UNIVERSITY OF CALGARY

The Acquisition of Phonological Features in a Second Language

by

Jennifer Mah

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

DEPARTMENT OF LINGUISTICS

CALGARY, ALBERTA

AUGUST, 2003

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "The Acquisition of Phonological Features in a Second Language" submitted by Jennifer Mah in partial fulfillment of the requirements for the degree of Master of Arts.

Supervisor, John Archibald, Department of Linguistics

Darin Howe, Department of Linguistics

Brian Gill, Department of French, Italian, and Spanish

26 Aug '03

Date

ABSTRACT

The research detailed in this thesis seeks to address two long standing questions in the field of L2 acquisition: the question of whether new structure can be acquired in a second language, and the question of how to account formally for the degree of difficulty of any new structure. With respect to the acquisition of segmental phonology, two existing proposals account for observed acquisition patterns by making reference to phonological features. These proposals are tested empirically using event-related potentials in order to assess and compare the neural activity elicited by native and non-native sounds among monolinguals, and L2 learners at various stages of acquisition. The results indicate that while young children with sufficient L2 exposure may acquire new features, adults cannot, regardless of the duration of L2 study; they do, however, show neural evidence of improving their perceptual abilities with respect to new L2 segments as proficiency improves.

ACKNOWLEDGEMENTS

Although mine is the only name that appears with the title of this thesis, it would be wrong for me to claim that the content of the pages that follow are the result of my own effort alone. Words cannot express just how thankful I am to those who have contributed along the way, but I hope that this section will give them a glimpse of my appreciation.

I must first thank my supervisor, John Archibald, who I am sure has gone above and beyond the call of duty of a supervisor in watching over me and my work. Not only did he supervise this effort, he has also supervised my undergraduate honours work, helped me in presenting some of my work at conferences, funded my ERP training in El Paso, provided me with the opportunity to teach an undergraduate linguistics course, and stepped in to play the part of career advisor when I began to contemplate what might come next upon completion of the M.A. I am eternally grateful to him for the immense role he has played in shaping my mind and my life, and for giving me every chance to explore the possibilities before me. I don't think I could ever say it enough, John: Thank you!

Also deserving of glowing praise are my committee members, Darin Howe and Brian Gill; both provided valuable comments and feedback on this project. Brian also lent me his voice for my experimental stimuli (although I only used a fraction of what was recorded – again, Brian, I apologize!) and helped me in recruiting participants. Darin has been a wonderful source of information about phonological theory, and I also had the pleasure of being his teaching assistant twice. I owe them both many, many thanks.

I would also like to thank Susan Bennett, Reyes Bertolín, Odile Rollin, and Ana Val for their assistance in recruiting participants for the experimental research, and the staff at Neuroscan, Inc., particularly Pete Delgado and Curtis Ponton, for the wonderful training and technical support for the all-important equipment.

The Department of Linguistics at the U of C has really come to feel like a second home over the past two years, and I think that that is due in large part to the faculty, who have taught me so much and always made me feel welcome. In addition to John and Darin, I must thank: Craig Chambers, Michael Dobrovolsky, Martha McGinnis, Robert Murray, Amanda Pounder, Betsy Ritter, Hotze Rullmann, and Doug Walker. I am also deeply grateful to Linda Toth, the department's administrative assistant, for her endless assistance with all things related to university life (and teaching me how to boil water), and to Susan Bennett, who started my academic addiction.

The 'homey' feel of the department is also due to my fellow grad student 'siblings': my most amazing classmates/officemates Jodi Edwards, Rebecca Hanson, and Sheena Van Der Mark; my 'older sibs' Jen Abel, Antonio Gonzalez Poot, Ilana Mezhevich, Cory Sheedy, Teresa Vanderweide, Andrea Wilhelm, Aili You, and Youri Zabbal; my 'younger sibs' Claudia Escartin Ortiz, Elizabeth Stacy, and Miko Summerell; as well as noteworthy 'pseudo-sibs' Heather Bliss, Corey Telfer, and Erika Thrift.

And then there is my family, who believed in me enough to make up for when I didn't. And then there is Luke Tilson, who was and will always be there, through both bad-tempered frustrations and elated epiphanies, just for me. A million times over, thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER FOUR

Chapter 1

A new approach in assessing L2 phonological acquisition

1.1. The challenge of new languages

Learning a language is much like piecing together a large, complex puzzle: language is a rule-governed, structured system composed of several subsystems – phonology, morphology, semantics, and syntax, each of which has its own structure and subdivisions. For successful language acquisition, then, the learner needs to identify all the rules and build appropriate mental representations for all the structures in this very complex system. Further complicating the task of language acquisition is the fact that the input available to the learner is in the form of continuous acoustic strings. There are no overt acoustic markers to indicate the boundaries of those units that are meaningful in analyzing and processing a given utterance, and the associations between sound and meaning are arbitrary: there is nothing in the phonetic form of the sentence *The cheese stands alone* that serves to inform that *cheese* is a noun referring to a solid dairy product, while the larger segment *The cheese* is a noun phrase functioning as the subject of the sentence. The learner hears a continuous sequence of speech sounds and must properly segment it into appropriate chunks. The enormity of the task of language acquisition becomes clear when we consider the outcome of adult second language (L2) acquisition.

Unlike first language (L1) learners, who uniformly attain a native level of proficiency, meaning that they are able to understand and produce an infinite number of L1 sentences, L2 learners display a great deal of variability in proficiency levels in their endpoint grammars; in fact, most do not achieve native-like proficiency in their second language (Bley-Vroman 1989). This situation arises in spite of the fact that L2 learners appear to have an advantage over L1 learners in approaching the task at hand in that many are explicitly taught about the rules and structures of the new language through classroom instruction, whereas infants learning a first language do not receive any such explicit training. Furthermore, some of the structures that are used in the learner's L1 may also be found in his L2, thus providing L2 learners with an additional jump-start in assigning structural representations to the acoustic input. Despite these apparent advantages, the success rate for L2 acquisition is considerably lower than that of L1 acquisition. This fact has led Bley-Vroman (1989) to propose the Fundamental Difference Hypothesis (FDH): second language acquisition is a process that is fundamentally different from first language acquisition, in that the latter is a specialized kind of learning, while the former operates much like other types of general learning. The variable outcome of L2 acquisition suggests a link to other things that are learned with varying outcomes, such as learning how to play chess, or perform mathematical operations. He argues that the greater variation among second language learners, with respect to both developmental paths and success of outcome, suggests that they are using varying strategies in their approach to the second language. According to the FDH, adults acquiring a second language differ from children acquiring a first language along two dimensions: they differ in their assumptions of what a possible grammar is, and they differ in their learning procedure, which allows them to arrive at a grammar based on the available data (Bley-Vroman 1989: 51). For children, Universal Grammar (UG) defines possible grammars, and domain-specific learning procedures enable them to arrive at native competence. For adults, the grammar of the first language defines possible grammars, and general problem-solving systems enable them to construct the second language grammar. Hence the reason that second language acquisition has such variable outcomes is due to the fact that these learners are more constrained in what they can acquire, and they are forced to resort to general learning methods in acquisition, where not all individuals are equally skilled.

White (1989, 2003) is a proponent of the opposite approach to L2 acquisition. She argues that while it may appear attractive to assume that L2 acquisition operates in a fundamentally different way than L1 acquisition, to do so would leave us with some important questions unanswered. Namely, we would need to find a way to account for the fact that interlanguages, which are the intermediate grammars that L2 learners establish and build upon as acquisition proceeds, appear to obey principles of UG. Indeed, if we consider that UG still operates beyond first language acquisition, a number of possibilities present themselves for its precise role: is UG fully available? Is only that portion that is active in the first language accessible? These two possibilities predict different outcomes: the first predicts that second language acquisition should proceed much like, if not in exactly the same fashion as first language acquisition, whereas the second predicts that native-like achievement in a second language should be restricted to those areas for which appropriate components of UG have been activated for the first language.

We must acknowledge that there is a difference between first and second language acquisition, though the precise nature and magnitude of this difference has yet to be determined. Importantly, L2 learners differ from L1 learners in that the former have at their disposal a complete L1 grammar. Analysis of L2 learner errors reveals negative transfer of L1 structures that are inappropriate to the L2 phenomenon being acquired in all domains of language: we find sound substitutions, inappropriate affix placement, incorrect word orderings, and errors with respect to vocabulary usage. Yet some aspects of the L2 are acquired with great ease, resulting in very few errors. This suggests that some parts of a second language are harder to learn than others, thereby raising the question: what formally determines that which will be easy to learn and that which will be difficult?

Broadly speaking, these are indeed the major questions of L2 acquisition that this thesis seeks to address: how does L2 acquisition proceed, and what is the relationship between the L1 grammar and the developing L2 grammar? The discussion and data presented here pertain to the acquisition of segmental phonology; that is, we will be looking at how new sounds are learned in a second language. Before moving into a discussion of our experimental research and results, we must first establish a place for this work in the existing literature by providing an overview of previous research on the acquisition of both L1 and L2 segmental phonology, along with some background information about our chosen research paradigm. Section 1.2 outlines previous work examining the role of the L1 grammar in the perception of non-native speech segments and L2 phonological acquisition. Section 1.3 presents our assumptions about phonological representations and the models of L1 interference that will be tested in the experimental portion of the thesis. Section 1.4 discusses the segments examined here and presents the hypothesis that drives our empirical investigation. Section 1.5 details some recent work in the field of electrophysiology which allowed us to devise the innovative tests reported in subsequent chapters of this thesis.

1.2. Previous approaches to the role of the L1 in perception and L2 phonological acquisition

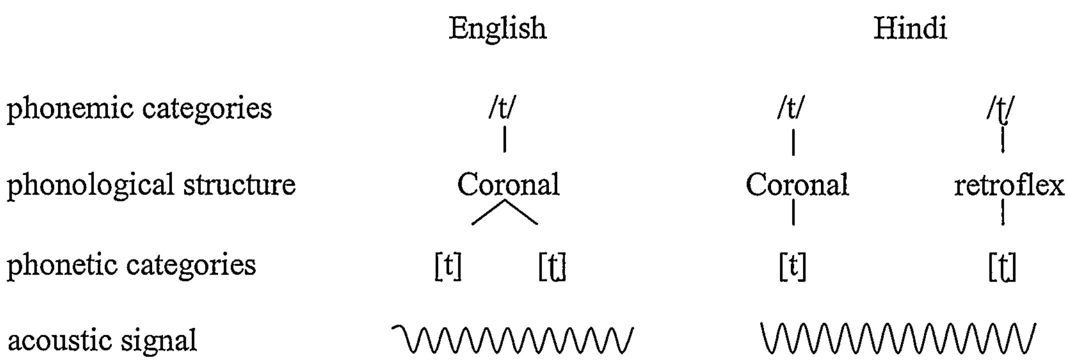
Following rationalist theories of mind and learning (Chomsky 1965), we assume here that all infants are born with some innate linguistic knowledge, or UG: they are born knowing that languages make use of structure in representation, and that these structures are limited in the variety of forms available. What they need to acquire, then, are the particular structures required by the ambient language in their environment. Indeed, a number of studies of infants' perceptual abilities indicate that all infants are born with the ability to perceive all possible segmental contrasts found in human languages, for both consonants and vowels, as any number of these may prove to be required by the phonological system of the ambient language (see Jusczyk 1997 for an overview). The results indicate that while 4-month old infants are able to discriminate all contrasts, this ability undergoes serious decline: by the age of 6 months, infants are only able to discriminate L1 vowel contrasts (Polka and Werker 1994), and by the age of 10 months, they are only able to discriminate among L1 consonant contrasts (Jusczyk 1997). That is, while still in infancy children develop a sensitivity to the phonemic distinctions of the L1.

Thus, by the time a child has reached her first birthday, she has already sorted out which acoustic cues could potentially signal something meaningful in the speech stream, and which ones are to be attributed in perception to speaker variation (such as acoustic differences arising from age and sex of the speaker, the context) or allophony. This is clearly demonstrated through categorical perception, evidenced by infants as well as by adults: individual speech sounds are sorted into categories in such a way that two tokens belonging to the same category tend to be identified as identical sounds (or, two tokens of a single phoneme), whereas two tokens belonging to different categories are identified as different sounds (or, tokens of two separate phonemes) (Best 1995). The speech stream is thus filtered and segments are identified in accordance with the phonemic inventory of the language being acquired; individual phonetic tokens are grouped into categories that correspond to phonemes. As a consequence, performance on tasks that involve making distinctions between sounds belonging to the same category undergo the observed decline described above (Jusczyk 1997).

A study by Werker and Tees (1984) demonstrates that the effects of categorical perception not only persist through adulthood, but they are also language specific; that is, the inability of native English speakers to discriminate various non-native contrasts is not due to a more general inability to discriminate between different types of sounds, but rather due to the fact that the two sounds are grouped into the same single phonemic category in English. This was examined in testing the perceptual abilities of adult native English speakers on two non-native contrasts: a dental vs. retroflex stop contrast (such as the /t/ vs. /ʈ/ contrast found in Hindi), and a plain vs. ejective stop contrast (such as the /k/ vs. /k'/ contrast found in Thompson). They found that while their subjects performed poorly on the non-native contrasts in nonsense syllables, performance increased dramatically when the syllables were truncated so that the resulting test sounds were a series of clicks and noise bursts, and no longer resembled human language. These findings strongly suggest that categorical perception and decreased abilities on within-category discrimination are not the result of general inexperience with various types of sounds. Rather, when it comes to language, we are biased in that we experience any and all speech input in terms of our L1 phonemes.

Brown (1997) proposes a model to account for categorical perception in an approach that views this phenomenon as the mapping of the acoustic speech stream into phonemes. This is schematized in Figure 1.1 below, comparing mapping of speech segments into phonemes of English and Hindi (Brown 1997: 146).

Figure 1.1. Brown's model of acoustic-to-phoneme mapping



Under this view, phonetic categories are identified in the acoustic input and then mapped onto phonemic categories in accordance with available phonological structure. Given the same acoustic input and phonetic categories, then, native English speakers will map these onto a single phonological representation, resulting in the perception of one phonemic category in the input. Native Hindi speakers, on the other hand, will map these onto two distinct phonological representations, thereby allowing them to perceive two phonemic categories in the input.

Given that categorical perception of speech sounds is observed both in young children as well as adults, as previous work indicates, then the identification of L1 phonemes must have a significant impact on L2 phonological acquisition. The diagrams given in Figure 1.1 above suggest that a native speaker of English acquiring L2 Hindi would not be able to perceive the /t/ vs. /t̪/ distinction, as the corresponding phonetic categories [t] and [t̪] extracted from the acoustic signal would be funneled into a single L1 phonemic category. Indeed, L2 learners are often observed to struggle with both the perception (Yamada 1995, Matthews 1997, Brown 2000) and production (Han 1992, Mah 2002) of novel sound contrasts, but recall from discussion above that not all that is new in the L2 is equally hard for learners. With respect to segmental phonology, the finding is that not all new sounds are equally hard. A classic example of this finding is found in L1 Japanese L2 English learners: while these individuals have little or no difficulty in distinguishing some English contrasts, such as /b/ vs. /v/, they are considerably worse at distinguishing other contrasts, such as /l/ vs. /ɹ/ (Brown 2000, Matthews 1997). Early work examining L1 transfer patterns led to the formulation of Contrastive Analysis (Lado 1957) as a predictive and explanatory tool for L2 acquisition: the structures of both the first and second languages were compared and contrasted, and the prediction was that those L2 structures which were similar to existing L1 structures would be easy to learn, whereas those that were different would be harder. Essentially, those structures that could not be transferred to the developing L2 grammar from the L1, and were thus "new", would present difficulty to learners. In the case of segmental acquisition, the phonemic inventories of the L1 and L2 were compared, and those L2 segments that did have a sufficiently similar counterpart in the L1 inventory were considered more difficult to acquire. To return to our L1 Japanese L2 English example, English /p/ vs. /b/ is easy to acquire since both phonemes are also phonemes in Japanese. English /l/ vs. /ɹ/ is difficult because Japanese only has one liquid phoneme in its inventory (see Mah 2002). Contrastive Analysis was heavily criticized, however, as researchers began to find that not all that was new seemed to be hard for L2 learners, and not all that was hard seemed to be equally hard (see Major 2001 for a review). Again returning to our L1 Japanese L2 English learner, a Contrastive Analysis approach predicts that English /b/ vs. /v/, /s/ vs. /θ/, and /l/ vs. /ɹ/ should all be (equally) difficult, yet studies have shown that /b/ vs. /v/ presents little difficulty, /s/ vs. /θ/ is more difficult, and /l/ vs. /ɹ/ is the most difficult[1] (Brown 2000, Matthews 1997). It would appear, then, that L1 transfer alone cannot account for the observed patterns in L2 segmental acquisition.

Flege (1995) proposes the Speech Learning Model (SLM) to account for L1 interference effects. The SLM is primarily concerned with the effect speech perception has on phonological acquisition. In this approach, a mechanism of equivalence classification identifies L2 segments as either "similar" or "different" and is thus responsible for L2 learners' difficulties with new L2 phonemes. In this approach, the

[1] Although the difference in performance on perception tests between /s/ vs. /θ/ and /l/ vs. /ɹ/ for these speakers does not reach statistical significance (Matthews 1997).

notion of "category" applies to a level of representation that is more abstract than that of phonetic categories, but not as abstract as phonemic categories. In order for the learner to establish separate categories for two L2 phonemes, he must first detect the phonetic difference between the two. If both segments under consideration can be assigned to a single existing L1 category, they will be deemed to be equivalent ("similar"), or two tokens of a single phoneme, and no new category will be formed. On the other hand, if a new L2 segment cannot be assigned to any existing L1 category, its phonetic difference from existing categories will be detected ("different"), and an appropriate new category can be built.

