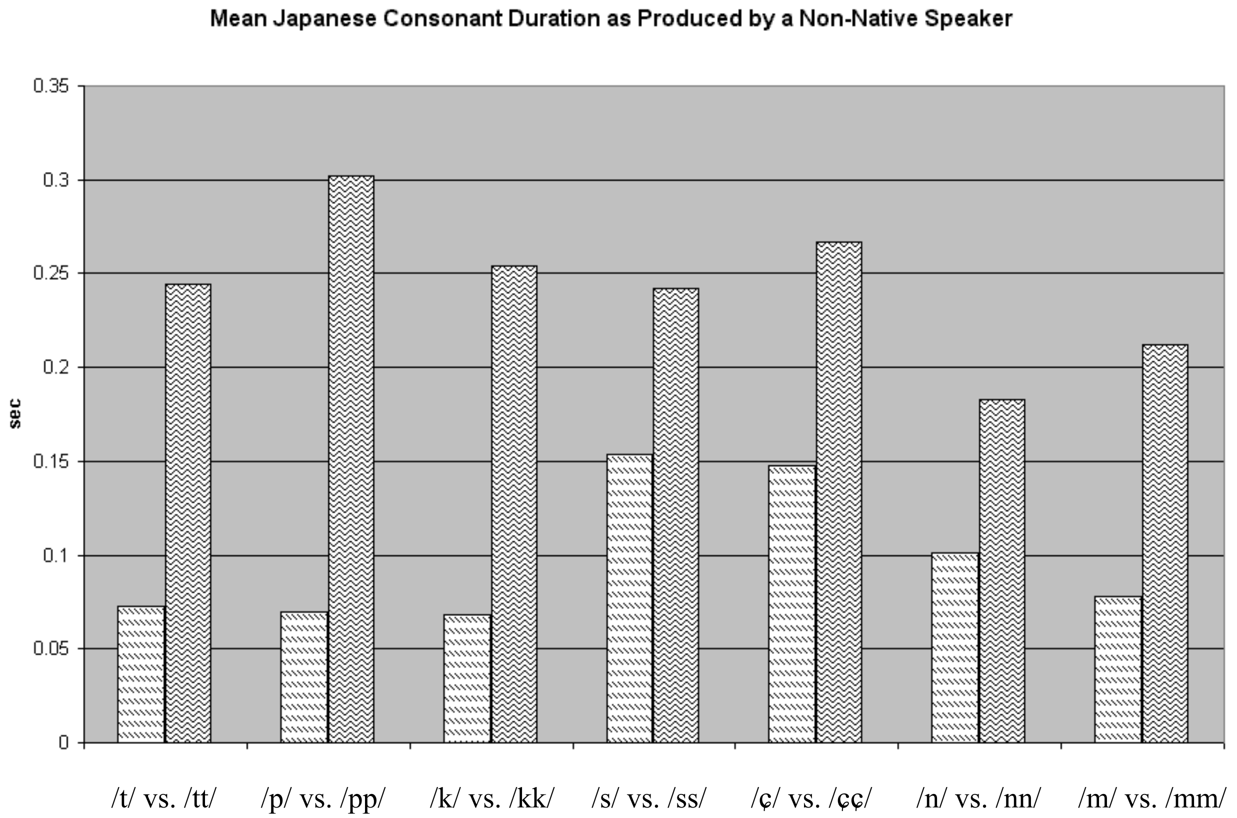


Table 12: Mean Consonant Closure Duration Ratios

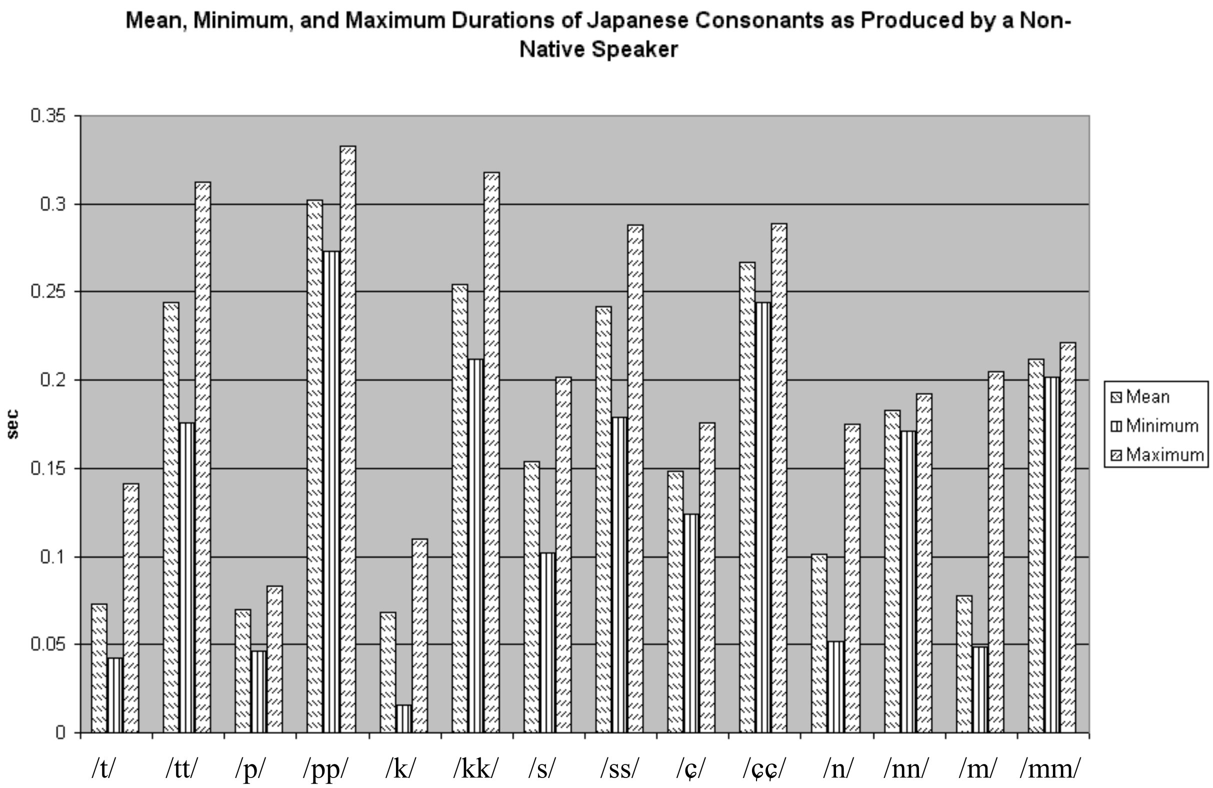
	Single	Geminate	Ratio
/t/ vs. /tt/	0.073	0.244	3.34
/p/ vs. /pp/	0.070	0.302	4.13
/k/ vs. /kk/	0.068	0.245	3.74
/s/ vs. /ss/	0.154	0.242	1.57
/ç/ vs. /çç/	0.148	0.267	1.80
/n/ vs. /nn/	0.101	0.183	1.81
/m/ vs. /mm/	0.078	0.212	2.72

* Recall that an additional sentence was recorded at Time II in order to obtain the /s/ vs. /s:s/ contrast, which had been omitted from Time I due to vocabulary constraints.