Work by Major and Kim (1996) on their Similarity Differential Rate Hypothesis (SDRH) builds on Flege's (1995) SLM, but with a focus on the rate of acquisition of L2 segments. In this approach, "similarity" and "difference" are determined with respect to the allophonic inventory of the L1, such that L2 segments that are phonetically present in the L1 will be considered as being "similar", whereas those that never surface in the L1 will be considered as being "different". Data was gathered from L1 Korean L2 English speakers on a series of production tasks with respect to the English segments /dʒ/ and /z/: [dʒ] is classed as a "similar" L2 segment because it occurs in Korean as an allophone of /tʃ/, whereas [z] is classed as a "different" L2 segment because it never occurs in Korean. Their results showed that L2 learners improve at a faster rate on segments that are different from existing L1 sounds, whereas improvement is slower on segments that are similar to existing L1 sounds. Initial performance was shown to be better on similar segments. Furthermore, their study found that the markedness of the segments in question play a role, where markedness is defined as follows: "A phenomenon A in some language is more marked than B if the presence of A in a language implies the presence of B; but the presence of B does not imply the presence of A." (Eckman 1987). Thus, [dʒ] is more marked than [z] as affricates are more marked than fricatives (Major and Kim 1996); languages whose segmental phonemic inventories contain affricates will also contain fricatives, but we do not find any languages that contain only affricates and no fricatives. We do, however, find languages that contain only fricatives and no affricates, such as French (Walker 2001). Markedness, then, influences the rate of acquisition in such a way that increased markedness of a given segment decreases the rate of acquisition for that segment. In other words, for equal degrees of similarity or difference, it takes L2 learners longer to master more marked segments.

One problem common to both the SLM and the SDRH lies in the fact that both approaches make crucial reference to notions of similarity and difference of segments, yet these notions are either poorly defined (SLM) or the proposed definition cannot be supported empirically (SDRH). For Flege (1995), similarity seems to be determined on the basis of IPA transcription: those sounds which are transcribed using very similar-looking symbols are similar (i.e., [k] and [k'] are similar whereas [k] and [g] are not). For Major and Kim (1996), similarity seems to be determined on the basis of L1 allophones, yet Brown (2000) shows that this definition cannot be maintained: Korean contains two phonetic liquid segments ([l] and [ɾ]), yet L1 Korean L2 English speakers are unable to acquire a phonemic distinction between /l/ and /ɹ/.

A common theme to the approaches to L1 phonological interference discussed thus far is that they all make reference to the segmental level of representation of speech sounds: that is, each segment is an indivisible whole in its interaction with other segments. Given that some form of inadequacy has been identified for each approach considered thus far, it may be the case that the segmental level is inappropriate for developing an account of L1 phonological interference. Phonological theory supplies us with a number of representational levels that can be called upon to account for various phenomena, and we now turn our discussion to models of L1 phonological interference that appear more promising in that they find more empirical support: feature-level models.

1.3. Feature Geometry and L1 phonological interference

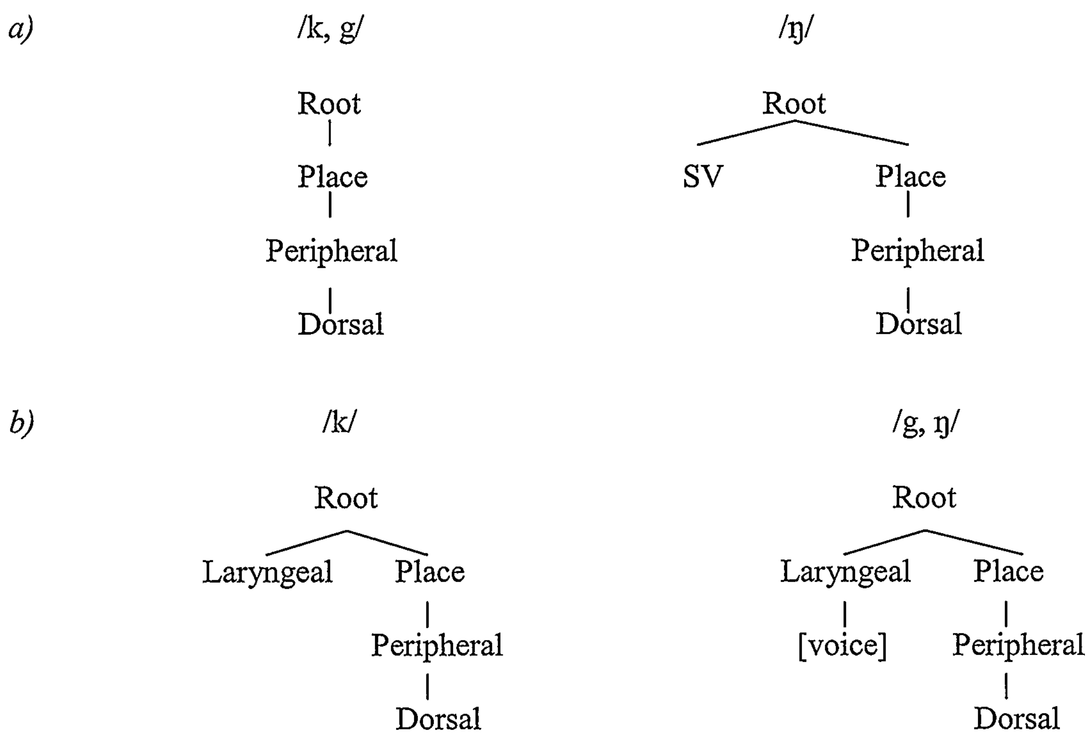
We will be assuming a model of Feature Geometry for the representations of phonemes, where the phonological features of a given segment are organized into a hierarchical tree structure. Rice and Avery (1995) show that such a model provides an elegant account for the patterns observed in the L1 acquisition of segmental phonology. The claim is that infants begin with very minimal feature geometry trees, and as acquisition proceeds, additional structure is added, thus allowing for representation of phonemic contrasts. This is illustrated in Figures 1.2 and 1.3 below (Rice and Avery 1995: 40). The appeal of such an approach lies in the fact that it captures both local variation and global uniformity in L1 segmental acquisition: children may vary in the order in which they add on to the feature geometry tree, but they all show evidence of being restricted in the ways that they are able to add to the tree, in that they do not violate any of the dependency relationships encoded in the tree structure. That is, a child with a representation such as that given in 1.2 below must be capable of contrasting coronal (Place), labial (Peripheral), and velar (Dorsal) segments, as all the features required to represent these contrasts are present in the feature geometry tree, and we would not expect to find a child who maintained a contrast between velar and labial segments, but not coronal segments (i.e., a child who had Peripheral and Dorsal, but not Place, in his feature geometry tree). Further, children cannot "invent" features to create new sounds. Thus, while a child with representations as in 1.3a below has added the Sonorant Voice (SV) node to contrast obstruents and sonorants (obstruents do not have an SV node), another child could just as easily add and elaborate the Laryngeal node to contrast voiced and voiceless segments, as in 1.3b.

Figure 1.2. L1 segmental acquisition as structure building – Stage I

/k, g, ŋ/

Root
|
Place
|
Peripheral
|
Dorsal

Figure 1.3. L1 segmental acquisition as structure building – Stage II



The feature trees that result from L1 acquisition, then, serve to define the phonemic categories as they are observed in categorical perception. The L1 feature geometry may, then, be the source of L1 interference in L2 segmental acquisition.

This possibility was the focus of work by Brown (2000), who ran a series of perceptual tasks in order to examine the abilities of speakers of three distinct L1s – Mandarin Chinese, Japanese, and Korean – with respect to a number of English sound contrasts that were absent for all three L1s, listed in Table 1.1 below along with the phonological feature required to represent the distinction.

Table 1.1. Sound contrasts examined by Brown (2000)

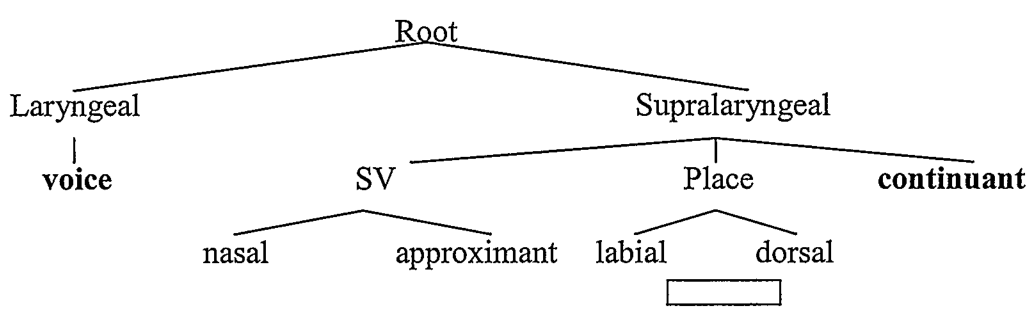
English Contrast	Feature Required
/p/ vs. /f/	[continuant]
/f/ vs. /v/	[voice]
/s/ vs. /θ/	[distributed]
/l/ vs. /ɹ/	[coronal]

Brown (2000) hypothesized that it was the L1 phonological features that determined the interference effects observed in L2 acquisition. That is, if a given speaker's L1 supplied the necessary features to represent a distinction between two L2 segments, then that contrast could be learned; if the required feature was absent from or inactive in the L1, however, then the learner would not perceive the novel contrast. Segmental approaches, on the other hand, do not make reference to the featural level, and thus make predictions on the basis of a segment's presence or absence in the L1 grammar. The feature-based hypothesis makes a number of testable predictions: first, it predicts that L1 Japanese, Mandarin, and Korean speakers will all be able to perceive the new English contrasts /b/ vs. /v/, /p/ vs. /f/, and /f/ vs. /v/, as all three L1s contain the features [voice] and [continuant]. Second, it predicts that none of the three groups should perceive the contrast /s/ vs. /θ/, as none of the L1s contain the feature [distributed]. Third, it predicts that neither the Japanese nor the Korean groups should perceive the contrast /l/ vs. /ɹ/, but the Mandarin group should be capable of perceiving this contrast, owing to the fact that while neither Japanese nor Korean contains the feature [coronal], Mandarin does as it maintains a plain /s/ vs. retroflex /ʂ/ contrast (Brown 2000). Further, this selection of L1s allows testing of Major and Kim's (1996) definition of similarity: in Korean, [l] is in complementary distribution with an apical alveolar flap [ɾ]. Korean speakers, then, have had phonetic experience with two liquids, one lateral and the other non-lateral, whereas Japanese and Mandarin speakers have only had experience with one. If it is true that (phonetic) similarity is the key in L2 segmental acquisition, then we would expect the Korean speakers to initially perform better than Japanese and Mandarin speakers on the English liquid contrast.

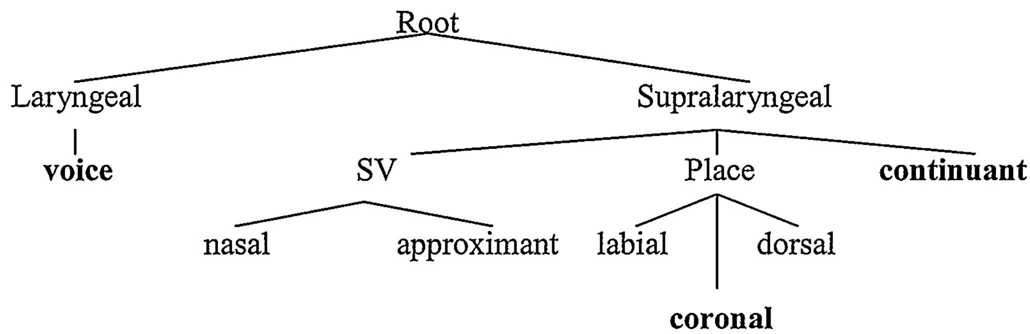
The feature geometry trees given in Figure 1.4 below represent the full adult feature geometries that learners bring to the task of L2 segmental acquisition (Brown 2000: 25). The features examined through the sound contrasts given in Table 1.1 above appear in boldface; the empty boxes in the Japanese (3a) and Korean (3c) geometries reflect the absence of the feature [coronal] in these languages.

Figure 1.4. Adult L1 feature geometries

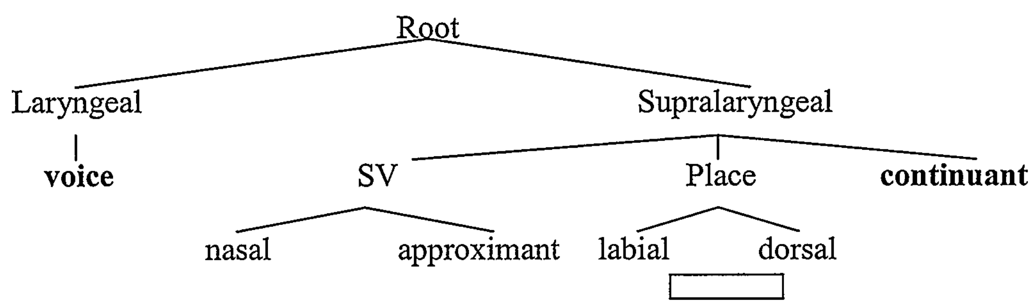
a) Japanese



b) Mandarin



c) Korean

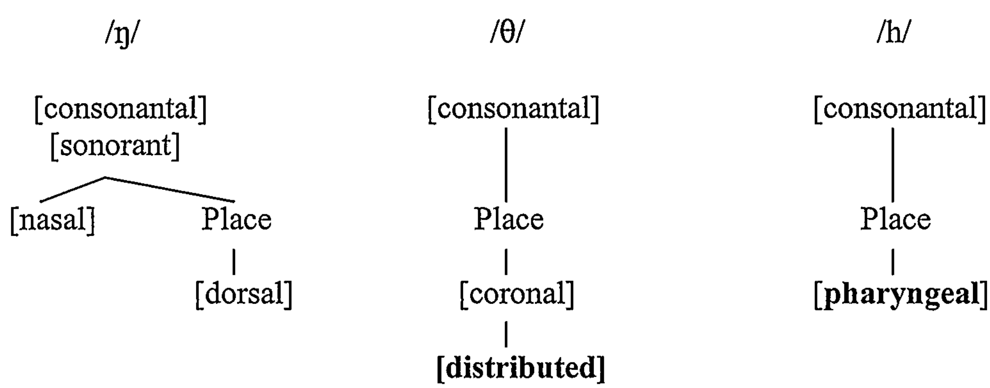


The predictions were tested using a 4IAX auditory discrimination task, in which subjects are presented with two pairs of words and are asked to indicate which of the two pairs contains two different words, as well as a two-way forced-choice picture selection task. The learners' performance was compared against that of a group of native speaker controls. The results showed that for all groups, performance on /p/ vs. /f/ and /f/ vs. /v/ was not significantly different from that of the native speaker controls, suggesting that the learners were able to construct appropriate phonological representations for these segments using the features supplied by the L1 grammar. Additionally, all three groups' performance was significantly worse than that of native speaker controls on the /s/ vs. /θ/ contrast, indicating that the absence of the required feature [distributed] from all three L1 grammars prevented appropriate representations from being constructed. With respect to the /l/ vs. /ɹ/ contrast, the groups are divided: the Japanese and Korean speakers' performance was significantly worse than both the Mandarin group and the native speaker controls. The Korean speakers' difficulty with this contrast is the precise opposite result of that predicted by Major and Kim's (1996) SDRH; recall that this approach to L1 phonological interference predicted that these individuals should be able to acquire the English liquid distinction as Korean has two liquid allophones. In sharp contrast to the outcome obtained for the Japanese and Korean speakers, there was no significant difference between the Mandarin speakers and the native controls: the Mandarin speakers were thus performing at a native-like level. This is attributed to the presence of the feature [coronal] in the Mandarin grammar and its absence in both Japanese and Korean. Thus, the L1 features are the source of interference in L2 segmental acquisition: those segments for which an appropriate representation can be built will be perceived and acquired, whereas those segments that require features which are lacking in the L1 grammar will not.

We can refer to this proposal as the strong L1 feature hypothesis, as it states that if the required feature is absent in the learner's L1 grammar, then that learner will *never* learn the new contrast as accurate perception of the sounds in question is blocked. In order to motivate such a strong claim, Brown (2000) presents data comparing a group of lower-level proficiency L1 Japanese L2 English speakers with a group of higher-level proficiency on two English contrasts not found in Japanese: /b/ vs. /v/ and /l/ vs. /ɹ/. The results show that while performance on /b/ vs. /v/ improves with increased proficiency to levels that are not significantly different from native speakers, performance on /l/ vs. /ɹ/ remains poor. That is, the absence of a required phonological feature completely rules out the possibility that a given contrast will be acquired. This claim seems to be further supported by work done by Matthews (1997) which examined the effects of articulatory training on the perceptual abilities of Japanese speakers on a number of English contrasts. Given the above discussion of Brown's (2000) work, of particular relevance are the pairs /b/ vs. /v/, /s/ vs. /θ/, and /l/ vs. /ɹ/. Subjects did not receive any perceptual training in that they did not listen to repeated native speaker productions of the targeted segments in class, but rather were given explicit articulatory instructions on producing the relevant speech segments in five sessions over a period of five weeks. They were also given the opportunity to practice pronunciation with real lexical items, and were given positive feedback when correct, and additional instruction and silent demonstration of articulation when incorrect. Testing was done before and after training using an auditory AX discrimination task, in which subjects are presented with pairs of words or syllables and are asked to indicate whether the two members of the pair are the same or different. Results showed a significant improvement on /b/ vs. /v/, but no improvement on /l/ vs. /ɹ/, which is consistent with the strong L1 feature hypothesis. Interestingly, however, performance on /s/ vs. /θ/ also improved with training, but did not reach significance. The fact that any improvement on this pair was seen at all hints that the strong L1 feature hypothesis may be too strong.

This possibility was further investigated by LaCharité and Prévost (1999). They looked at the acquisition of the English sounds /ŋ/, /θ/, and /h/ by native speakers of Quebec French (QF), which lacks all three sounds in its phonemic inventory. According to the strong L1 feature hypothesis, these speakers should experience perceptual difficulty with /θ/ and /h/, but not /ŋ/: all the features required to build an appropriate representation for /ŋ/ are available in the L1 grammar, whereas the features [distributed], required for /θ/, and [pharyngeal], required for /h/, are absent. Diagrams of the relevant portions of the representations of these segments is presented in Figure 1.5 below (LaCharité and Prévost 1999: 375); the features that are missing from the L1 grammar appear in boldface.

Figure 1.5. Representations of /ŋ/, /θ/, and /h/



Results from an AX discrimination task indicate that these three new sounds present different levels of difficulty: /ŋ/ is, not surprisingly, the easiest to perceive with the L2 learners performing at near-native levels, /θ/ is harder to perceive, and /h/ is the hardest.

With respect to the two latter segments, performance on /θ/ was lower than that of a native speaker control group but this difference was not statistically significant; performance on /h/, on the other hand, was significantly worse than the controls. These findings suggest a hierarchy of difficulty among new sounds: /h/ is hardest due to the absence of [pharyngeal], /θ/ is moderately hard due to the absence of [distributed], and /ŋ/ is easy as the L1 grammar supplies all the required features. A similar hierarchy was observed in Matthews' (1997) results: /l/ vs. /ɹ/ was the most difficult (and thus was not improved with training) due to the absence of [coronal], /s/ vs. /θ/ was moderately hard (and thus showed some improvement with training) due to the absence of [distributed], and /b/ vs. /v/ was the easiest (and thus showed significant levels of improvement) as the L1 grammar supplied all the necessary features. Taking into consideration both their own findings as well as those of Matthews (1997), LaCharité and Prévost (1999) propose that the relative difficulty of novel L2 segments is due to the kind of feature required by the L2 that is absent from the L1 grammar, with articulator nodes (such as [coronal] and [pharyngeal]) presenting a greater challenge to acquisition than terminal nodes (such as [distributed]). Furthermore, these findings suggest that, contra to the strong L1 feature hypothesis, it is possible to acquire new features in L2 segmental acquisition: improvement on segments involving the terminal feature [distributed] was seen in both studies, and LaCharité and Prévost's (1999) L1 QF subjects performed at a level that was not significantly worse than that of native speaker controls on /h/ in an ABX word identification task, in which subjects are presented with pairs of different words and then asked if a third word matches the first or second member of the pair; performance on /θ/ was still higher than /h/. We can therefore refer to this model of L1 interference as the weak L1 feature hypothesis.

While the weak L1 feature hypothesis is supported by the data described here, it also makes an interesting new prediction that takes us beyond the existing discussion: if the L1 feature geometry determines not only which L2 segments will be problematic to the learner but also what degree of difficulty the learner will experience with each, then we would expect this to be reflected in the learner's neural encoding and processing of the new sounds. In other words, the model implies that L2 learners are able to detect and make reference to features that are absent from their L1 grammar; we would expect this, in turn, to be reflected in the patterns of neural activity elicited in response to the new sounds. Thanks to developments in neuroimaging technology, as stated this is an empirically testable claim. The goal of the present research, then, is to test out the weak L1 feature hypothesis and explore issues related to the acquisition of new segmental structure in a second language. Namely, we seek to find evidence in support of the predictions made by the model, as well as evidence of change or restructuring of phonological categories that would suggest acquisition of previously absent phonological features.

1.4. Testing the model

In order to properly test out the predictions of the weak L1 feature hypothesis as set out above, it will be necessary to examine the behaviour of individuals with a common L1 who are at various stages in acquiring a given L2, as well as a group of individuals with the same L1 who have not been exposed to the selected L2. As all data collection was to be carried out in Calgary, Alberta, Canada, the most readily accessible population meeting these criteria was L1 English speakers, and the L2s that were selected for examination were French and Spanish. More specifically, the research was designed to examine L1 English speakers' acquisition of two new rhotic sounds: the French uvular trill /ʀ/ and the Spanish alveolar trill /r/.

The selection of French and Spanish trills as the segments of interest for the current investigation was made on the basis of several factors. First, there is much anecdotal evidence that L1 English speakers encounter difficulty in producing both trills, and it was of interest to see whether this difficulty extends to perceptual abilities as well. Second, the two trills are well-suited to testing LaCharité and Prévost's model with L1 English speakers, as they are predicted to present differing levels of difficulty in acquisition due to their phonological representations: Spanish alveolar trill /r/ differs from corresponding English alveolar approximant /ɹ/ along the dimension of manner of articulation, while French uvular trill /ʀ/ differs from English alveolar approximant /ɹ/ with respect to both place and manner of articulation[2].

Before moving into a discussion of the representations assumed for /ʀ/ and /r/, it is of use to briefly discuss which features are assumed to be present in English, as our assumptions differ from those of LaCharité and Prévost (1999) with respect to the feature

[2] Walsh Dickey (1997) notes that cross-linguistically, the alveolar trill /r/ is much more common than the uvular trill /ʀ/.

[pharyngeal] and its use in the representation of English /h/. Recall from discussion above that LaCharité and Prévost (1999) assumed that the articulator node [pharyngeal] was present in English to allow representation of /h/ and that this feature was absent in French, as they claim that there are no pharyngeal sounds in French. Notable, however, is the fact that /h/ does not involve any constriction of the pharyngeal cavity, as would be suggested by a [pharyngeal] feature. Therefore instead of following this convention, we will follow that of Halle et al. (2000) in assuming that /h/ has a [glottal] articulator node feature[3], rather than [pharyngeal], thereby allowing a representation for /h/ such as that found in 1.6 below. Further note that this difference in assumptions does not create any conflict with the L2 learner results found by LaCharité and Prévost (1999) with respect to the acquisition of /h/ by L1 QF speakers, as these individuals are still missing an articulator node required for an appropriate representation of /h/, given in 1.6 below.

Figure 1.6. Revised representation of /h/

/h/

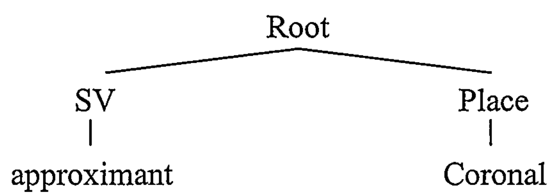
Root
|
Laryngeal
|
[glottal]

The representation of trills in feature geometric representations is not entirely straightforward, however, due to a wide range of proposed representations for trills in the

[3] Although Howe (2003) states that [glottal] is a Laryngeal feature (i.e., dependent of the Laryngeal node of the feature geometry), and not a Place feature, this detail is not crucial to our argument here as it is still considered an articulator node.

literature. English /ɹ/ is less controversial: both Brown (1997) and Blevins (1994) agree that /ɹ/ should have a representation like that given in Figure 1.7 below (Brown 1997: 67).

Figure 1.7. Representation of English /ɹ/



Walsh Dickey (1997) provides the representation in Figure 1.8 for the alveolar trill /r/ and Figure 1.9 for the uvular trill /ʀ/; however, she also notes that no language contrasts trills and approximants at the same place of articulation, and concludes that trills and approximants are therefore phonologically identical. That is, both Spanish alveolar trill /r/ and English alveolar approximant /ɹ/ would have a representation such as that given in 1.8 below.

Figure 1.8. Representation of /r/ (Walsh Dickey 1997: 173)

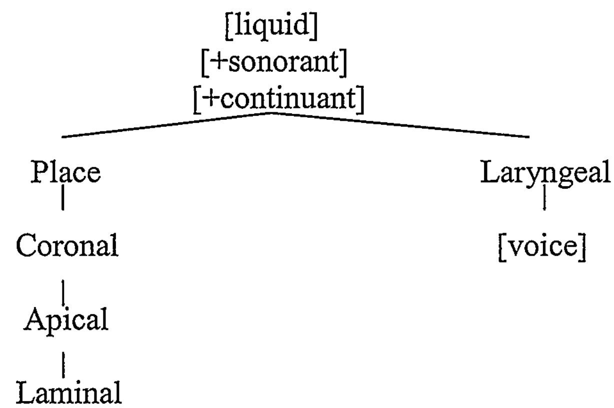
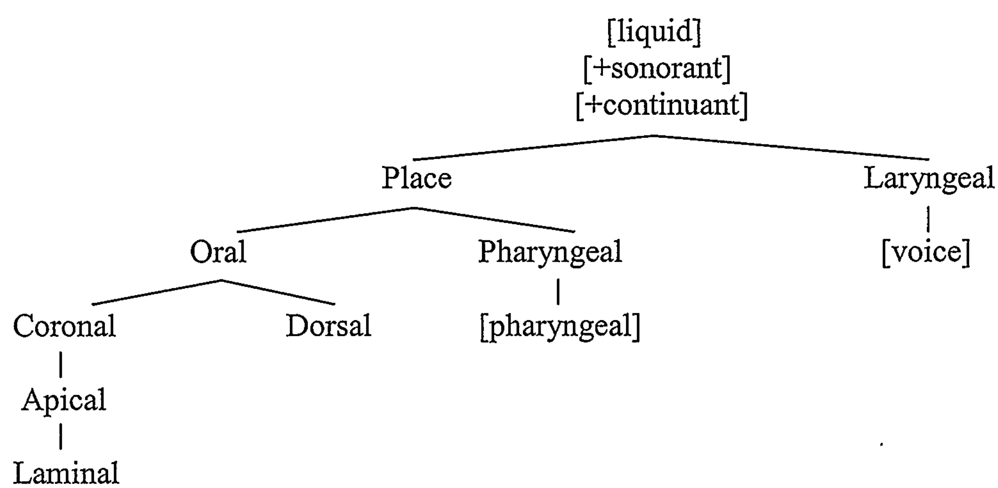


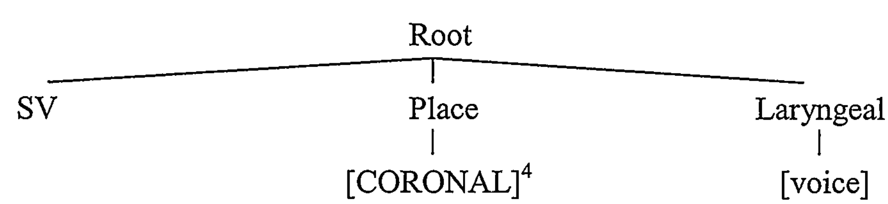
Figure 1.9. Representation of /ʀ/ (Walsh Dickey 1997: 171)



Bakovic (1994) presents arguments that trills and approximants should be differentiated in the phonology, as trills are articulatorily more complex than approximants: the configuration of the tongue must be such that will allow a Bernoulli effect to bring about the repeated closures of the trill. His account, then, is captured in phonological representations in terms of Aperture Theory (Steriade 1993). Under this view, /ɹ/ would be specified as A_{max} (oral release) while /r/ would be specified as having both A_0 (oral closure) and A_{max} aperture settings (Bakovic 1994: 33). This view is also put forth by Bradley (2001). Colantoni (2001) provides a way to capture this distinction in the feature geometry by proposing the terminal feature [vibrant] for trills, claiming that this feature characterizes the vertical movement of the tongue associated with trilled sounds.

At this point, phonological theory does not provide a directly obvious way to select which of these representations should be assumed here: all find support in data from various languages. As a result, we will be making the following assumptions about the representation of /ɹ/, /r/, and /ʀ/, as diagrammed in 1.10, 1.11, and 1.12 below, respectively. These structures follow those proposed by Walsh Dickey (1997), with the addition of the terminal feature [vibrant] for the trills following Colantoni (2001); the dependents of [CORONAL] (Apical and Laminal) as proposed by Walsh Dickey (1997) are not crucial to the discussion here and thus have been pruned for reasons of space. We are assuming these representations as they allow us to capture the distinctions among rhotic segments that we have described in the languages we have chosen to examine.

Figure 1.10. Assumed representation of /ɹ/



[4] It should be noted that the use of capital letters for articulator nodes in the representations given here and throughout the remainder of the thesis does not carry any theoretical implications. It is a notational convention that is only used to highlight the distinction being maintained between articulator and terminal node features.

Figure 1.11. Assumed representation of /r/

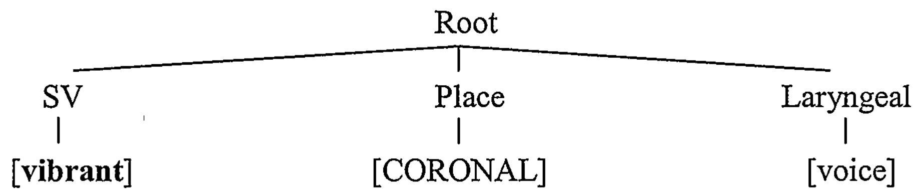
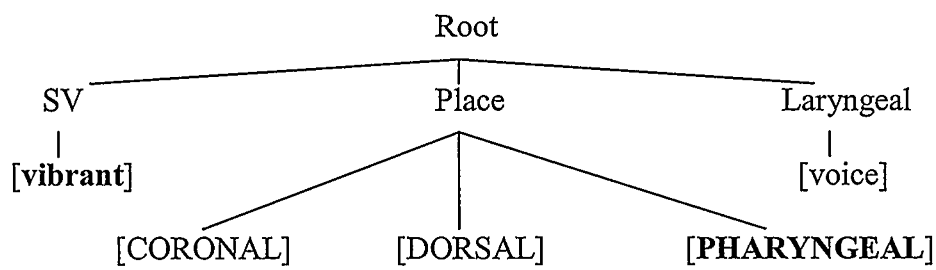


Figure 1.12. Assumed representation of /ʀ/



It is worth commenting on the appearance of [CORONAL] in the representation of the uvular trill /ʀ/, as this proposal cannot be said to be motivated by articulation. We have included it in our representations here so as to adhere to Walsh Dickey's (1997) proposals as closely as possible; however, Walsh Dickey (1997) suggests that the presence of [CORONAL] and its dependents for all rhotics serves to unify the class of rhotic sounds, and further notes that there is no existing phonological data to either support or disprove this claim.

According to the L1 interference model being tested here, then, Spanish /r/ should prove difficult for L1 English speakers to acquire due to the fact that the L1 feature geometry does not supply the terminal node [vibrant]; French /ʀ/, however, should present greater difficulty to these learners as in this case, the L1 feature geometry crucially fails to supply the articulator node [PHARYNGEAL].

A previously unexplored implication of the weak version of our model of L1 interference is that the hypothesis should have neurological support: that is, if learners are able to (sub-consciously) make judgments about new sounds on the basis of phonological features that are absent from their L1, then this should be reflected in the neural encoding and processing of new sounds. The hypothesis we are testing is that those sounds which are deemed to be easier to acquire should elicit different neural responses than those sounds which are harder to acquire. More specifically, /r/ should elicit different neural response patterns than /ʀ/ in native English speakers. This is the claim through which we seek to test the weak feature hypothesis, using data collected through event-related potentials, or ERPs. We now turn to a discussion of these.

1.5. ERPs

The term event-related potential (ERP) refers to measures obtained through continuous sampling of the electrical component of the body's electromagnetic fields, or the electroencephalogram (EEG), and these are used in neuroimaging. Neural activity can be measured as voltage deflections in the EEG: a neuron at rest will have a negative charge, and a stimulus generates a nerve signal by altering the plasma membrane of a neuron so that it becomes positively charged (Campbell et al. 1997). The brain is never completely at rest, as it monitors and controls every bodily function, from coordinated muscle movements to less voluntary bodily functions, such as heart rate and breathing. In order to properly identify deflections in the EEG and reliably associate these with a specific stimulus, we must compare these EEG signals with those generated when no specific stimulus is present; that is, any EEG deflections must be considered relative to an averaged baseline. This will allow us to extract meaningful deflections from "noise" that occurs as a result of normal, unavoidable brain activity.

ERPs, then, are deflections in neural activity relative to an averaged baseline that are time-locked to a given stimulus; they thus reflect the neural activity related to the processing of that particular stimulus (Rugg 1999, Rugg & Coles 1995). Their use in the present research is due to a number of desirable characteristics of ERP measurement: first, they are sensitive to very brief periods of neuronal activity, on the order of tens of milliseconds (Rugg & Coles 1995); second, they offer great temporal precision: van Turennout et al. (1997, 1998) have used ERPs to establish the relative time course of semantic and phonological processing in lexical access. These properties suggest that ERPs may provide a useful measure in examining the processing of segmental structure, due to the temporally short nature of individual speech segments. ERPs do not, however, offer a good deal of spatial resolution (Rugg 1999), therefore any conclusions drawn from the work presented here will not bear on regions of activation.

A number of ERPs have been identified for use in research, and each reflects a different aspect of the brain's response to a stimulus. Some responses are automatically elicited by sensory stimuli, in that they do not require that the subject attend to the stimuli being presented, and are thus said to be obligatory. Other responses are related to attention, comprehension, and decision-making processes. The discussion that follows here will be limited to two obligatory evoked potentials, the mismatch negativity, or

MMN, and the N100, as these are the ERPs which have been used in existing investigations of phonological processing and speech sound perception.