Overall, the subject produced a mean closure duration ratio of 2.59:1.0. A two-tailed *t* test revealed that geminate consonants were produced with a significantly longer closure duration than their corresponding single consonants [/t/ vs. /tt/: $t = -15.691$, $p < 0.001$; /p/ vs. /pp/: $t = -18.836$, $p < 0.001$; /k/

vs. /kk/: t = -21.076, $p < 0.001$; /s/ vs. /ss/: t = -6.772, $p < 0.001$; /ç/ vs. /çç/: t = -16.953, $p < 0.001$; /n/ vs. /nn/: t = -4.529, $p < 0.001$; /m/ vs. /mm/: t = -8.876, $p < 0.001$], suggesting once more that the subject has indeed acquired new phonological structure enabling her to deal with the Japanese length contrast in a consistent and significant manner. Once again, performance groups coincide with natural sound classes: among the obstruents, the timing ratio patterns differently for stops (larger ratio than that reported for native speakers) than it does for continuants (smaller ratio than that reported for native speakers). Again, an inspection of the raw data reveals some variation among the tokens produced, illustrated in the graph given in (15).

(15)



6.2. Vowel Duration

The graph in (16) illustrates the subject's mean timing control of vowel duration; Table 13 details the timing control ratios produced illustrated in (16).

(16)

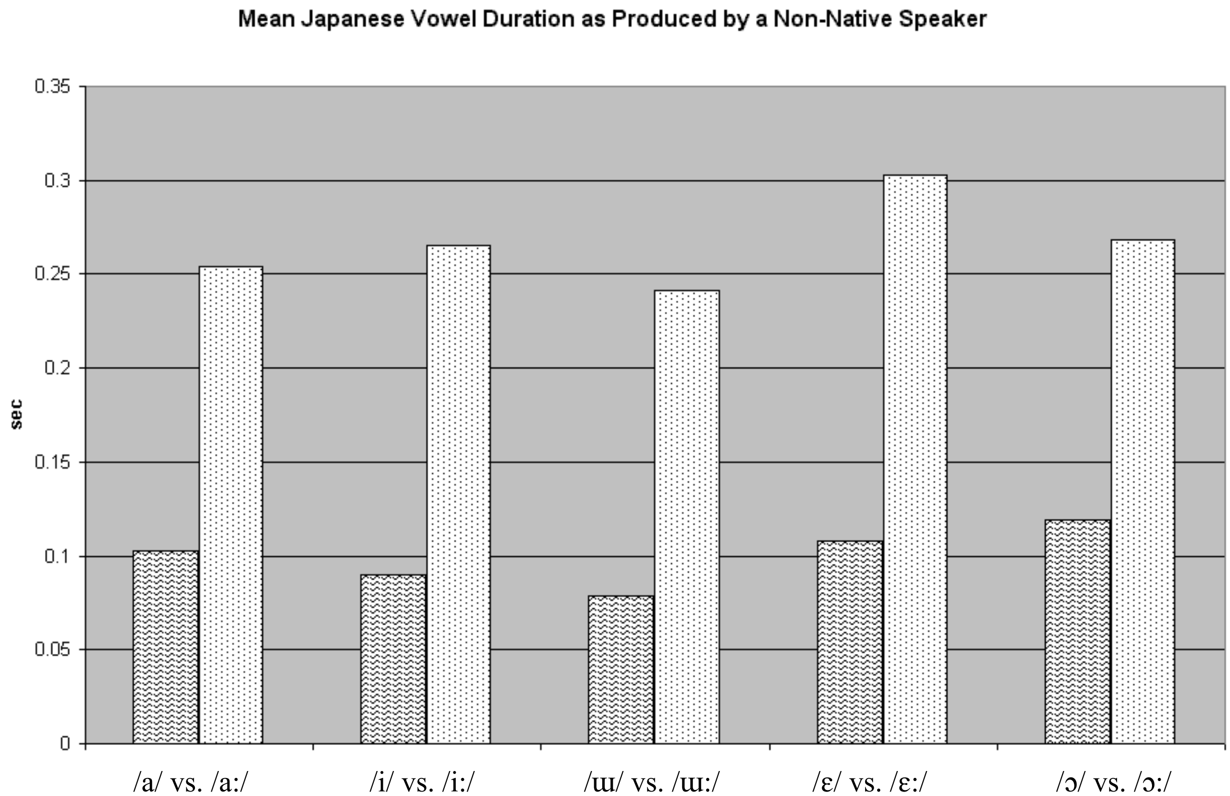
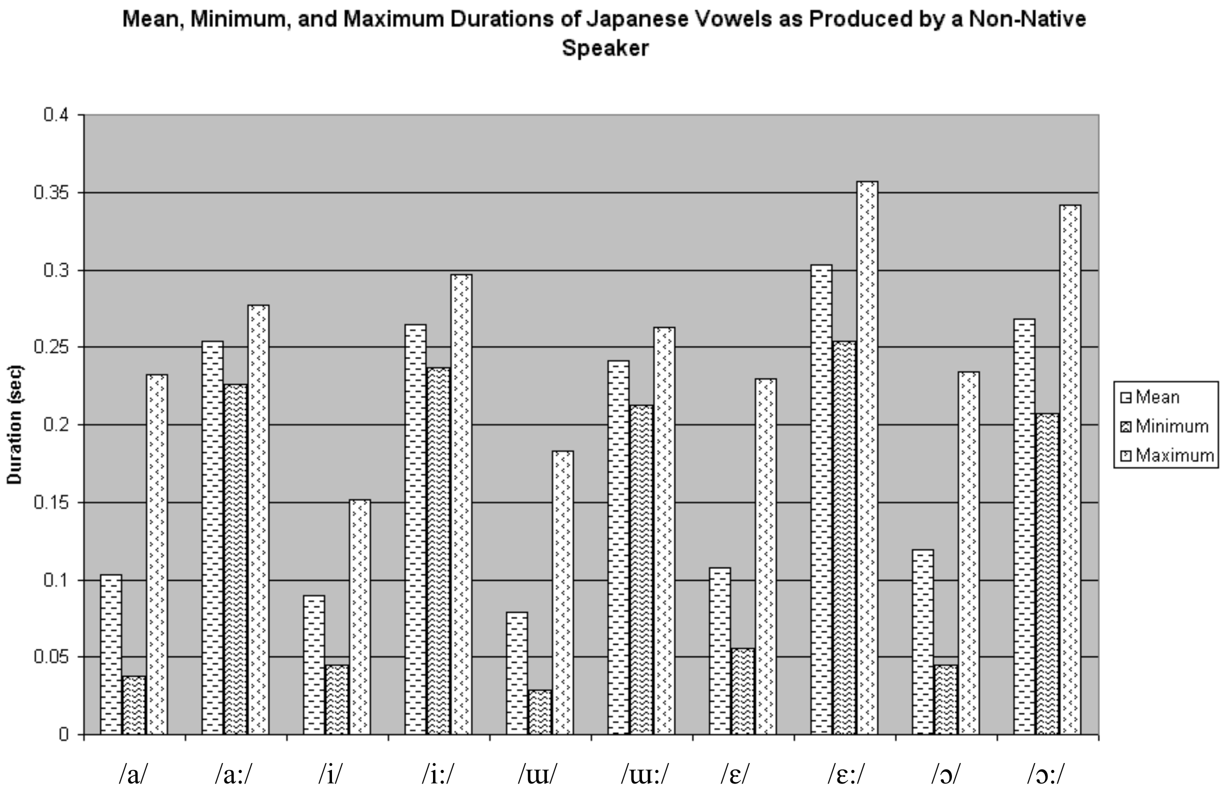