1.5.1. MMN

The mismatch negativity (hereafter MMN) is also known as the N2 or the N200. It is a large negative (hence, N) evoked potential that occurs approximately 200 milliseconds (msec) after presentation of a visual or auditory stimulus (Phillips 2001, Poeppel and Marantz 2000, Dehaene-Lambertz et al. 2000, Dehaene-Lambertz 1997). It is elicited by a deviant stimulus presented in a stream of standard stimuli, and thus reflects the automatic detection of physical deviance (Coles & Rugg 1995). Furthermore, the MMN has been argued to be sensitive to the degree of deviance: the more physically disparate the deviant is from the prevailing context, the greater the amplitude of the MMN (Coles & Rugg 1995).

A number of studies have relied on the MMN as evidence of subjects' ability to perceive various speech sound contrasts. Dehaene-Lambertz et al. (2000) examined the perceptual abilities of Japanese speakers with respect to consonant clusters. Japanese obeys a number of constraints on syllable structure, and as a result does not permit obstruent consonant clusters (Shirai 2001). When words containing these clusters are borrowed into the language, they are repaired with an epenthetic vowel, typically /u/, so as to ensure that all constraints on syllable structure are satisfied: the consonant cluster is broken up by inserting a vowel between the consonants. This is exemplified in loanword adaptations from English: /ski/ 'ski' is adapted into Japanese as /sukii/ (Mah 2001). Dehaene-Lambertz et al.'s (2000) research was carried out using minimal pair nonsense words in what is referred to as a mismatch paradigm: participants were presented with a repetitive stream consisting of some nonsense word, labeled the standard stimulus, that was followed by a single token that was either another repetition of the standard stimulus, or a different nonsense word, labeled the deviant stimulus. When the Japanese subjects heard deviants that consisted of a nonsense word whose internal cluster had been repaired with an epenthetic vowel /u/ (such as /ebuzo/) in the context of a minimally paired nonsense word with the cluster intact (such as /ebzo/), no MMN was elicited. The same stimuli presented to a control group of French speakers, however, yielded a large and robust MMN response. Dehaene-Lambertz et al. (2000) take these findings to suggest that the French speakers were able to detect the presence or absence of the epenthetic vowel, owing to the fact that French allows these consonant clusters. The Japanese speakers, on the other hand, could not detect this difference, which is argued to suggest that for these individuals, the consonant sequences were automatically repaired.

In another study, Dehaene-Lambertz (1997) studied the ability of native French speakers to perceive a native place of articulation contrast (/ba/ vs. /da/) as compared with these same subjects' ability to detect a non-native dental vs. retroflex contrast found in Hindi (/da/ vs. /ɖa/). The results showed that while the native contrast elicited a large MMN, the non-native contrast did not. That is, the non-native contrast was not readily detectable to subjects, whereas the native contrast was easily perceived. Since the stimuli were drawn from a synthesized continuum, Dehaene-Lambertz (1997) was able to control for acoustic distance between standards and deviants used, and thus hypothesized that the different detection of contrasts was due to the storage of a phonemic representation of the standard, against which the deviants were compared. Where the deviant belonged to a separate phonemic category, an MMN was elicited. Where the deviant did not, no MMN was found.

Similar results were found in an MEG study examining the MMN's magnetic counterpart, the magnetic mismatch field (MMF) (Phillips et al. 2000). This particular study sought to demonstrate that the MMF is sensitive to phonological categories and reflects categorical perception of acoustic tokens into phonological categories. An adapted oddball paradigm was used, in which several tokens are used as a standard stimulus, and several tokens are used as a deviant stimulus. Strictly acoustically speaking, all stimuli used in an adapted oddball paradigm are different; however, more abstract categories of mental representation allow these physically distinct items to be grouped together into categories. Stimuli drawn from a synthesized VOT continuum were manipulated so that in one experimental condition, a many-to-one ratio of phonological categories was present, i.e., the standard stimuli were syllables that were identified as [ta] while the deviant stimuli were syllables that were identified as [da]. In the other experimental condition, there was no many-to-one ratio of phonological categories: there were equal numbers of [ta] and [da] syllables. The acoustic distribution of stimuli remained identical across both conditions, in such a way that there was never a many-to-one ratio of acoustic tokens. The results showed that the many-to-one ratio of phonological categories condition elicited an MMF, while the other condition did not.

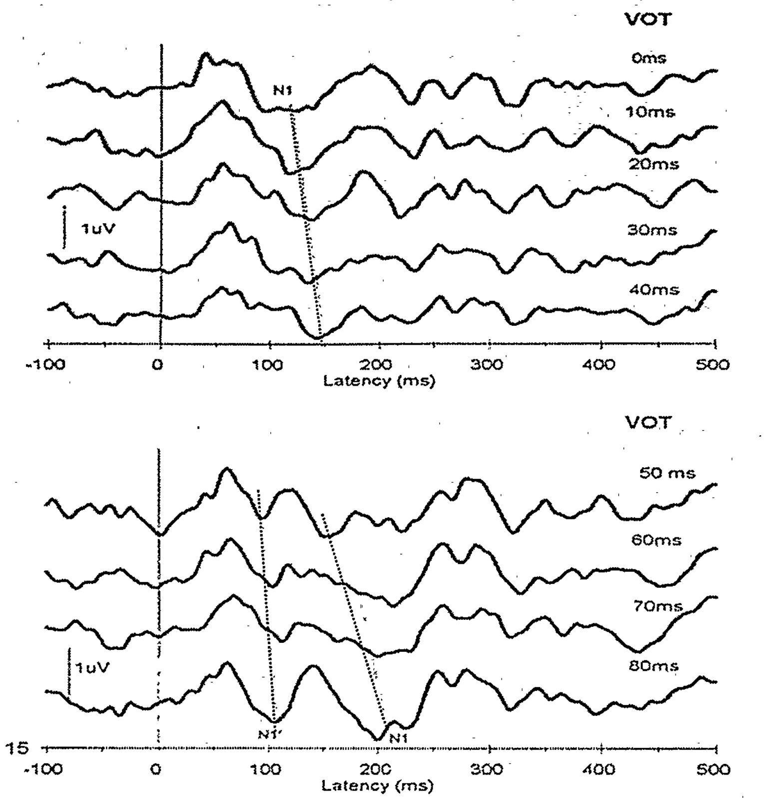
The finding that the MMF/MMN is sensitive to categorical perception does not seem to be consistent. Sharma et al. (1993) found that an MMN was elicited in both within and across phonological category conditions; that is, regardless of whether the standard and deviant stimuli were both tokens of one phoneme, or tokens of separate phonemes. Work by Maiste et al. (1995) suggests that the MMN is dependent on the acoustic content of standard and deviant stimuli being examined. No MMN was elicited when the deviant stimulus frequency content was a subset of that of the standard stimulus: for example, a deviant stimulus consisting of a pure tone will not elicit a MMN when presented in the context of a standard stimulus consisting of a complex tone where the frequency of the deviant pure tone is also found in the complex tone. This suggests that presentation of new frequencies is required to elicit an MMN. A study by Sharma and Dorman (1999) revealed that while an MMN was elicited in both within and across category conditions, the MMN was larger and more robust in the across category condition.

1.5.2. N100, N1 – P2 complex

The N100, or N1, is a major negative deflection following presentation of a stimulus (Sams et al. 1990, Maiste et al. 1995, Sharma & Dorman 1999, Tremblay et al. 2001). It typically occurs between 80 and 140 msec (Maiste et al. 1995) and has been argued to reflect sensory encoding of stimulus features (Sharma & Dorman 1999), as it is sensitive to changes in frequency and amplitude (Poeppel & Marantz 2000, Phillips 2001). In examining behaviours related to speech sounds it responds more readily to changes in formant structure, or the higher frequencies (i.e., F2, F3), than it does to changes in the fundamental frequency (F0) itself; that is, it seems to reflect segment identification rather than speaker identification (Poeppel & Marantz 2000, see also Phillips 2001). It also responds to temporal cues, such as VOT (Sharma & Dorman 1999, Tremblay et al. 2001), and has also been argued to serve as an indicator of categorical perception (Sharma and Dorman 1999). Furthermore, it is most sensitive to stimulus onsets (Maiste et al. 1995).

Sharma & Dorman (1999) examined the N100 as a potential neural correlate of categorical perception of VOT by native English speakers using synthesized syllables along a /da/ - /ta/ continuum. Their findings suggest that the encoding of VOT is reflected in morphological changes to the waveform: stimuli with short VOTs that were consistently identified behaviourally as /da/ (0 – 30 msec) elicited a single N100 component, or a single negativity, while stimuli with longer VOTs that were consistently identified behaviourally as /ta/ (50 – 80 msec) elicited an N100 with two distinct components, as illustrated in Figure 1.13 below (Sharma & Dorman 1999: 1082). Furthermore, in the long VOT condition, the second N100 peak (labelled N1) demonstrated a high correlation with VOT while the first peak (labelled N1') did not, which is taken to suggest that the first component occurred in response to a noise burst at syllable onset, while the second component occurred in response to the onset of voicing.

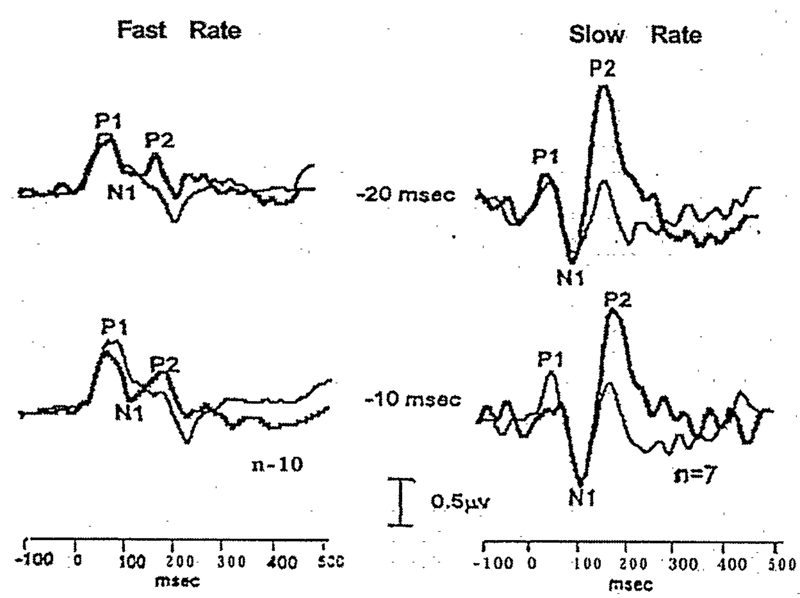
Figure 1.13. N100 responses along VOT continuum



Tremblay et al. (2001) also examined the behaviour of the N100 with respect to VOT, but not in the context of categorical perception: they sought to determine if the N100, along with the following positive deflection P200 (together the two responses are referred to as the N1 – P2 complex), reflects training-induced improvements in the perception of VOT. That is, they were investigating whether the N1 – P2 complex reflected neural plasticity, as evidenced by changes according to experience (here, training on a speech sound contrast). Native speakers of English were tested, trained, then re-tested on discrimination of two synthetic syllables representing a non-native contrast, /mba/ and /ba/. Two different rates of presentation were also used: a slow rate, with an interstimulus interval (ISI) of 910 msec, and a fast rate, with an ISI of 390 msec. This manipulation was motivated by a desire to determine whether the N1 – P2 complex could be elicited at faster rates typical of other research paradigms.

The results showed that improved perception of the VOT distinction on a behavioural auditory discrimination task was accompanied by a significant increase in amplitude of the N1-P2 complex; however, the waveform pattern obtained for /mba/ was not significantly different from that obtained for /ba/.

Figure 1.14. N1 – P2 complex before (thin line) and after (thick line) training



Tremblay et al. (2001) argue that these changes in the N1 – P2 complex reflect increases in neural synchrony, or an increase in the amount of neural activity occurring in response to a given speech segment, thereby reflecting improved speech perception on the novel contrast. The findings do not, however, suggest that this training has encouraged the formation of two phonological categories, as no changes were noted in the waveform shape, only its amplitude. That is, training brought about improved within-category perception, but not the formation of new categories with new representations, and the N1 – P2 complex allows us to observe this distinction.

These findings suggest that the N100 and the N1 – P2 complex are ideally suited to the investigation set out here: the N100 will provide information on whether second language learners are able to build novel phonological representations for new sounds, or are categorizing new sounds with "similar" L1 sounds; the N1 – P2 complex will allow us to observe any improvements on their within-category perception abilities.

1.6. Research goals

The research described in the remainder of this thesis was carried out in order to provide answers to the following questions:

1) Do monolingual English speakers differentiate between English /ɹ/ and other rhotics?
2) Do monolingual English speakers make distinctions based on absent phonological features?
3) Can language learners trigger new structure (here, features) and build appropriate phonological representations for new sounds?

As work progressed, it became clear that the data allowed us to address an additional question:

4) How does experience with a new sound in a second language affect its acquisition in a third language?

Questions (1) and (2) are addressed in Chapter 2; question (3) is addressed in Chapter 3; question (4) is addressed in Chapter 4. Chapter 5 presents a summary of findings and further discussion.

Chapter 2

On the initial state of adult L2 phonological acquisition

2.1. To test a model of L1 interference

The discussion of the previous chapter addressed two major issues in the field of second language acquisition. First, it has long been noted that not everything that is new in the L2 seems to present the same degree of difficulty to L2 learners. With respect to phonological acquisition, we find that not all new segments are equally hard to acquire. Our goal, then, is to determine whether theoretical constructs developed in linguistic theory are able to account for this. Success in this endeavour would not only provide a satisfying answer to our questions about L2 acquisition, it would also provide evidence that our assumed theoretical constructs have some psychological reality and explanatory adequacy. An approach to L1 interference that based itself on current theoretical approaches in phonological theory was developed in work by Brown (2000) and LaCharité and Prévost (1999).

Under examination here are the predictions of both the weak L1 feature hypothesis (LaCharité and Prévost 1999) and the strong L1 feature hypothesis (Brown 2000). The weak L1 feature hypothesis predicts that for novel L2 segments, those whose phonological representation requires only those features which are present in the L1 will be easy to acquire; those segments which require a terminal node feature that is absent in the L1 will present greater difficulty but will still be acquirable, and those segments which require an articulator node feature that is absent in the L1 will the most difficult, if not impossible to acquire. The strong L1 feature hypothesis, on the other hand, predicts that novel L2 segments are either acquirable or not, depending on whether the required features are available in the L1 grammar.

While LaCharité and Prévost's (1999) discussion does not address this, we observed that crucial to their model of interference is the assumption that L2 learners are able to detect novel phonological features in the input and make reference to these in classifying novel segments according to their degree of difficulty. If relative difficulty in acquisition is due to phonological features and the weak L1 feature hypothesis is to be supported, then any judgments made on the basis of these should appear in the earliest stages of acquisition, from the moment the learner first hears the new sound. This is not to say that we are assuming that learning is instantaneous; what we do assume, however, is that the distinguishing characteristics of speech sounds on which phonological features are based (i.e., place of articulation) have some acoustic manifestation that universally allows them to be assigned to an appropriate phonetic category (Borden et al. 1994).

The present chapter reports on an electrophysiological test of the weak L1 feature hypothesis: section 2.2 reviews the model, its predictions, and the testing method used here; section 2.3 provides the details of the data collection procedure; section 2.4 presents the results obtained; section 2.5 provides some discussion of these results.

2.2. Predictions of the strong and weak L1 feature hypotheses

Unlike the strong version, the weak L1 feature hypothesis suggests that L2 learners are able to detect features that are absent from the L1 grammar and make distinctions among segments on the basis of these. That is, L2 learners are able to recognize acoustic manifestations of novel phonological features. We would expect that these abilities are present at the initial state of acquisition, where learners first hear the new segments, as observed improvements on various segments are taken as evidence that the learners are acquiring the required structure. These assumptions, in turn, allow us to predict that those segments which, according to the L1 grammar, are relatively more difficult to acquire (due to a missing articulator node), will be subject to different neural processing and encoding than those segments which are relatively less difficult for acquisition (due to a missing terminal node). That is, under the assumption of this particular model of L1 interference, we would expect to see different neural responses elicited by the two types of segments. Under the assumption of the strong version, however, we would not expect to see any differences in neural responses, as both segments would be predicted to be difficult for acquisition.

These predictions are tested here by examining the neural responses elicited in response to trilled rhotics in French and Spanish in monolingual English speakers. The feature geometry structures for the two trills, given in Figures 2.1 and 2.2 below, serve to illustrate the appropriateness of these two segments for the task at hand: for both trills, English lacks the terminal feature [vibrant], however, for the French trill, English crucially lacks the articulator node [PHARYNGEAL]. For expository purposes, those features absent from English in the representations below appear in boldface. The feature geometry structure for the English approximant /ɹ/ appears in Figure 2.3.

Figure 2.1. Feature geometry of French trill /ʀ/

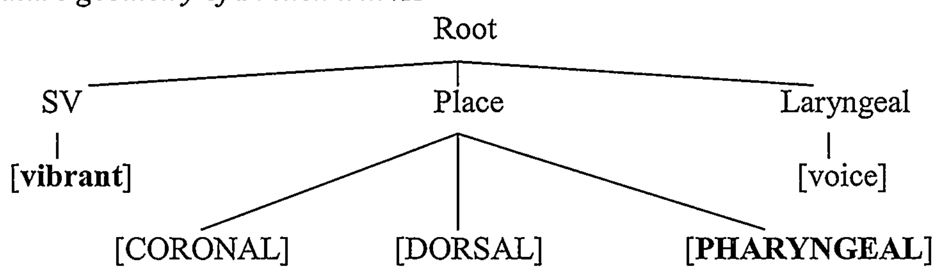


Figure 2.2. Feature geometry of Spanish trill /r/

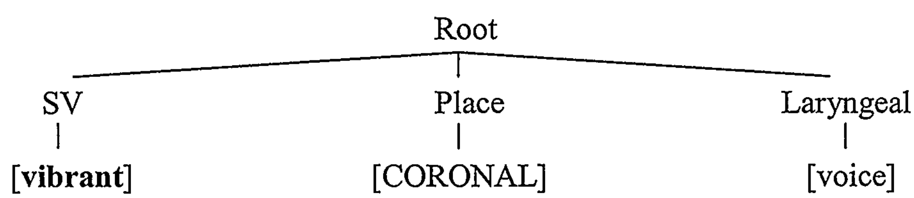
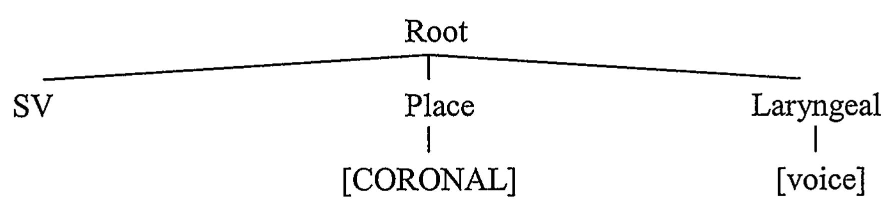


Figure 2.3. Feature geometry of English approximant /ɹ/



According to the weak L1 feature hypothesis, then, the Spanish trill /r/ should be difficult for L1 English speakers as they lack the terminal feature [vibrant]; however, the French trill /ʀ/ should be significantly more difficult for these same individuals due to the lack of the articulator node [PHARYNGEAL]. Following our prediction discussed above, if the weak version is to be supported, we would expect to find significantly different patterns of neural activity occurring in response to these two segments in monolingual English speakers who are not learning either French or Spanish. If the strong version is to be supported, then the two segments should not differ in their patterns of neural activity in these individuals.

This prediction is tested by collecting event-related potential (ERP) data from monolingual English speakers as they listen to the three rhotic sounds discussed above. ERPs are well-suited to investigations of phonological processing as they record data in a near-continuous fashion, allowing the observed responses to be reliably associated with the presentation of particular stimuli; responses are said to be time-locked to the stimuli (Rugg 1995). The patterns of neural activity that occur in response to each rhotic are evaluated here with respect to two evoked potentials: the N100 (also called the N1), and the N1 – P2 complex. The N100 has been described as a major negative deflection in the waveform with a peak latency between 80 and 140 milliseconds post-stimulus (Maiste et al. 1995) whose peak latency reflects categorical perception of speech sounds (Sharma and Dorman 1999): speech segments identified as belonging to the same category show N100 deflections with similar peak latencies, whereas speech segments identified as belonging to different categories show N100 deflections with different peak latencies. The N1 – P2 complex is the portion of the waveform spanning from the N100 peak to the peak of the following positive deflection, referred to as P2, which has a peak latency of around 175 milliseconds post-stimulus (Tremblay et al. 2001). The amplitude of the N1 – P2 complex has been shown to be related to synchronous neural activity, such that increased amplitude indicates increased neural activity associated with the stimulus (Tremblay et al. 2001) (see section 1.5 above for further discussion).

More specifically, then, under the assumption of the weak L1 feature hypothesis, our predictions can be stated as follows:

1) All three rhotics should elicit N100 evoked potentials with similar peak latencies: monolingual English speakers are unable to build appropriately distinct representations for /ʀ/ and /r/ and thus assign these to the existing structure for non-lateral approximants: /ɹ/.

2) The three rhotics should differ with respect to the amplitude of the N1 – P2 complex each elicits: the amount of neural activity associated with each sound should reflect that /ʀ/ and /r/ differ with respect to the type of feature that is missing from the L1 grammar.

Under the assumption of the strong L1 feature hypothesis, the prediction in (1) above still stands, the prediction in (2), however, is replaced by (3):

3) The two novel trills should differ from the native English approximant with respect to the amplitude of the N1 – P2 complex each elicits: the amount of neural activity associated with these sounds should reflect that both trills require a feature that is missing from the L1 grammar.

2.3. Data collection

2.3.1. Stimuli

The stimuli used in this experiment were the syllables /ɹi/, /ʀi/, and /ri/, along with the control condition syllables /ɹo/ and /bo/. The syllables were elicited in isolation from a male native speaker of English who is also a near-native speaker of both French and

Spanish. Each syllable was repeated three times, and the second repetition was used in preparing the task items. The syllables were recorded using a Sony TCD-D100 DAT recorder and Sony ECM-MS908C electret condenser microphone. These were then edited using Peak LE 2.62 digital audio editing software on a Macintosh desktop computer to eliminate empty portions of the sound files. The syllables were then transferred to a desktop PC computer, which was used to present the stimuli during the experiment using Neuroscan Inc.'s STIM stimulus presentation software package.

2.3.2. Subjects

Eight subjects participated in this experiment: all were undergraduate students in an introductory Linguistics course at the University of Calgary, with a mean age of 22.5. All were native speakers of English, none had previously studied Spanish; Subject E5, however, had previously studied Italian and had spent one academic year on an exchange program in Italy. Some variation was found with respect to previous exposure to French: although none claimed to understand any French, subjects E1, E2, E3, and E8 all reported having taken a French class in either elementary (E1, E2, E3) or junior high school (E8), but all report that class time was limited to one or two classes per week, and that the classes lasted one academic year (except for E8, whose French class lasted one semester). Subject E7 reported that he had completed three years of elementary school in a French immersion program, but was pulled out of the program after Grade 4 due to poor performance. Additionally, Subject E4 is a native speaker of both English and Cantonese Chinese. All subjects gave informed consent for their participation in this research, and were paid $20 for their participation.

The data presented and analyzed here will not include data obtained from Subjects E5 and E7; these will be discussed in Chapter 4. The discussion here, then, will focus on those individuals with minimal or no exposure to languages other than their first language. Although Subject E4 is a bilingual English/Cantonese speaker, her data are included in this group as exposure to Cantonese is not expected to have any effect on her perception of the French and Spanish trills as Cantonese has no trills (Killingley 1993). Due to technical difficulties encountered in recording, not all data that was collected was available for subsequent off-line analysis. In an attempt to remedy this problem, data from Subjects E6 and E7 was recorded at an Analog-to-Digital (A/D) rate of 1 kHz, which was a deviation from the procedure followed for all other subjects. This had the unfortunate result of making the resulting data files incompatible with those obtained from other participants, and Subject E6's data could not be included in preparing the figures presented below. It was possible, however, to include his data in the statistical analyses.

2.3.3. Procedure

The experiment was conducted at the Language Research Centre at the University of Calgary. Surface electrodes were placed using a Neuroscan Quik Cap electrode cap, which arranges the electrode array according to the International 10-20 system (Jasper 1958, cited in Tremblay et al. 2001), along with a forehead ground and right mastoid reference (A2). Data recording was done using a Neuroscan NuAmps digital EEG amplifier and Neuroscan's Scan 4.2 Acquisition software package, run on a Dell Inspiron laptop computer. Stimuli were presented using Neuroscan's Stim stimulus presentation software package on a desktop PC. Analysis of raw data was performed off-line using Neuroscan's Scan 4.3[5] Edit software package, run on a Dell Inspiron laptop computer.

Subjects watched a video while being fitted with the electrode cap. Once all the equipment had been properly set up and configured, the experiment began. A mismatch paradigm was used, in which a stream of identical syllables is occasionally interrupted by a different syllable, which is referred to as the deviant stimulus. The stimulus syllables were paired as described in Table 2.1 below; for all conditions, 85% of the tokens consisted of the syllable labelled as the standard, and 15% of the tokens consisted of the syllable labelled as the deviant. Syllables were presented with an onset-to-onset interstimulus interval of 850 milliseconds. For the control conditions /ɹo/ vs. /bo/ and /bo/ vs. /ɹo/, two blocks of 250 syllables were presented, for a total of 500 tokens. For all other conditions, eight blocks of 250 syllables were presented, for a total of 2000 tokens. Stimuli were presented to the subject's right ear via earphone at a volume of 75 dB, and subjects were asked to indicate the occurrence of the deviant stimulus by pressing a button on a response pad held in their left hand using their left thumb.

[5] While recording was done using version 4.2 of Neuroscan's Scan software, an upgrade became available before analysis began, and it was thought that upgrading the system might allow us with future subject groups to avoid some problems in data acquisition that had been encountered.

Table 2.1. Stimulus pairings

Standard	Deviant
/ɹo/	/bo/
/bo/	/ɹo/
/ɹi/	/ʀi/
/ʀi/	/ɹi/
/ɹi/	/ri/
/ri/	/ɹi/

Due to technical difficulties in recording[6], not all data collected was available for off-line analysis. Table 2.2 details the average number of individual stimulus presentations (data sweeps) analyzed per participant in each condition, as well as the total number of sweeps analyzed for all participants in each condition.

Table 2.2. Average number of sweeps analyzed

	Average Number of Sweeps per Participant	Total Number of Sweeps Analyzed
/ɹo/	219	1314
/bo/	217	1299
/ɹi/	808	4847
/ʀi/	840	5037
/ri/	821	4926

A number of methodological standards were followed in recording the data: electroencephalogram (EEG) channels were analog bandpass filtered on-line from 0.5 Hz to 70 Hz, amplified with a gain X 1338 and converted using an Analog-to-Digital (A/D)

[6] One of the cables connecting the various components of the system was found to be faulty, the result being that not all stimulus event triggers were being recorded in the continuous data files.

rate of 500 Hz[7]. Eyeblink artifact was removed off-line, and the continuous data files were segmented into recorded sweeps consisting of 100 millisecond pre-stimulus, 500 millisecond post-stimulus epochs. Epochs containing voltage deflections greater than 100 µV were deemed to be contaminated by artifact and were rejected off-line. Remaining sweeps were baseline corrected and filtered off-line using a 30 Hz low pass filter (12 dB/octave). All data and resulting waveforms were recorded from the electrode Cz, following Sharma and Dorman (1999) and Tremblay et al. (2001).

2.4. Results

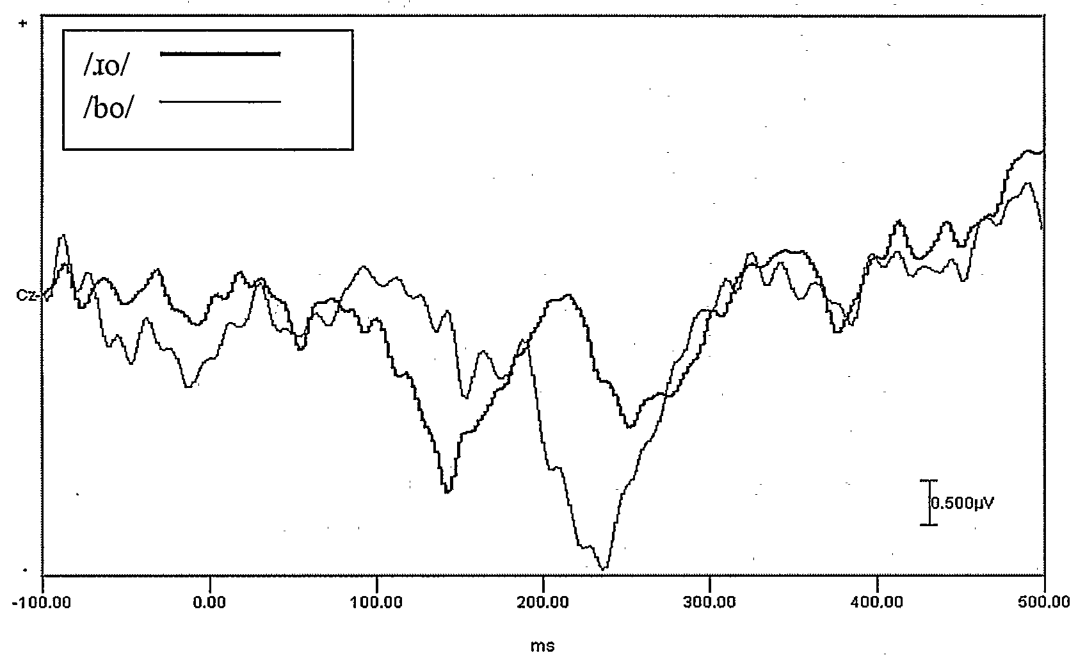
Figure 2.4 presents the group-averaged waveform response data for the control condition of /ɹo/ (thick line) vs. /bo/ (thin line); in (a) the two waveform responses are overlaid with a common baseline, in (b) they are rastered out for expository purposes, with the dashed horizontal lines indicating the baseline for each waveform. While we present the results using both graphing formats here and in Figure 2.5 below, for reasons of space and expository clarity only the rastered versions will be presented throughout the remainder of the thesis. It should further be noted that negative and positive peaks are identified in terms of relative negative and positive voltage shifts, and not in terms of absolute voltage values. That is, in 2.4 below, while both P2 peaks appear below the baseline and thus have negative absolute voltage values, they are clearly positive shifts in voltage following the negative N100 peak.

---

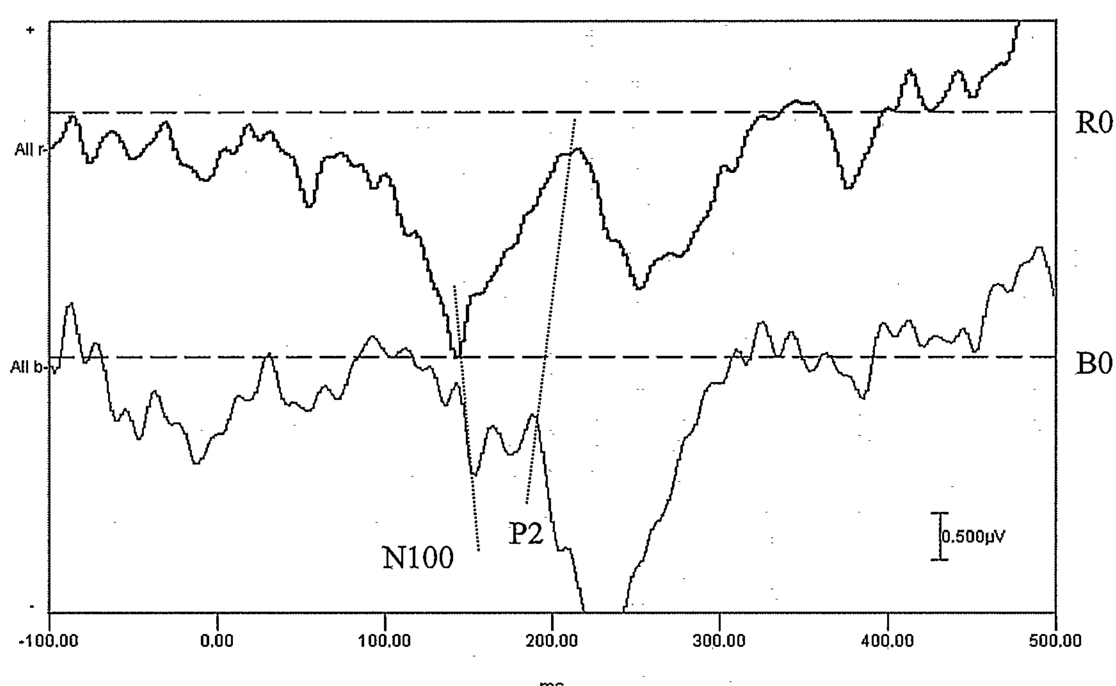
[7] As mentioned above, Subjects E6 and E7 were recorded using an A/D rate of 1 kHz.

Figure 2.4. Control condition responses

a)



b)

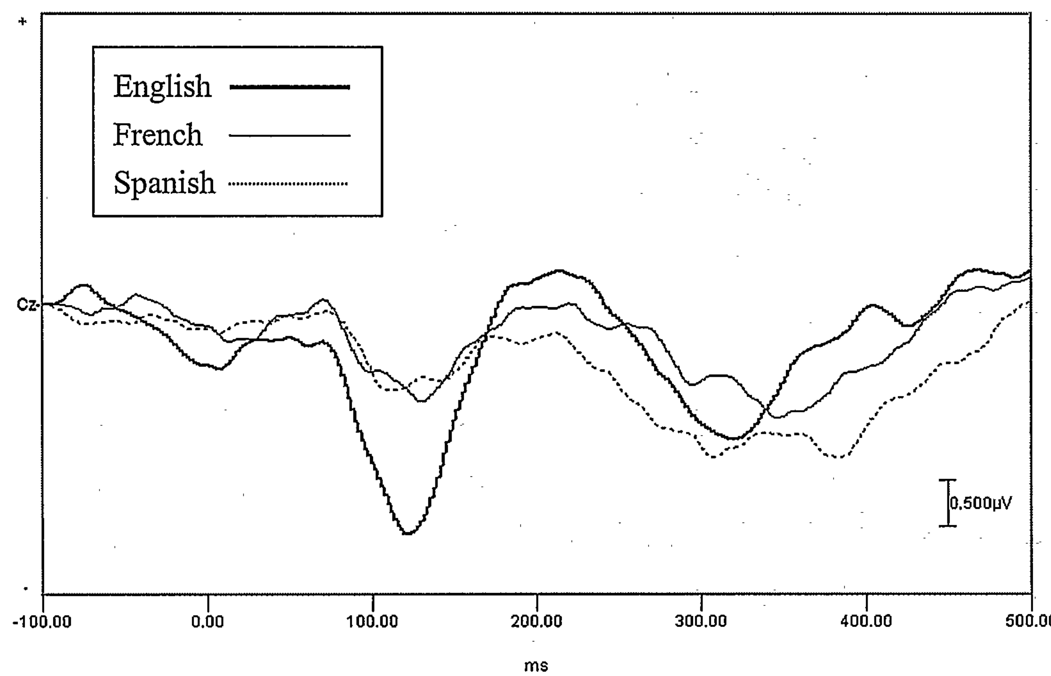


Note that the waveform responses are very different: /ɹ/ elicits a single N100 peak, whereas /b/ elicits a slightly later N100 that appears to have a second peak occurring at 176 msec post-stimulus. These differences indicate that the two segments are perceived as separate categories, and thus have distinct phonological representations. These sorts of differences in waveform morphology, then, will indicate the presence of across-category distinctions for the three rhotic sounds being examined here.