Table 13: Mean Vowel Duration Ratios

	Short	Long	Ratio
/a/ vs. /a:/	0.103	0.254	2.47
/i/ vs. /i:/	0.090	0.265	2.94
/ɯ/ vs. /ɯ:/	0.079	0.241	3.05
/ɛ/ vs. /ɛ:/	0.108	0.303	2.81
/ɔ/ vs. /ɔ:/	0.119	0.268	2.25

Overall, the subject is producing a vowel duration contrast ratio of 2.63:1.0. A two-tailed *t* test revealed that long vowel duration was significantly longer than the corresponding short vowel [/a/ vs. /a:/: $t = -10.695$, $p < 0.001$; /i/ vs. /i:/: $t = -16.203$, $p < 0.001$; /ɯ/ vs. /ɯ:/: $t = -12.921$, $p < 0.001$; /ɛ/ vs. /ɛ:/: $t = -12.587$, $p < 0.001$; /ɔ/ vs. /ɔ:/: $t = -10.002$, $p < 0.001$]. The problematic variation that was found at Time I is still present, however; the graph in (17) illustrates this point.

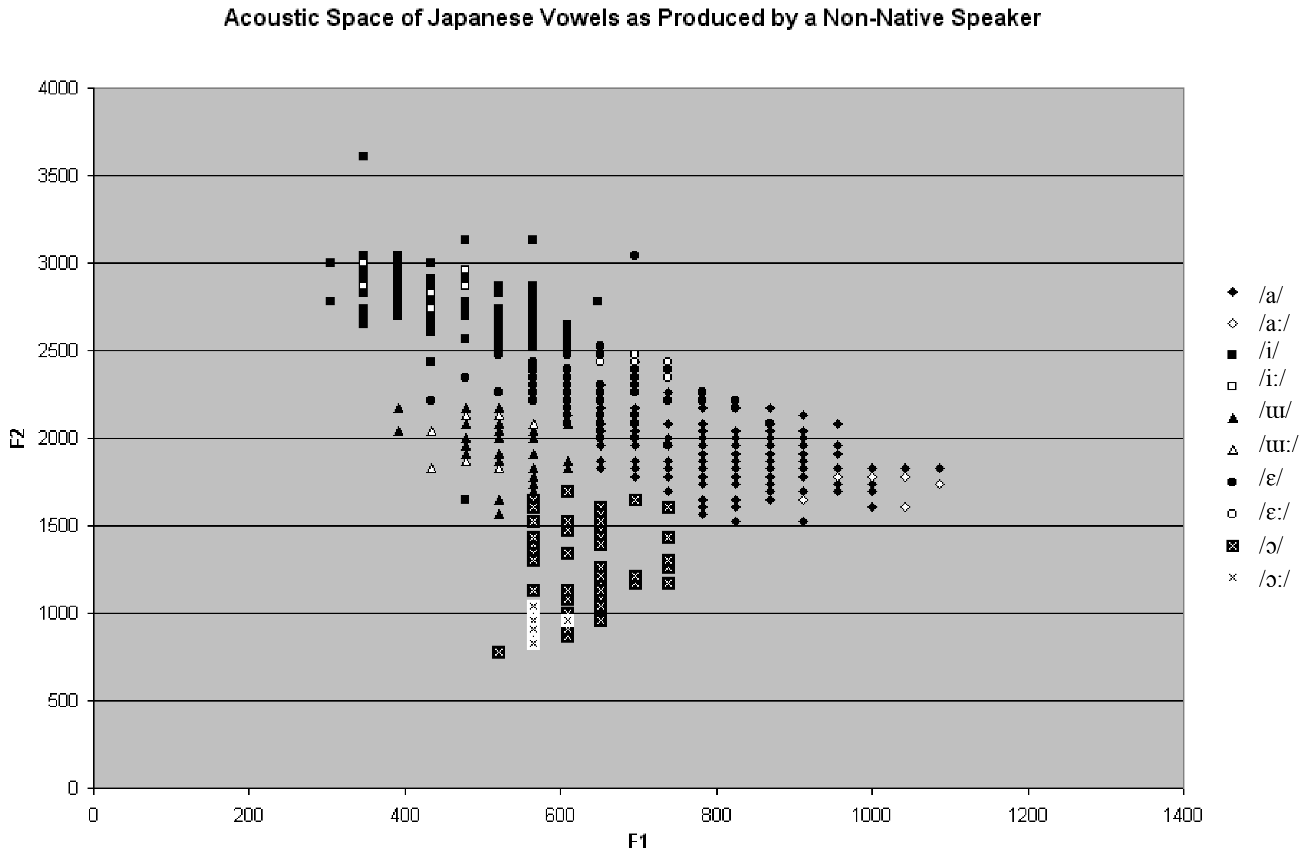
(17)



6.3. Vowel Quality

The scatter plot presented in (17) details the subject's vowel productions in acoustic space.

(17)



The mean F1 and F2 frequencies are given in Table 14 below, along with the mean F1 and F2 frequencies produced by a native speaker.

Table 14: Mean F1 and F2 of Japanese Vowels

	F1 – Native	F1 – Non-Native	F2 – Native	F2 – Non-Native
/a/	1046	833	2075	1893
/a:/	1046	1007	2075	1724
/i/	354	454	2886	2775
/i:/	354	420	2886	2876
/ɯ/	367	529	2060	1914
/ɯ:/	367	502	2060	1980
/ɛ/	655	641	2209	2292
/ɛ:/	655	702	2209	2427
/ɔ/	659	628	977	1244
/ɔ:/	659	575	977	927