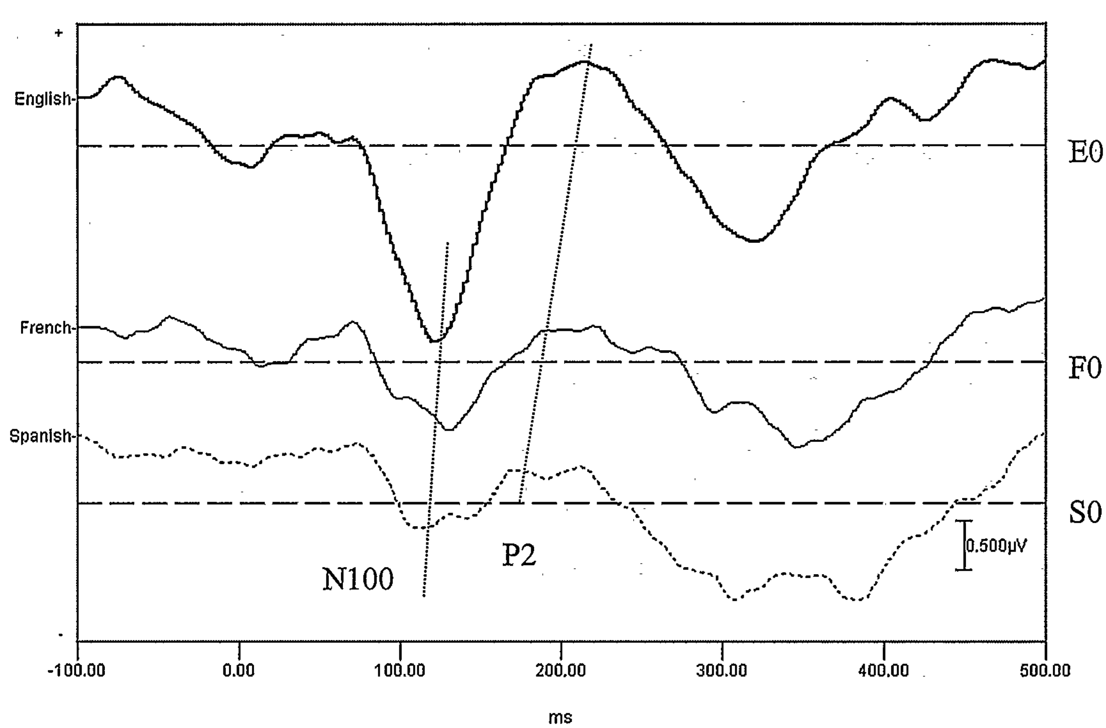
Figure 2.5 presents the grand-averaged waveform response data for all three rhotics obtained from subjects E1, E2, E3, E4, and E8. As above, both the overlaid and rastered formats of the waveform graph are presented; the averaged response to /ɹi/ is indicated by a thick line, the averaged response to /ʀi/ is indicated by a thin line, and the averaged response to /ri/ is indicated by a dashed line. Figure 2.6 presents individual averages, including the data obtained from E6.

Figure 2.5: Grand-averaged responses to /ɹ/, /ʀ/ and /r/

a)

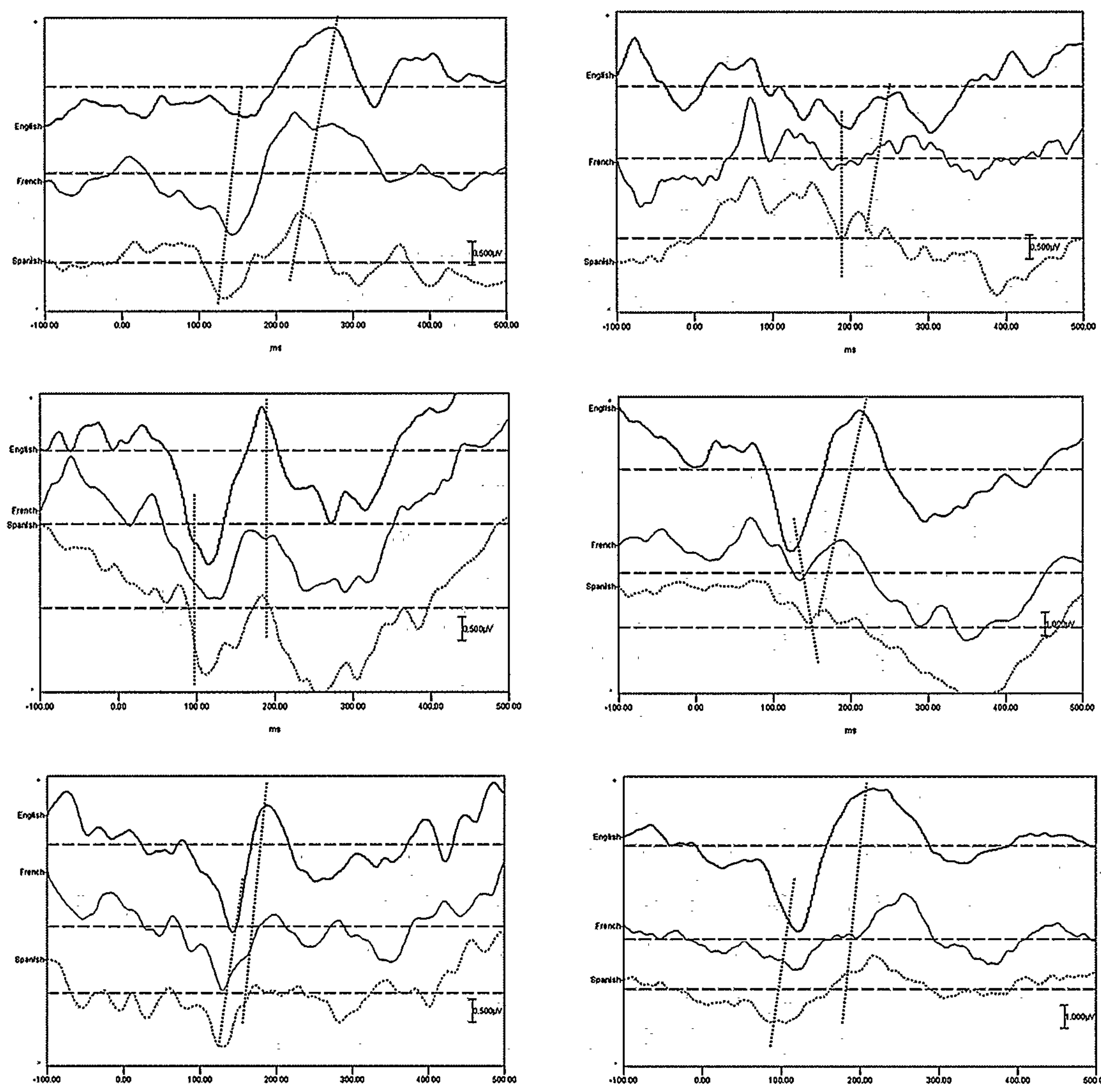


b)



The grand-averaged waveform data presents a fairly clear picture. For all three rhotic sounds, the N100 peak occurs within the same time frame. With respect to the N1 – P2 complex, /ɹ/ clearly differs from /ʀ/ and /r/, but these latter two do not appear to differ from each other in N1 – P2 amplitude.

Figure 2.6. Individual averaged waveforms



Two-tailed *t*-tests on the group's pooled data revealed no significant differences in peak latency for the N100 across the three rhotics examined (for /ɹ/ vs. /ʀ/ p = 0.535; for /ɹ/ vs. /r/ p = 0.556; for /ʀ/ vs. /r/ p = 0.272). Significant differences were found, however, with respect to N1 – P2 amplitude: the N1 – P2 complex elicited by /ɹ/ was found to have a significantly larger amplitude than that elicited by both /ʀ/ ($p < 0.05$) and /r/ ($p < 0.05$). No significant difference was found in N1 – P2 amplitude between /ʀ/ and /r/ (p = 0.186). Table 2.3 below presents the mean values obtained for each rhotic on each of the measures examined.

Table 2.3. Mean values

	N100 Latency	N1 – P2 Complex Amplitude
/ɹ/	124 msec	3.99 μV
/ʀ/	127 msec	2.46 μV
/r/	121 msec	2.17 μV

These results indicate that while all three rhotics were grouped into one single category, as evidenced by the homogeneous N100 latencies, the monolingual English speakers were making some distinctions among rhotic sounds. The larger N1 – P2 amplitude elicited by /ɹ/ is indicative of increased synchronous neural activity associated with this speech sound. The lack of difference in N1 – P2 amplitude between /ʀ/ and /r/, on the other hand, indicates that no distinction is made between the two sounds. This suggests, then, that these individuals are making a two-way within category distinction for rhotic sounds: /ɹ/ elicits neural behaviour that is distinct from both /ʀ/ and /r/, which in turn are not distinct from each other, yet all three sounds are recognized as belonging to a single category.

2.5. Discussion

Not all of the predictions set out in Section 2.2 above were borne out: while the three rhotics were categorized together with the native /ɹ/ being distinguished from the two non-native segments, no evidence was found to suggest that any distinctions were being made between the two non-native rhotics. These results, then, do not support the weak L1 feature hypothesis.

We cannot, however, directly compare the present results to those obtained by LaCharité and Prévost (1999), as the groups from which the data were collected vary considerably. LaCharité and Prévost's (1999) data came from a group of very advanced L2 English learners who were training to become English as a Second Language (ESL) teachers and had completed a graduate level course in English phonetics. Our data, on the other hand, came from individuals who had not previously studied a language containing either /ʀ/ or /r/, nor were they studying any such language at the time of testing. The fact that our present electrophysiological findings do not support the weak L1 feature hypothesis does not give us cause to reject it. Rather, we can interpret these results as indicating that the weak L1 feature hypothesis does not accurately account for the initial processing of novel L2 segments. Instead, our findings about the initial state of L2 phonological acquisition are more in keeping with the strong L1 feature hypothesis.

Recall from previous discussion in section 1.3 that the strong L1 feature hypothesis (Brown 2000) was formulated to account for the observation that not all new L2 segments were troublesome for L2 learners in the sense that some contrasts were perceived with ease in spite of the fact that they involved a new segment, while others were notoriously difficult to perceive. It also elegantly accounted for the fact that the L1 itself plays an important role in determining which new segments will present difficulty in acquisition: while both Mandarin Chinese and Japanese lack the English liquid contrast, only the L1 Japanese speakers encountered difficulty discriminating these sounds, while L1 Mandarin speakers performed at a level that was not significantly different from that of native English speakers. This difference was attributed to the feature [CORONAL], which is crucially required to maintain the contrast among English liquids. This feature is absent in Japanese and present in Mandarin. Furthermore, for L1 Japanese speakers, neither increased proficiency (Brown 2000) nor explicit training (Matthews 1997) were shown to bring about any improvement in the ability to perceive the English liquid contrast. In order to account for these observations, the strong L1 feature hypothesis maintains that the L1 grammar's influence in L2 acquisition is absolute.

The weak L1 feature hypothesis (LaCharité and Prévost 1999) was formulated in response to the observation that we find that some new segments which would otherwise be predicted to be unacquirable do show some improvement over time – an outcome not predicted by the strong L1 feature hypothesis. As previously noted, the results of the current electrophysiological experiment seem to support the predictions of the strong version: neural responses to the new rhotic trills were significantly different from that elicited by the native rhotic approximant, yet they were not significantly different from each other. Although these monolingual English speakers were categorizing all three together, or assigning all three rhotics to a single 'rhotic representation', they were maintaining a distinction within this category of '/ɹ/' vs. 'non-/ɹ/ rhotic'.

Our extrapolation of the weak L1 feature hypothesis to the properties of the initial state is not supported by the present electrophysiological data. We can still, however, evaluate its claims about the rate at which acquisition of novel segments proceeds by examining data from language learners of varying levels of proficiency. The weak L1 feature hypothesis may be taken to apply to the same domain of acquisition as the Similarity Differential Rate Hypothesis (SDRH) (Major and Kim 1996). The SDRH only makes predictions about the rate at which L2 learners will improve in producing new segments based on notions of similarity and difference between the L2 segments and existing segments in the L1 grammar, where "similar" L2 segments improve more slowly than do "different" L2 segments. Similarly, the present data suggest that while the claims of the weak L1 feature hypothesis might apply to the rate of acquisition, they do not apply to the initial state. While it was noted in the previous chapter that Brown's (2000) work demonstrates that Major and Kim's (1996) notions of similarity and difference cannot be maintained as they make incorrect predictions about L1 Korean speakers' abilities with the English liquid contrast, Brown (2000) only reports on a single group of L1 Korean speakers, all of whom had studied English for a number of years (the mean number of years for the group was 9.9), and the results came from a single battery of tests administered one time only. That is, we have no way of evaluating Major and Kim's claims about differences in rate of improvement with Brown's data, as it tells nothing of the developmental path followed by the learners she examined. This is perhaps the domain to which the weak L1 feature hypothesis applies: features representing terminal nodes are those that allow a learner to show more improvement on a given segment as acquisition proceeds, whereas those features representing articulator nodes are those that show decidedly less improvement for a particular segment. In order to evaluate this new hypothesis, we will want to examine how the neural responses to the rhotic trills examined here change as acquisition proceeds. The experiments described in the next chapter of this thesis achieve that end, in addition to addressing the broader question of whether it is possible to acquire new structure in L2 acquisition at all.

In sum, we have seen in this chapter that electrophysiological evidence suggests that the strong L1 feature hypothesis is better suited to accounting for the initial abilities of language learners with novel segments. There does not appear to be any evidence that learners make distinctions among new segments (for which their L1 grammar does not supply a required feature) on the basis of the type of feature required; there is no significant difference between the neural responses elicited by segments requiring absent articulator nodes and those neural responses elicited by segments requiring absent terminal nodes. It was further suggested that the weak L1 feature hypothesis might be best applied to the rate of acquisition, or improvement, rather than initial interference effects. This is the matter at hand in the following chapter.

Chapter 3

On the developmental path of adult L2 phonological acquisition

3.1. The results thus far

The previous chapter reported on an electrophysiological test of a proposed model of L1 interference in L2 phonological acquisition – the weak L1 feature hypothesis (LaCharité and Prévost 1999). This model claims that the observation that some novel L2 segments present greater difficulty in acquisition than others is due to the type of phonological features required by the L2 which are absent from the L1 grammar. More specifically, the claim is that those segments which require a missing articulator node will be more difficult to acquire than those segments which require a missing terminal node. Our prediction was that if L2 learners are indeed sensitive to these kinds of distinctions, they should be reflected in the neural activity elicited by novel segments at the initial state, even before acquisition begins. This was tested by eliciting event-related potential (ERP) data from monolingual English speakers in response to two novel rhotic trills: the alveolar trill /r/ found in Spanish (Dalbor 1997), and the uvular trill /ʀ/ found in French (Walsh Dickey 1997). It was predicted that these two sounds would elicit different neural responses, as English lacks the terminal node [vibrant] required for both /r/ and /ʀ/, but crucially lacks the articulator node [PHARYNGEAL] required for /ʀ/.

Two evoked potential responses were examined: the N100, which reflects the categorization of speech segments (Sharma and Dorman 1999) and, in turn, underlying phonological representations, and the N1-P2 complex, which reflects the amount of synchronous neural activity associated with a speech segment (Tremblay et al. 2001). The N100 results indicated that monolingual English speakers assign all three rhotics to a single category. The N1-P2 complex results suggest that while these individuals are making some distinction between the native rhotic and the two novel rhotic segments, no distinction is found to differentiate the French /ʀ/ and the Spanish /r/ from each other. Our extension of the weak L1 feature hypothesis predicted that even at the initial state, we should find evidence that some distinction is maintained between novel L2 segments on the basis of the type of feature required for an appropriate representation that is absent from the L1 grammar. That is, contrary to our predictions, the featural composition of the new segments does not seem to have any neural consequences, as both novel rhotic segments elicit nearly identical neural responses.

The present chapter reports on a series of ERP investigations that sought to test the predictions of the weak L1 feature hypothesis as acquisition proceeds. Our prediction is that, if terminal node features are more easily acquired than articulator node features, we will see changes in the neural responses elicited by /r/ at an earlier stage in L2 Spanish learners than those elicited by /ʀ/ in L2 French learners. More specifically, we would expect to see earlier evidence of a new category being established for /r/, when comparing beginner and advanced L2 learners, as indicated by differences in the N100 response with respect to peak latency as well as the general shape of the deflection (i.e., single peak vs. multiple peaks, as seen in the control condition in Chapter 2). We will also be looking for indications of increased neural activity associated with this segment as indicated by an increase in the N1 – P2 complex amplitude seen when comparing L2 learner responses with monolingual responses. Again, we expect that increases in N1 – P2 complex amplitude will occur at earlier stages of acquisition for /r/ than for /ʀ/; that is, we expect to find this increase in beginner L2 Spanish learners, but not in beginner L2 French learners.