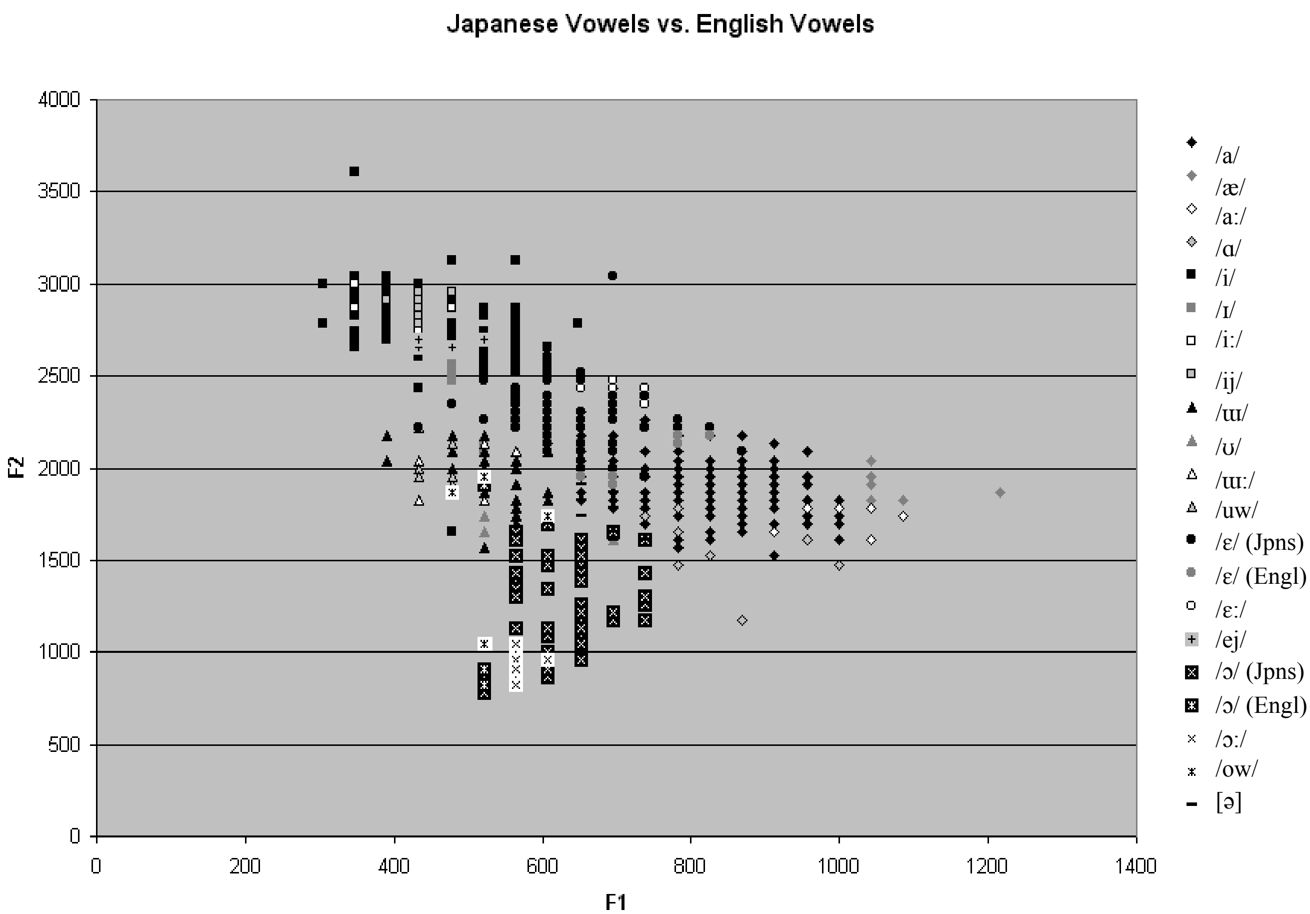
A two-tailed *t* test revealed that /a/, /ɛ/, and /ɔ/ were all pronounced with a significantly different quality than their long counterpart, while /i/ and /ɯ/ showed no significant difference between long and short counterparts. Table 15 below details the results of the statistic analysis.

Table 15: Results of Two-Tailed t Test for Japanese Vowel Quality – Time II

Vowel Contrast		t value	p value
/a/ vs. /a:/	F1	-4.642	p < 0.001
	F2	2.888	0.004
/i/ vs. /i:/	F1	1.049	0.297
	F2	-1.184	0.239
/ɯ/ vs. /ɯ:/	F1	1.408	0.165
	F2	-0.946	0.348
/ɛ/ vs. /ɛ:/	F1	-1.949	0.055
	F2	-5.057	p < 0.001
/ɔ/ vs. /ɔ:/	F1	5.686	p < 0.001
	F2	7.783	p < 0.001

The measurement data were plotted against the data for English vowels to check for substitutions, giving the scatter plot in (18).

(18)



A two-tailed *t* test revealed that with the exception of /i:/ and /ɯ:/, all Japanese vowels were pronounced with a significantly different quality than their corresponding English vowels. This suggests that the subject did not substitute English vowels for Japanese vowels in her productions, and is instead attempting to modify her vowel space so that her Japanese vowels more closely resemble those of a native speaker. The full details of statistic analysis are listed in Appendix B.

6.4. Comparing Time I to Time II

6.4.1. Consonant Closure Duration

There was no universal change in consonant closure duration from the first recording session to the second; a two-tailed *t* test revealed that many of the changes were in fact not significant. The exceptions that were found will be discussed here, the full details of the statistic analysis are listed in Appendix B.

While the change in closure duration of /t/ was not significant [$t = 1.665$, $p = 0.102$], the change in closure duration of /tt/ was found to be significantly shorter at Time II [$t = 4.202$, $p < 0.001$]. This suggests that the subject is adjusting the timing control of the geminate stop in order to obtain a more native-like ratio (3.91:1.0 at Time I vs. 3.34:1.0 at Time II). Conversely, both /p/ and /pp/ were found to be significantly shorter at Time II [/p/: $t = 3.050$, $p = 0.008$; /pp/: $t = 3.147$, $p = 0.010$], suggesting a general adjustment to the timing control of closure duration for the bilabial stop, resulting in only a slight detrimental change to the mean ratio (4.00:1.0 at Time I vs. 4.13:1.0 at Time II). A third scenario presents itself in the case of /k/ vs. /kk/ and /ç/ vs. /çç/: here, the single consonant was found to be significantly shorter at Time II [/k/: $t = 3.717$, $p < 0.001$; /ç/: $t = 2.878$, $p = 0.007$], while the change in the corresponding geminate consonant was not significant [/kk/: $t = 2.168$, $p = 0.055$; /çç/: $t = 2.213$, $p = 0.051$]. In both cases, however, the mean ratio is only slightly improved by the change (/k/ vs. /kk/: 3.87:1.0 at Time I vs. 3.74:1.0 at Time II; /ç/ vs. /çç/: 1.78:1.0 at Time I vs. 1.80:1.0 at Time II).

In spite of the improvement seen with respect to the /t/ vs. /tt/ contrast, these results suggest that the subject was unable to make use of information about her performance from the first recording session in the later recording. Her timing control of the geminate – single consonant contrast in most cases did not change significantly; where it did, it did not result in a great improvement to the mean ratio.

6.4.2. Vowel Duration

An examination of the changes in vowel duration from the first recording session to the second yielded non-uniform results, much like the analysis for consonant closure duration described in Section 6.4.1. above. A two-tailed *t* test showed that three of the five short vowels were significantly shorter at Time II [/a/: $t = 3.425$, $p = 0.001$; /i/: $t = 3.030$, $p = 0.003$; /ɔ/: $t = 2.868$, $p = 0.005$], whereas two of the five long vowels were found to be significantly shorter at Time II [/ɛ:/: $t = 2.628$, $p = 0.025$; /ɔ:/: $t = 3.146$, $p = 0.006$] (Full details of the statistic analysis are presented in Appendix B). These changes brought about slight improvements to the mean vowel duration ratios, as illustrated in Table 17 below; the only exception to this is the increase in the /ɯ/ vs. /ɯ:/ contrast ratio, which is puzzling since neither vowel in this contrastive pair was found to have significantly either increased or decreased in duration.

Table 17: Mean Vowel Duration Ratios, Time I vs. Time II

	Ratio – Time I	Ratio – Time II
/a/ vs. /a:/	2.50	2.47
/i/ vs. /i:/	3.22	2.94
/ɯ/ vs. /ɯ:/	2.67	3.05
/ɛ/ vs. /ɛ:/	3.08	2.81
/ɔ/ vs. /ɔ:/	2.29	2.25

Although slight improvement was observed across two of the five contrastive pairs of vowels, these results still suggest that the subject was unable to make use of insight on her performance gained after the first recording session. This claim is based largely on the fact that both the /ɯ/ vs. /ɯ:/ and /ɔ/ vs. /ɔ:/ contrasts worsened, in opposite directions, with respect to native speaker performance. If the new explicit linguistic knowledge had been available to the subject, we would have expected to see improvement across all five contrasts, not just a select few.

6.4.3. Vowel Quality

Turning now to an examination of changes in vowel quality from one recording session to the next, the results are again varied. A two-tailed *t* test revealed that none of the long vowels underwent a significant change in vowel quality from Time I to Time II. Among the short vowels, three did undergo a significant change, but only one did so in a manner that resulted in a more native-like vowel. The lower-than-native F1 of /a/ was significantly lower at Time II [$t = 2.438$, $p = 0.015$] and the higher-than-native F1 of /i/ was significantly higher at Time II [$t = -3.805$, $p < 0.001$]; while the lower-than-native F2 of /ɯ/ was significantly higher (but still not native-like) at Time II [$t = -3.310$, $p = 0.001$]. The full details of the statistic analysis are presented in Appendix B.

Again, these findings suggest that the subject was not able to make use of insight on her performance gained after the first round of recordings to make her vowel production more like that of a native speaker. Only one vowel was found to have improved, two worsened, and the remainder did not undergo any significant change from one recording session to the next.

7. Discussion

The results of the experiment show that the subject's Japanese productions are not merely English phonemes funnelled into Japanese words and sentences, but they are not native-like either. Indeed, the data collected do show some intriguing properties, which will be addressed here, both in the context of the experiment itself, and in the context of phonological acquisition and L2 acquisition theory. Since our main research goal was to determine if explicit linguistic knowledge facilitated L2 acquisition by allowing the learner to bypass the L1 knowledge, we begin with a discussion of the results from that perspective.

In spite of some variability in the data, the single – geminate consonant contrast produced by the subject did reach significance in statistical analyses, suggesting that the subject has acquired something allowing her to distinguish reliably and consistently between single and geminate consonants. Given that this type of length contrast is not found in English, it is reasonable to assume that the subject's explicit knowledge of length contrasts in Japanese is what enabled her to build the necessary phonological structure to represent them and produce them reliably, even after only four months of classroom exposure to Japanese. When we also consider that the subject in this research outperformed one of Han's (1992) fluent subjects, who did not produce any single – geminate stop contrast at all, our conclusion grows stronger. The behaviour observed for the vowels do not provide any opposition to this conclusion, as the subject was able to produce a significant short – long vowel contrast in her Japanese productions, an outcome that was expected even under the assumption that explicit linguistic knowledge plays no role in L2 phonological acquisition.

How, then, are we to interpret the following observations: although significant, the timing ratio for consonants was never native-like, and both consonants and vowels showed variability in the tokens produced. I propose that the representation of a geminate consonant or long vowel, like all segments, has two parts: a phonological representation, and a phonetic implementation. The phonological representation would contain the abstract information that allows that segment to be distinguished phonemically from other phonemes; in the case of both geminate consonants and long vowels in Japanese, this would be a feature relating to weight. The phonetic implementation, on the other hand, would deal with the aspects pertaining to the specific articulatory details of that segment, such as timing control. The motivation for this distinction is not merely one of convenience: studies of languages that make use of a single – geminate consonant contrast found closure duration is found to be the single distinguishing cue, but the actual timing control ratios employed by the languages may differ. Thus, while Bengali and Italian both employ a 2:1 timing ratio, Turkish patterns with Japanese in making use of a 3:1 ratio (Esposito and Di Benedetto 1999; Lahiri and Hankamer 1988). Indeed, Han (1992) also distinguishes between these two levels, arguing that of the four fluent American subjects, one is evidencing a phonological difficulty and thus is unable to produce a single – geminate stop contrast, while the remaining three are evidencing a phonetic difficulty and thus are producing a contrast, but not one that resembles that produced by native speakers. It is reasonable to assume a similar situation for vowels. This distinction, then, between phonology and phonetics, allows us to account for the variability not only in the timing ratios produced but also the variability in tokens of a single segment: the subject is struggling with the phonetic representation of geminate consonants and long vowels. This may be rooted in a difficulty in perceiving the finer details of segment length, which would allow us to account for the differences in performance on stops, fricatives, and vowels. The subject may be relying on tactile cues to determine segment length: in the articulation of a stop, there is definite contact between the tongue and the articulators; a fricative is articulated with less contact; vowels are articulated with no contact at all. This accounts for the greater variability in vowel production compared with consonant production and the greater variability in fricative production compared with stop production.

Thus, explicit linguistic knowledge has been shown to be useful in building novel phonological representations, but not in developing novel phonetic implementation for the newly acquired phonological representations. This conclusion is further supported by the fact that production was not seen to improve from Time I to Time II, despite the presentation of explicit linguistic knowledge with reference to the phonetic details of Japanese length contrasts and the shortcomings of the performance at Time I: although the subject was aware of the timing ratios for consonants and vowels produced by native speakers and how her productions fell short of native-like levels at Time I, she was unable to adjust her performance so that it more closely resembled that of a native speaker at Time II. Therefore, explicit linguistic knowledge cannot be the only factor in attaining native-like proficiency in a second language.