3.2. Data Collection

3.2.1. Stimuli

The stimuli used in this experiment were the syllables /ɹi/, /ʀi/, and /ri/, previously used in the experiment described in Chapter 2. The syllables were elicited in isolation from a male native speaker of English who is also a near-native speaker of both French and Spanish. Each syllable was repeated three times, and the second repetition was used in preparing the task items. The syllables were recorded using a Sony TCD-D100 DAT recorder and Sony ECM-MS908C electret condenser microphone. These were then edited using Peak LE 2.62 digital audio editing software on a Macintosh desktop computer to eliminate empty portions of the sound files. The syllables were then transferred to a desktop PC computer, which was used to present the stimuli during the experiment using Neuroscan Inc.'s STIM stimulus presentation software package.

3.2.2. Subjects

Nine subjects participated as beginner L2 learners: all were undergraduate students at the University of Calgary with a mean age of 20. Four were registered in an introductory French language course, five were registered in an introductory Spanish language course; both courses provided 8 months of classroom exposure to the L2. Our goal was to recruit individuals whose first language was English and who had begun studying either French or Spanish as an adult; furthermore, we were looking for individuals who had not acquired or studied another language with a trilled rhotic in its phonemic inventory prior to beginning their studies in French or Spanish. In spite of our best efforts, our participants were found to vary considerably in terms of their experience with language; this is summarized in Table 3.1 below. Moreover, due to the time-consuming nature of our data collection methods, we were unable to recruit additional participants for study upon recognition of this variation.

Table 3.1. Summary of beginner L2 learners examined

	L1	**Previous L2**	**Age Started/Years Studied**	**Immersion**
Beginner French (BF)				
BF1	English	German	7/9	none
BF2	English	none	--/--	--
BF3	English	French	5/9	6 years
BF4	English	French German	5/7 15/1	3 years
Beginner Spanish (BS)				
BS1	English	French Russian	12/6 20/3	none
BS2	Polish	English French	5/13 9/9	13 years none
BS3	English	French Spanish	9/6 15/4	none none
BS4	English	French	9/6	none
BS5	English	French Finnish	12/2 18/1	none 1 year

One subject, BS2, had acquired Polish as a first language but claimed that English has been her dominant language since childhood. Subject BS5 was a native speaker of

English but had spent a year in Finland and thus had an intermediate level of proficiency in Finnish. Subject BF1 was a native speaker of English but had studied German as a second language for nine years and thus had an intermediate-to-advanced level of proficiency in German. Subjects BF3 and BF4 were both native speakers of English who had spent a minimum of two years enrolled in French immersion programs in elementary school, but had not continued with the program through junior and senior high school. Subject BS3 indicated that her mother's L1 was Tagalog, and although she did not consider herself to be a speaker of Tagalog, her data were not included in the group averages calculated for analysis purposes. Such variation means that we cannot pool the data obtained from all subjects and accurately compare it against the previously obtained monolingual data. Therefore, the discussion here will be based on the data obtained from subjects BS1, BS4, and BF2. Discussion of the results obtained from subjects BS2, BS5, BF1, BF3, and BF4 can be found in Chapter 4.

Three subjects participated as advanced L2 learners: all were undergraduate students at the University of Calgary with a mean age of 25. One was registered in a senior-level French language course, two were registered in a senior-level Spanish language course. All three were native speakers of English and had begun acquiring their L2 in either late childhood or as adults: subject AF1 studied French in a non-immersion classroom setting starting at the age of 10, subject AS1 studied Spanish at the University of Calgary for three years and spent three months in Spain, subject AS2 studied Spanish at the University of Calgary for two years and spent 10 months studying abroad in Mexico. An additional two subjects participated as near-native L2 speakers: one (VAS) was a graduate student in the Department of French, Italian, and Spanish at the University of Calgary, age 29, the other (VAF) was a teacher-librarian with the Calgary Public School Board, age 55, who has spent many years both teaching French in public schools as well as serving on bilingual library committees for the school board. Both were native speakers of English, VAS had begun acquiring her L2 as an adult and reported having lived in Spain for 3 years, VAF had begun acquiring her L2 in a junior high school French class that was not part of an immersion program. All subjects gave informed consent for their participation in this research, and were paid $20 for their participation.

Our discussion of the results will be based on a comparison of the elicited neural responses, so as to identify any changes that occur as proficiency increases. This type of analysis also allows us to identify any developmental paths that may emerge. It should be noted that the previous chapter's findings came from an analysis of group-averaged, or grand-averaged data, which means that the data obtained from all suitable participants was pooled into one large group for averaging purposes. This method of analysis was not possible here; where more than one suitable subject was found, the data is pooled into a group, however we also report here on responses obtained from individual participants. In any group a certain amount of individual variation is found, so it should not be assumed that the following graphs all illustrate robust group responses. They do, however, serve our purpose here in establishing a developmental path in phonological acquisition.

3.2.3. Procedure

The experiment was conducted at the Language Research Centre at the University of Calgary. Surface electrodes were placed using a Neuroscan Quik Cap electrode cap, which arranges the electrode array according to the International 10-20 system (Jasper 1958, cited in Tremblay et al. 2001), along with a forehead ground and right mastoid reference (A2). Data recording was done using a Neuroscan NuAmps digital EEG amplifier and Neuroscan's Scan 4.2 and 4.3 Acquisition software package[8], run on a Dell Inspiron laptop computer. Stimuli were presented using Neuroscan's Stim stimulus presentation software package on a desktop PC. Analysis of raw data was performed off-line using Neuroscan's Scan 4.3 Edit software package, run on a Dell Inspiron laptop computer.

Subjects watched a video while being fitted with the electrode cap. Once all the equipment had been properly set up and configured, the experiment began. A mismatch paradigm was used, in which a stream of identical syllables is occasionally interrupted by a different syllable, which is referred to as the deviant stimulus. The stimulus syllables were paired as described in Table 3.2 below; for all conditions, 85% of the tokens consisted of the syllable labelled as the standard, and 15% of the tokens consisted of the syllable labelled as the deviant. Syllables were presented with an onset-to-onset interstimulus interval of 850 milliseconds. For all conditions, four blocks of 250 syllables were presented, for a total of 1000 tokens. Stimuli were presented to the subject's right ear via earphone at a volume of 75 dB, and subjects were asked to indicate

[8] Recording for subject BS1 was done using Scan 4.2, an upgrade became available shortly after, allowing us to record all additional data from other subjects using Scan 4.3.

the occurrence of the deviant stimulus by pressing a button on a response pad held in their left hand using their left thumb. In contrast with the method used for testing the monolingual English speakers, the L2 learners only listened to pairs containing the rhotic trill used in their L2; that is, French learners only heard pairs consisting of /ɹi/ and /ʀi/, while Spanish learners only heard pairs consisting of /ɹi/ and /ri/.

Table 3.2. Stimulus pairings

Standard	Deviant	
/ɹi/	/ʀi/	French learners
/ʀi/	/ɹi/	
/ɹi/	/ri/	Spanish learners
/ri/	/ɹi/	

A number of methodological standards were followed in recording the data: electroencephalogram (EEG) channels were analog bandpass filtered on-line from 0.5 Hz to 70 Hz, amplified with a gain X 1338 and converted using an Analog-to-Digital (A/D) rate of 500 Hz. Eyeblink artifact was removed off-line, and the continuous data files were segmented into recorded sweeps consisting of 100 millisecond pre-stimulus, 500 millisecond post-stimulus epochs. Epochs containing voltage deflections greater than 100 μV were deemed to be contaminated by artifact and were rejected off-line. Remaining sweeps were baseline corrected and filtered off-line using a 30 Hz low pass filter (12 dB/octave). All data and resulting waveforms were recorded from the electrode Cz, following Sharma and Dorman (1999) and Tremblay et al. (2001).

3.3. Results

Figure 3.1 below presents the group-averaged responses elicited by /ɹ/ and /r/ from the beginner L2 Spanish learners, Figure 3.2 presents the averaged response elicited by /ɹ/ and /ʀ/ from the beginner L2 French learner.

Figure 3.1. Group-averaged responses to /ɹ/ and /r/ by beginner Spanish learners

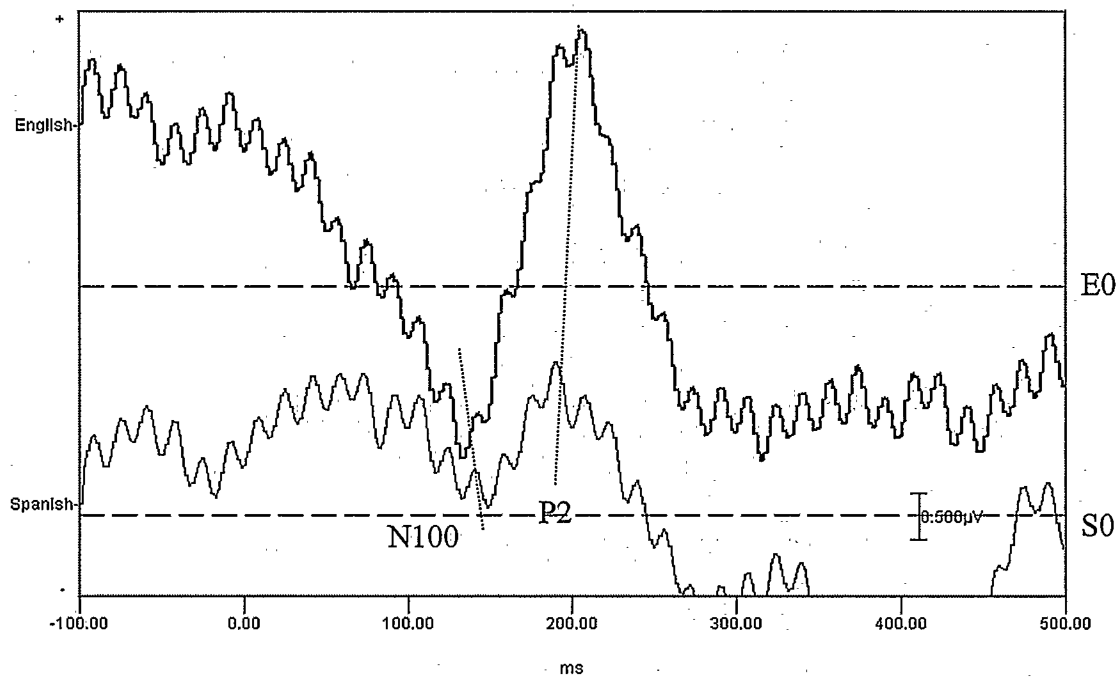
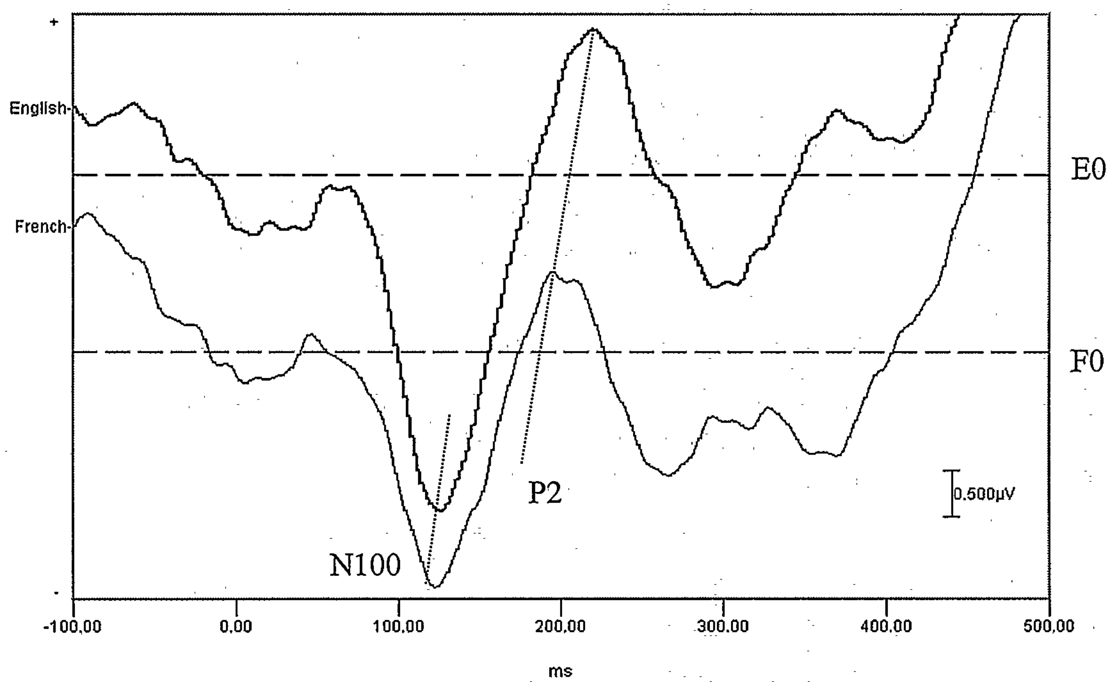


Figure3.2. Averaged responses to /ɹ/ and /ʀ/ by a beginner French learner



We can see that for both groups of L2 learners, the N100 peak latency occurs in the same time range for both the native and novel rhotics, which was the finding for monolingual English speakers when tested on these same segments. This suggests that the learners are still assigning the novel rhotic segment to the existing category established for /ɹ/.

The N1 – P2 complex, however, does show some differences between the two groups. The L2 Spanish group shows an N1 – P2 complex that is smaller for /r/ (amplitude of 1.562 µV) when compared to that obtained for /ɹ/ (amplitude of 4.434 µV), a result that was also obtained among the monolingual group previously examined. For the L2 French learner, however, a fairly large N1 – P2 complex was elicited in response to /ʀ/ (amplitude of 3.358 µV), though it does not appear to be as large as that elicited in response to /ɹ/ (amplitude of 5.162 μV). This indicates that after 8 months' L2 exposure, greater synchronous neural activity is elicited in response to /ʀ/ than to /r/.

Figure 3.3 below presents the group-averaged responses elicited by /ɹ/ and /r/ from advanced Spanish learners; Figure 3.4 below presents the averaged responses elicited by /ɹ/ and /ʀ/ from the advanced French learner.

Figure 3.3. Group-averaged responses to /ɹ/ and /r/ by advanced Spanish learners

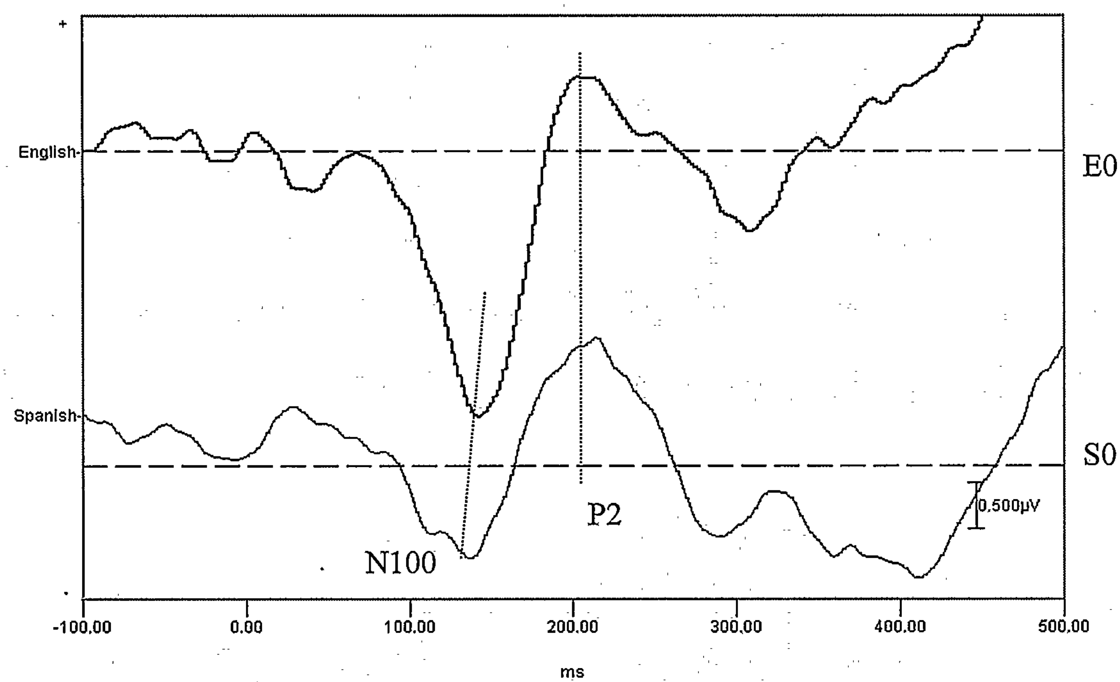
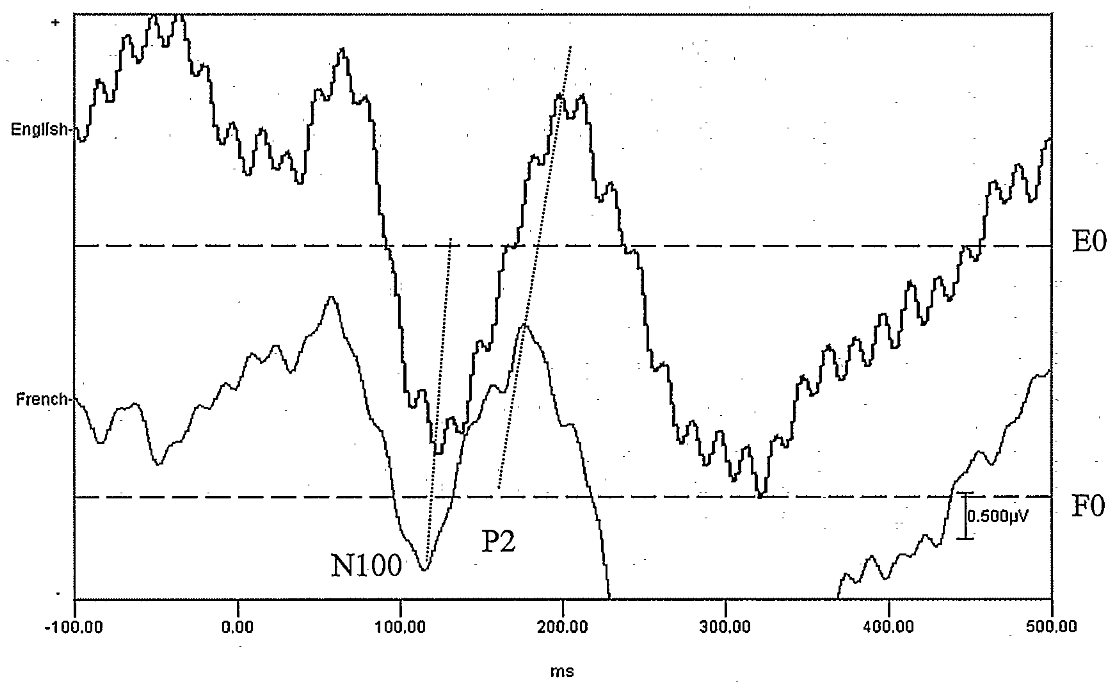


Figure 3.4. Averaged responses to /ɹ/ and /ʀ/ by an advanced French learner



Again, we see that for both L2 learner groups the N100 peak latency obtained in response to /r/ occurs within the same time frame as that obtained in response to /ɹ/, suggesting that the two segments are being assigned to a single category.