Turning now to vowel quality, the subject was observed to produce vowels that were significantly different from her own English vowels, yet they were not like those produced by native Japanese speakers in a crucially important way: short vowels were found to differ significantly from their corresponding long vowels. Although the subject is not directly substituting English vowels in place of Japanese vowels (with the important exception of /ʊ/ for /ɯ/ and /uw/ for /ɯː/, a substitution that is expected under Brown’s (2000) model but not under Han’s (1992) explicit linguistic knowledge proposal), she seems unable to make use of only one distinguishing acoustic cue in Japanese where English makes use of two. This is in sharp contrast with her ability to make use of a novel distinguishing cue in the case of geminate consonants. Furthermore, the subject was aware of the details of the Japanese vowel system, so under our explicit linguistic knowledge hypothesis she should have been able to accurately produce short vowels that were of the same vowel quality as the corresponding long vowels. We cannot, however, attribute this difficulty to the phonetic level as we did with variability of vowel duration; rather, the subject seems unable to merge what would be two vowel qualities in English into one in Japanese, and as such represents a true phonological difficulty: the subject differentiates between long and short vowels along a dimension that is not used by native speakers. Contrastively, the substitution of the rounded high back vowels of English for the unrounded high back vowels of Japanese can be accounted for as a phonetic difficulty, as roundedness is not contrastive for back vowels in either Japanese or English, therefore the subject is not motivated to build novel phonological structure and instead treats roundedness as a phonetic detail, which we have argued is not improved with explicit knowledge. Given these results, it appears that vowel quality is a domain that is not well addressed by the present explicit knowledge model and requires further research in order to fully explain the behaviour observed here. It may be the case that the problem is linked more closely to a relative difficulty of merging two phonological categories of the L1 into one single phonological category in the L2: that is, suppressing an L1 contrast that is not active in the L2. More research is required to fully examine this possibility, as the present research was aimed at examining the acquisition of new contrasts, and not the suppression of existing ones.

On the whole, the results of this research indicate that explicit linguistic knowledge is useful to the L2 learner in that it allows the learner to bypass his L1 knowledge and build novel phonological structure to accommodate the contrasts of the L2; it is not, however, useful in attaining native-like phonetic control of those contrasts. This is not the result expected under Han's (1992) proposal, which predicted that learners should be able to produce a native-like timing control ratio of geminate and single consonants if they receive explicit instruction on the nature of the timing control. Nor is this the result expected under Brown's (2000) model, which predicted poor performance on the single – geminate consonant contrasts at all stages of acquisition due to the absence of the feature for contrastive consonant length in English, thus predicting that the Japanese consonant length contrast should be impossible for the subject to acquire. This is not to say that we should dismiss Brown's (2000) model, as it does make accurate predictions about those segments in the L2 that will present a difficulty to L2 learners with different L1s. Nor do we want to dismiss Han's (1992) proposal, since it was borne out to some degree in the results of this study. Instead, the findings presented here suggest that the claims made by Brown's (2000) model about the impossibility of acquisition of certain L2 contrasts are too strong, and that it is possible for L2 learners to overcome phonological difficulties that are caused by the L1, even at very early stages of acquisition. Likewise, Han's (1992) proposal is shown to be too strong, and explicit instruction alone will not give rise to native-like performance. I would like to propose a compromise: explicit linguistic knowledge of the novel contrast will enable the learner to acquire the contrast by assisting in building the required phonological structure, but it cannot fill in the phonetic details.

If explicit linguistic knowledge is only helpful in acquisition of the phonological side of the contrast, how then is it possible for learners to acquire native-like phonetic control of the contrast? Bongaerts et. al's (2000) findings force us to acknowledge that it must be possible, yet the results presented here clearly demonstrate that explicit linguistic knowledge is not the key to success in this domain. Although more research is needed in this area, I would tentatively suggest that it is a matter of experience with the L2 that determines ability in this domain. The subject in this research performed poorly in the phonetic domain, but under this suggestion, this can be attributed to lack of experience with Japanese: at Time II, she had had only six months of classroom exposure to Japanese, which is an extremely short time in the process of language acquisition. The subjects in Han's (1992) and Bongaerts et. al's (2000) research were all advanced learners who had acquired their L2 as adults, yet only those in Bongaerts et. al (2000) were reported to perform at a level that was indistinguishable from native speakers. The crucial difference is the amount of time spent either studying the L2 or being immersed in it. Bongaerts et. al (2000) report that their star subjects had lived in the Netherlands for several years (unfortunately, no specific time frame is specified) and communicated almost exclusively in Dutch (the L2) with their families and in the work place. The subjects in Han's (1992) study, on the other hand, had spent considerably less time in a Japanese-speaking environment: one had studied Japanese intensively for two months before living in Japan for two years, one had studied Japanese in a university classroom setting for three years before spending one year studying at a Japanese university, one had spent several years living in Japan on an intermittent basis, and the fourth had studied Japanese in a classroom setting for one and a half years before spending a year living in Japan in a university student exchange program. Clearly the two groups differ with respect to the amount of experience each has had with the L2, which has impacted their ability to control the phonetic details of the language (interestingly, Han's (1992) fourth subject displayed a phonological difficulty and was unable to produce any single – geminate stop contrast in testing).

This distinction also works well when applied to the results of research on training in L2 phonological acquisition. Matthews (1997) reports on a study in which subjects whose L1 was Japanese were explicitly trained in the articulation of novel segmental contrasts in English (the L2), then tested on their perceptual abilities of these contrasts (/b/ vs. /v/, /s/ vs. /θ/, /θ/ vs. /f/, and /l/ vs. /ɹ/). In our terms, they were given phonetic training on the contrasts. The results showed improvement, but not across all categories, and not to native-like levels, despite the subject's ability to correctly pronounce the targeted segments during training sessions. Under the analysis presented with this research, the subjects in Matthews' (1997) study were being given explicit linguistic information on articulatory (phonetic) details, which would not be helpful beyond the training session as they would not be helpful in building the phonological structure required to distinguish the novel contrasts.

Of course, the research presented should not be taken as an endpoint for investigations into this area. Of particular interest and importance to the argument drawn from this study would be an investigation into the perceptual abilities of L2 learners with explicit linguistic knowledge of the structures of the L2; the fact that the data was gathered using a reading task allows for the possibility that the subject in this particular study is responding to a visual cue in the Japanese *hiragana* script and artificially lengthening the required segments while the phonological contrast is, in fact, absent from her L2 grammar. Also of interest would be investigations into the range of timing control ratios that are deemed to be native-like by native speakers and their perceptions of differing timing ratios, both greater and lower than those of native speakers, as well as investigations into the acquisition of segmental contrasts following this same model, for although we have argued that the same mechanisms should govern both segmental and prosodic domains of phonological acquisition, this may not be the case in actual fact.

A few cautionary remarks are in order: although the results of this study suggest that explicit linguistic knowledge is useful in L2 phonological acquisition, it is not suggested that only explicit knowledge will assist the L2 learner, nor is it suggested that the presence of explicit linguistic knowledge will guarantee successful acquisition of L2 phonology. The acquisition of a second language is a complex procedure, involving many factors, such as practice and motivation, and not all factors have an equal effect on the final state of the L2 grammar. The purpose of the present research was to determine if explicit linguistic knowledge could be counted as one of these factors: does knowing about the L2 assist the learner in using it?

It should also be noted that the findings of this research are not meant to be interpreted as evidence against the Critical Period Hypothesis, nor is it suggested that all linguists will be good at learning languages. In spite of all the linguistic knowledge about Japanese at her disposal, the subject still did not achieve native-like control of the timing ratios examined, nor was she able to correct known errors in performance. L2 acquisition is still a hard and demanding task, while L1 acquisition is remarkably easy.

Thus, the present study is really just one of many steps in research into ultimate attainment in L2 acquisition. Explicit linguistic knowledge was shown to play an important role in assisting the L2 learner in acquiring novel contrasts which are distinguished by a feature absent from the L1 grammar; however, the results also indicate that explicit linguistic knowledge cannot be the only factor governing L2 phonological acquisition, as the learner was unable to achieve full native-like control of the phonetic details involved in producing the new contrast. More research is needed in order to answer the questions of why and how some adult L2 learners are able to achieve a level of phonological competence in their L2 where their performance is indistinguishable from that of a native speaker.

References

Archibald, J., and G. Libben, eds. 1995. *Research Perspectives on Second Language Acquisition*. Copp Clark Ltd.

Berko-Gleason, J., ed. 1997. *The Development of Language*. Allyn & Bacon.

Birdsong, D. 1999. *Second Language Acquisition and the Critical Period Hypothesis*. Lawrence Erlbaum, 1 – 22.

Bley-Vroman, R. 1989. The logical problem of foreign language learning. In S. Gass and J. Schachter, eds. *Linguistics Perspectives on Second Language Acquisition*. Cambridge University Press, 41 – 68.

Bongaerts, T., S. Mennen and F. van der Slik. 2000. Authenticity of pronunciation in naturalistic second language acquisition: the case of very advanced late learners of Dutch as a second language. *Studia Linguistica* 54 (2), 298 – 308.

Borden, G. J., K. S. Harris, and L. J. Raphael. 1994. *Speech Science Primer: Physiology, Acoustics, and Perception of Speech*. Williams & Wilkins.

Brown, C. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald, ed. *Second Language Acquisition and Linguistic Theory*. Blackwell, 4 – 63.

Burnham, D. K. 1986. Developmental loss of speech perception: exposure to and experience with a first language. *Applied Psycholinguistics* 7, 207 – 240.

Esposito, A., and M. G. Di Benedetto. 1999. Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America* 106 (4), 2051 – 2062.

Friederici, A. D., and J. M. I. Wessels. 1993. Phonotactic knowledge and its use in infant speech perception. *Perception and Psychophysics* 54, 287 – 295.

Goodsitt, J. V., J. L. Morgan, and P. K. Kuhl. 1993. Perceptual strategies in pre-lingual speech segmentation. *Journal of Child Language* 20, 229 – 252.

Han, M. 1994. Acoustic manifestations of mora timing in Japanese. *Journal of the Acoustical Society of America* 96 (1), 73 – 82.

Han, M. 1992. The timing control of geminate and single stop consonants in Japanese: a challenge for nonnative speakers. *Phonetica* 49, 102 – 127.

Han, M. 1962. The feature of duration in Japanese. *Study of Sounds* 10, 65 – 80.

Hohne, E. A., and P. W. Jusczyk. 1994. Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics* 56, 613 – 623.

Homma, Y. 1981. Durational relationship between Japanese stops and vowels. *Journal of Phonetics* 9 (3), 273 – 281.

Homma, Y. 1973. An acoustic study of Japanese vowels: their quality, pitch, amplitude, and duration. *Study of Sounds* 16, 347 - 368.

Jusczyk, P. W. 1997. *The Discovery of Spoken Language.* MIT Press, 73 – 109.

Jusczyk, P. W. 1993a. From general to language specific capacities: the WRAPSA model of how speech perception develops. *Journal of Phonetics* 21, 3 – 28.

Jusczyk, P. W., and R. N. Aslin. 1995. Infants' detection of sound patterns of words in fluent speech. *Cognitive Psychology* 29, 1 – 23.

Jusczyk, P. W., A. Cutler, and N. Redanz. 1993. Preference for the predominant stress patterns of English words. *Child Development* 64, 675 – 687.

Jusczyk, P. W., E. A. Hohne, and D. R. Mandel. 1995. Picking up regularities in the sound structure of the native language. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. York Press, 91 – 115.

Lahiri, A., and J. Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics* 16, 327 – 338.

Long, M. 1990. Maturational constraints on language development. *Studies in Second Language Acquisition* 12 (3), 251 – 286.

Matthews, J. 1997. The influence of pronunciation training on the perception of second language contrasts. In J. Leather and A. James, eds. *New Sounds 97*. University of Klagenfurt Press.

Newport, E. 1990. Maturational constraints on language learning. *Cognitive Science* 14, 11 – 28. Reprinted in P. Bloom, ed. 1994. *Language Acquisition: Core Readings*. MIT Press, 543 – 560.

O'Grady, W., and M. Dobrovolsky. 1996. *Contemporary Linguistic Analysis: an Introduction*. Copp Clark Ltd.

Polka, L., and J. F. Werker. 1994. Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20, 421 – 435.

Port, R. F., J. Dalby, and M. O'Dell. 1987. Evidence for mora timing in Japanese. *Journal of the Acoustical Society of America* 81 (5), 1574 – 1585.

Rice, K., and P. Avery. 1995. Variability in a deterministic model of language acquisition: a theory of segmental elaboration. In J. Archibald, ed. *Phonological Acquisition and Phonological Theory*. Erlbaum, 23 – 42.

Rogers, H. 1991. *Theoretical and Practical Phonetics*. Copp Clark Pitman Ltd.

Spencer, A. 1996. *Phonology*. Blackwell.

White, L. 1989. *Universal Grammar and Second Language Acquisition*. J. Benjamins, 35 – 55.

Appendix A

Test Sentences

1. Sono zassi wa takakatta desu ka.

   "Was that magazine expensive?"

2. Tiisai susiya wa eki no tonari ni arimasu.

   "The small sushi bar is beside the station."

3. Yasumi wa mikka kara youka made desu.

   "The holidays are from the third until the eighth."

4. Kinou ookii zisyo wo kaimasita.

   "Yesterday I bought a large dictionary."

5. Kuroi empitu wo ippon katte kudasai.

   "Please buy one black pencil."

6. Eeto ammari warukunai desu.

   "Well, it's not that bad."

7. Yuubinkyoku dake itte kimasu.

   "I'm only going to the post office (and I'll be right back)."

8. Yuube konnani tempura wo tabemasita.

   "Last night I ate this much tempura."

9. Daigaku no tosyokan wa benri desu nee.

   "The university library is convenient, isn't it?"

10. Canada[3] no kitte ga yon mai irimasu.

    "I need four Canadian stamps."

11. San zi zyuppun mae desu.

    "It's ten minutes before three o'clock."

[3] Productions of *Canada* were not included in analysis due to the obvious English origins of the word; it is likely that the subject would pronounce *Canada* in Japanese in the same manner as she would in English, thus making it unrepresentative of her pronunciation in Japanese.

12. Kinou wa yokka zya nakatta desu. Tooka desita.

"Yesterday wasn't the fourth. It was the tenth."

13. Yasumi wa issyuukan desita.

"The holidays lasted one week."

14. Depaato[4] de furosiki wo katte kite kudasai.

"Please go buy a *furosiki* at the department store (and come right back)."

15. Sumimasen, zyaa mata.

"Excuse me, I'll see you later."

16. Kissaten de hon wo issatu yomimasita.[5]

"I read one book at the coffee shop."

[4] Although this word is borrowed from English (*department store*), it was included in the analysis because it was not deemed likely that the subject would pronounce the word as she would in English.
[5] This sentence was added at Time II in order to obtain measurements on the /s/ vs. /ss/ contrast, which had been absent in the Time I recordings due to vocabulary constraints.

Appendix B

Results of Statistic Analysis

Time I

Consonant Duration	t value	p value
/t/ vs. /tt/	-12.820	p < 0.001
/p/ vs. /pp/	-10.649	p < 0.001
/k/ vs. /kk/	-7.292	0.001
/ç/ vs. /çç/	-8.044	p < 0.001
/n/ vs. /nn/	-7.640	p < 0.001
/m/ vs. /mm/	-12.657	p < 0.001

Japanese Vowels		t value	p value
/a/ vs. /a:/	Duration	-9.716	p < 0.001
	F1	-3.825	p < 0.001
	F2	3.464	0.001
/i/ vs. /i:/	Duration	-5.818	0.002
	F1	2.860	0.016
	F2	-2.663	0.009
/ɯ/ vs. /ɯ:/	Duration	-9.276	p < 0.001
	F1	1.569	0.123
	F2	-2.285	0.027
/ɛ/ vs. /ɛ:/	Duration	-12.668	p < 0.001
	F1	-1.795	0.077
	F2	-0.845	0.401
/ɔ/ vs. /ɔ:/	Duration	-8.523	p < 0.001
	F1	3.256	0.004
	F2	2.504	0.015

Japanese Vowels vs. English Vowels		t value	p value
/a/ vs. /æ/	F1	-6.647	p < 0.001
	F2	-0.509	0.611
/a:/ vs. /ɑ/	F1	3.245	0.006
	F2	1.111	0.287
/i/ vs. /ɪ/	F1	-5.309	p < 0.001
	F2	7.905	p < 0.001
/i:/ vs. /ij/	F1	-5.437	p < 0.001
	F2	0.997	0.362
/ɯ/ vs. /ʊ/	F1	-2.114	0.083
	F2	1.032	0.307
/ɯ:/ vs. /uw/	F1	0.923	0.373
	F2	-1.209	0.248
/ɛ/ vs. /ɛ/	F1	-2.386	0.020
	F2	4.161	p < 0.001
/ɛ:/ vs. /ej/	F1	10.716	p < 0.001
	F2	-11.939	p < 0.001
/ɔ/ vs. /ɔ/	F1	2.673	0.010
	F2	0.890	0.377
/ɔ:/ vs. /ow/	F1	2.603	0.022
	F2	-2.908	0.033

Time II

Consonant Duration	t value	p value
/t/ vs. /tt/	-15.691	p < 0.001
/p/ vs. /pp/	-18.836	p < 0.001
/k/ vs. /kk/	-21.076	p < 0.001
/s/ vs. /ss/	-6.772	p < 0.001
/ɕ/ vs. /ɕɕ/	-16.953	p < 0.001
/n/ vs. /nn/	-4.529	p < 0.001
/m/ vs. /mm/	-8.876	p < 0.001

Japanese Vowels		t value	p value
/a/ vs. /a:/	Duration	-10.695	$p < 0.001$
	F1	-4.642	$p < 0.001$
	F2	2.888	0.004
/i/ vs. /i:/	Duration	-16.203	$p < 0.001$
	F1	1.049	0.297
	F2	-1.184	0.239
/ɯ/ vs. /ɯ:/	Duration	-12.921	$p < 0.001$
	F1	1.408	0.165
	F2	-0.946	0.348
/ɛ/ vs. /ɛ:/	Duration	-12.587	$p < 0.001$
	F1	-1.949	0.055
	F2	-5.057	$p < 0.001$
/ɔ/ vs. /ɔ:/	Duration	-10.002	$p < 0.001$
	F1	5.686	$p < 0.001$
	F2	7.783	$p < 0.001$

Japanese Vowels vs. English Vowels		t value	p value
/a/ vs. /æ/	F1	-6.579	$p < 0.001$
	F2	-0.191	0.849
/a:/ vs. /ɑ/	F1	4.015	0.001
	F2	2.262	0.041
/i/ vs. /ɪ/	F1	-3.316	0.001
	F2	6.952	$p < 0.001$
/i:/ vs. /ij/	F1	-0.578	0.586
	F2	-0.607	0.552
/ɯ/ vs. /ʊ/	F1	-1.600	0.167
	F2	2.814	0.007
/ɯ:/ vs. /uw/	F1	2.101	0.056
	F2	-0.693	0.501
/ɛ/ vs. /ɛ/	F1	-3.054	0.003
	F2	3.251	0.002
/ɛ:/ vs. /ej/	F1	12.106	$p < 0.001$
	F2	-12.359	$p < 0.001$
/ɔ/ vs. /ɔ/	F1	3.407	0.001
	F2	1.518	0.134
/ɔ:/ vs. /ow/	F1	2.863	0.013
	F2	-2.816	0.036

Time I vs. Time II

Consonant Duration	t value	p value
/t/	1.665	0.102
/tt/	4.202	$p < 0.001$
/p/	3.050	0.008
/pp/	3.147	0.010
/k/	3.717	$p < 0.001$
/kk/	2.168	0.055
/ç/	2.878	0.007
/çç/	2.213	0.051
/n/	0.129	0.897
/nn/	2.416	0.132
/m/	1.782	0.079
/mm/	2.591	0.061

Japanese Vowels		t value	p value
/a/	Duration	3.425	0.001
	F1	2.438	0.015
	F2	-1.435	0.152
/a:/	Duration	1.930	0.082
	F1	-0.497	0.630
	F2	-1.232	0.246
/i/	Duration	3.030	0.003
	F1	-3.805	$p < 0.001$
	F2	0.323	0.747
/i:/	Duration	1.856	0.116
	F1	-1.634	0.133
	F2	1.156	0.275
/ɯ/	Duration	0.430	0.669
	F1	-1.810	0.074
	F2	-3.310	0.001
/ɯ:/	Duration	-1.882	0.078
	F1	-1.362	0.192
	F2	-0.573	0.575
/ɛ/	Duration	0.959	0.339
	F1	0.929	0.354
	F2	0.631	0.529
/ɛ:/	Duration	2.628	0.025
	F1	0.632	0.541
	F2	-2.190	0.053
/ɔ/	Duration	2.868	0.005
	F1	-0.287	0.775
	F2	-0.635	0.527
/ɔ:/	Duration	3.146	0.006
	F1	0.372	0.715
	F2	-0.473	0.643