With respect to the N1 – P2 complex, here we see among the advanced Spanish speakers a much larger N1 – P2 complex obtained in response to /r/ (amplitude of 2.35 µV), though it is still smaller than that obtained in response to /ɹ/ (amplitude of 3.647 µV). Similarly, the N1 – P2 complex obtained from the advanced French speaker in response to /ʀ/ is quite large (amplitude of 2.635 µV), though not as large as that elicited by /ɹ/ (amplitude of 3.865 µV).

The data gathered thus far allow us to sketch out a progression in L2 segmental acquisition by comparing the neural responses elicited for a given segment at the three stages examined thus far. Figure 3.5 below compiles the results obtained for /r/; Figure 3.6 below compiles the results obtained for /ʀ/.

Figure 3.5. Comparison of initial state, beginner learners, and advanced learners – Spanish

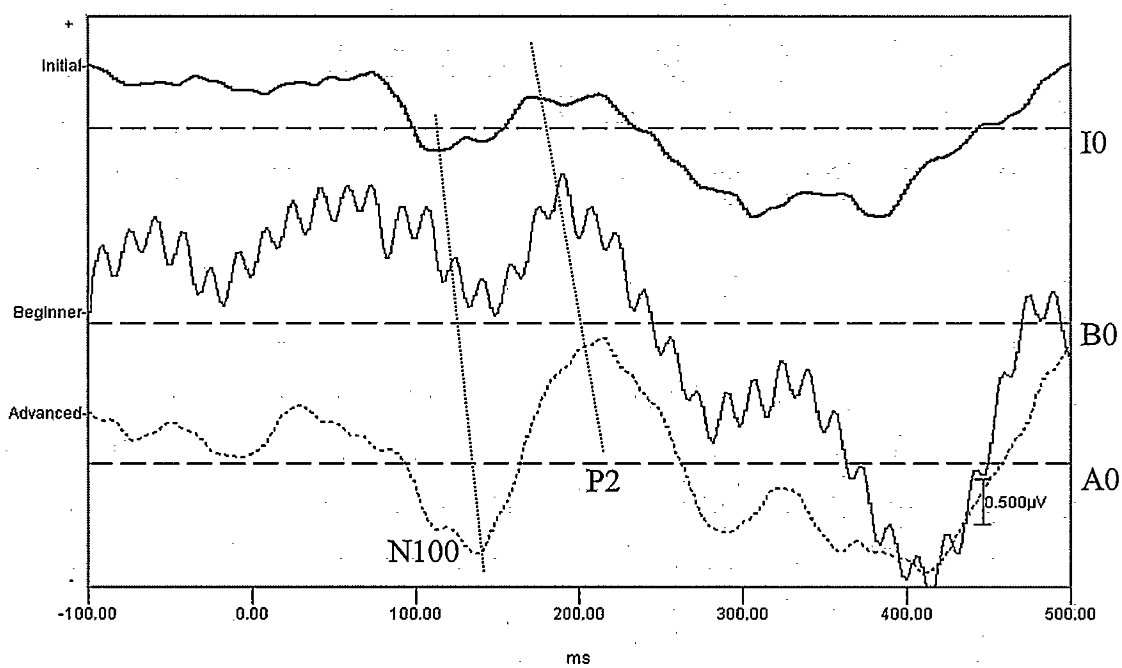
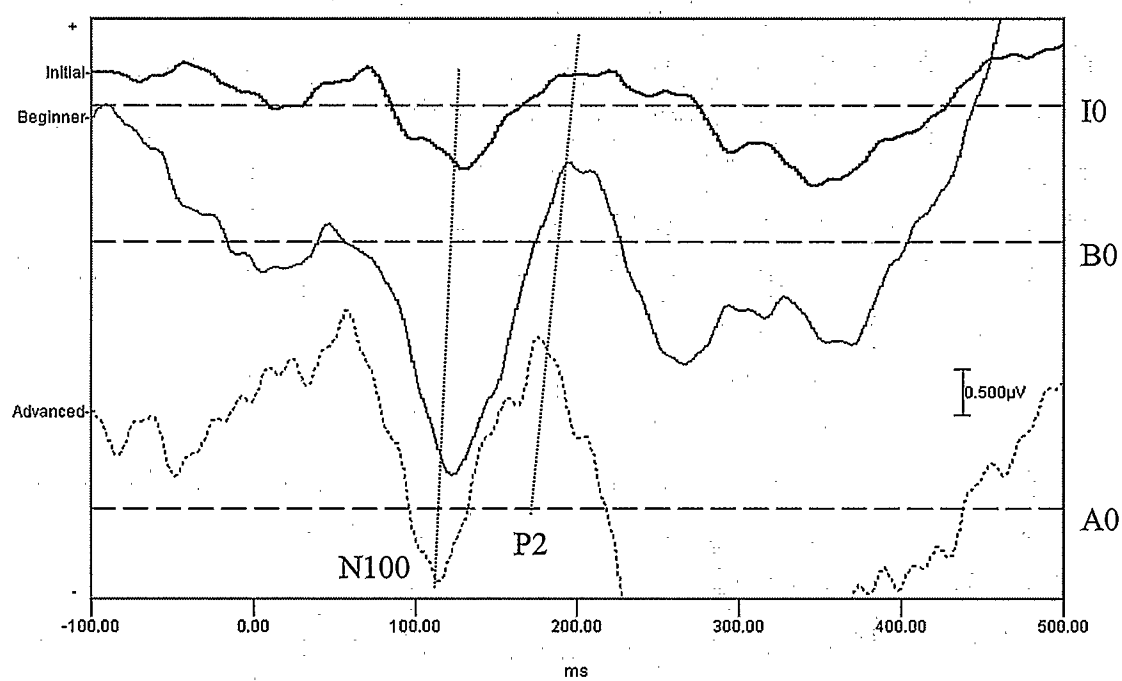


Figure 3.6. Comparison of initial state, beginner learner, and advanced learner – French



These figures show that the N100 peak latency does not appear to undergo any changes in either group; that is, at all stages observed here, the new rhotic segment is assigned to the existing rhotic category. Further, we also see that the increase in N1 – P2 complex amplitude occurs at a later stage of acquisition among L2 Spanish speakers for the alveolar trill /r/ than it does among L2 French speakers for the uvular trill /ʀ/. This reflects that the experience-related increase in synchronous neural activity observed for both trills occurs at an earlier stage of acquisition for the uvular trill.

The finding that the N100 peak latency does not appear to undergo any change as acquisition proceeds prompted us to gather additional data from near-native speakers of French and Spanish. The goal here was to examine whether it was possible for L2 learners to establish separate categories for new segments in their end-state grammars.

Figure 3.7 below presents the averaged responses obtained from a near-native Spanish speaker; Figure 3.8 below presents the averaged responses obtained from a near-native French speaker.

Figure 3.7. Averaged responses to /ɹ/ and /r/ by a near-native Spanish speaker

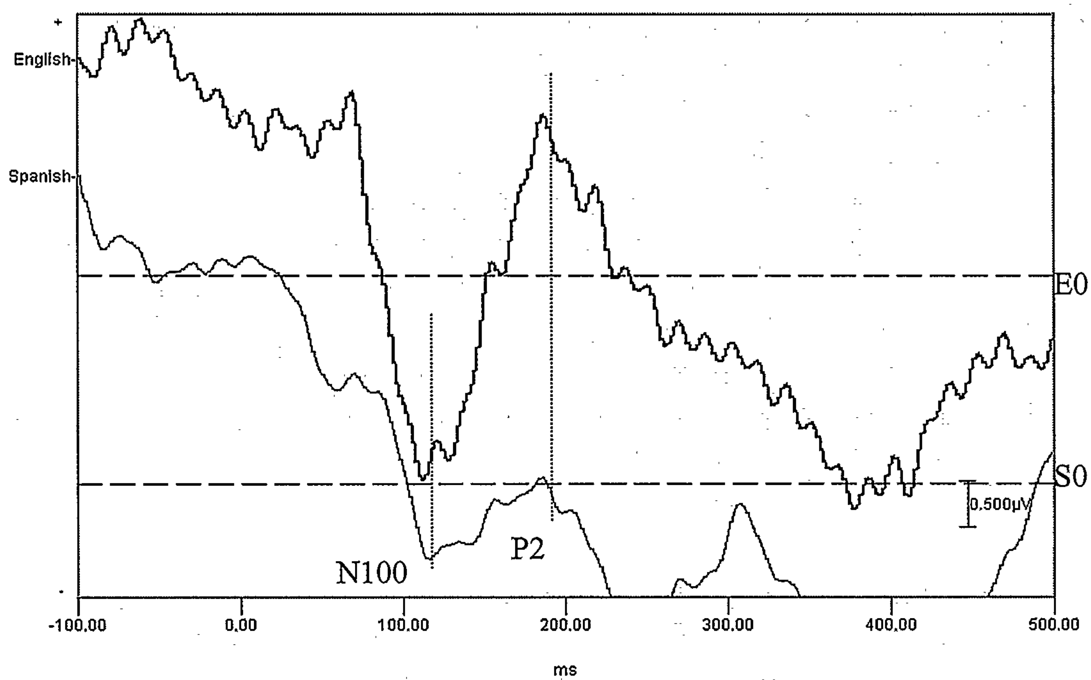
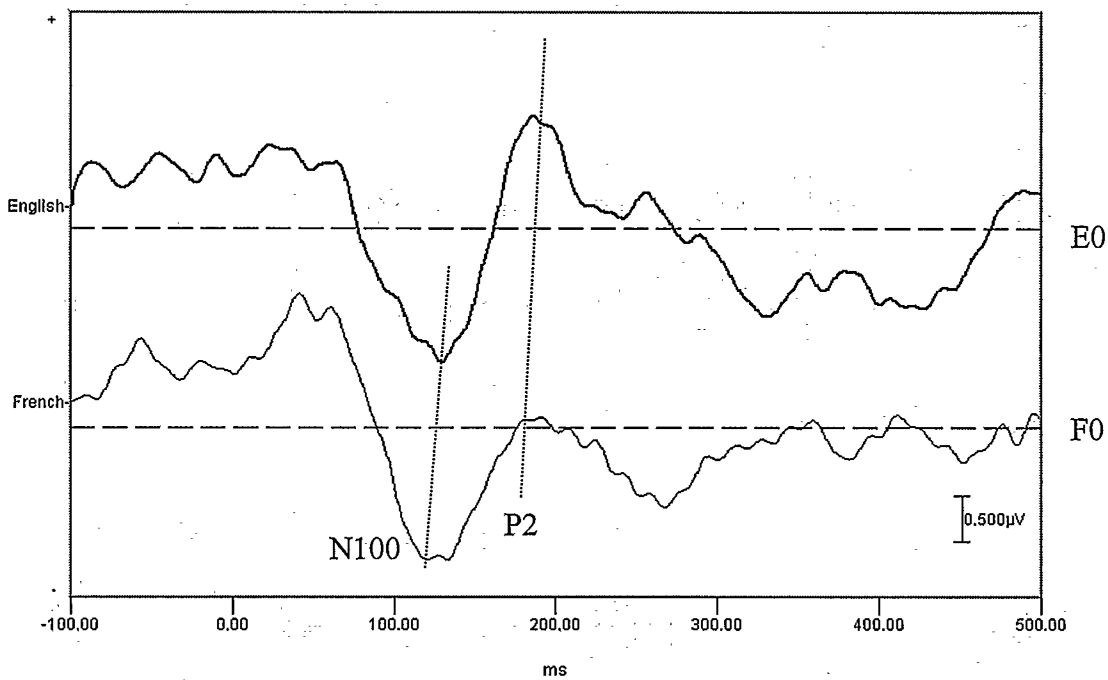


Figure 3.8. Averaged responses to /ɹ/ and /ʀ/ by a near-native French speaker



These results paint a picture much like that previously seen with other subject groups: the N100 peak latency occurs in the same time frame for both types of rhotic segments. Even near-native speakers assign the L2 rhotics to the L1 rhotic category. Further, for the near-native Spanish speaker, the N1-P2 complex obtained in response to /r/ is much smaller than that obtained for /ɹ/; these results are counter-intuitive given the pattern seen thus far in that they more closely resemble those obtained from our monolingual and beginner L2 learner subject groups. The near-native French speaker shows a pattern much like that obtained for the advanced L2 learner.

3.4. Discussion

The results obtained here show an intriguing pattern: they suggest that adult speakers cannot form new categories for novel L2 segments. Further, the data indicate that the neural responses elicited by /ʀ/ undergo a change at an earlier stage in acquisition than do those elicited by /r/. Neither of these findings support the weak L1 feature hypothesis.

Indeed, the N100 peak latency findings show support for the strong L1 feature hypothesis, which claims that new phonological features, regardless of whether they are terminal nodes or articulator nodes, cannot be acquired in L2 acquisition. Under this view, we would not expect to see any changes in neural responses to novel L2 segments even over time, as the claim is that learners cannot build appropriate representations for the new categories. This is precisely the pattern observed with respect to N100 peak latency: regardless of L2 proficiency, the N100 peak elicited by both /r/ and /ʀ/ occurs in the same time frame as that elicited by /ɹ/. That is, even at very advanced stages of acquisition, /r/ and /ʀ/ are funneled into the existing category for /ɹ/. As previously stated, this finding suggests that adult L2 learners are unable to form new categories for novel L2 segments, which in turn indicates that they are unable to construct appropriate representations for these segments. To make a broader claim, we can interpret these findings as indicating that adult L2 learners cannot acquire novel phonological features.

This is not to say that the interlanguage grammar does not undergo any changes as acquisition proceeds, as we do observe some changes in neural responses with increasing L2 proficiency. The amplitude of the N1 – P2 complex was found to increase with increasing proficiency (with the surprising exception of our near-native Spanish speaker VAS), and this increase was found to occur at different times for /r/ and /ʀ/; however, the results obtained are the opposite of those predicted by the weak L1 feature hypothesis. We saw an increase in N1 – P2 complex amplitude for /ʀ/ occurring at an earlier stage of acquisition than that observed for /r/. Recall that the weak L1 feature hypothesis predicts that here /ʀ/ should be more difficult to acquire than /r/ as L1 English speakers are crucially missing the articulator node [PHARYNGEAL], whereas for /r/ these same speakers lack only the terminal node [vibrant]. We therefore expected to find earlier increases in N1 – P2 complex amplitude for /r/.

The N1 – P2 complex reflects the amount of synchronous neural activity associated with a given speech segment (Tremblay et al. 2001). Increases in its amplitude therefore reflect increases in neural activity for the perception of that particular segment. Tremblay et al. (2001) showed that this increase co-occurred with an increase in subjects' ability to perceive a non-native speech segment contrast, and argued that the increased neural activity underlies improved within-category perception. The same argument can be made here: we have evidence that L1 English speakers are not forming new categories for novel L2 segments, but rather are improving their ability to make a within-category distinction. The notion of degree of difficulty, which the weak L1 feature hypothesis attributes to the type of feature that is missing from the L1 grammar, is assumed to be reflected in the electrophysiology by the rate of change to the N1 – P2 complex: those segments that are easier to acquire will elicit responses with larger N1 –

P2 complex amplitudes at earlier stages of acquisition, as easier segments will show earlier improvement.

The question remains, however, why /ʀ/ should show improvement before /r/, as its featural content predicts that it should be more difficult for L1 English speakers to acquire. It is unlikely that the weak L1 feature hypothesis is completely wrong, as it provides an elegant account of the behaviour observed in L1 Japanese L2 English speakers (Matthews 1997) as well as L1 Quebec French L2 English speakers (LaCharité and Prévost 1999). This finding may be due to differences in the phonemic inventories of French and Spanish with respect to rhotic sounds: French has one rhotic while Spanish maintains an intervocalic contrast among two rhotics, the alveolar trill /r/ and the alveolar tap /ɾ/. This is illustrated by the phoneme inventory charts given below in Table 3.3 for French (Walker 2001: 119), and Table 3.4 for Spanish (adapted from Dalbor 1997: 103)[9].

Table 3.3. Consonant phoneme inventory of French

		Labial	*Apical*[10]	*Palatal*	*Velar*	*Uvular*
Stops	voiceless	p	t	k		
	voiced	b	d	g		
Fricatives	voiceless	f	s	ʃ		
	voiced	v	z	ʒ		
Sonorants	nasals	m	n	ɲ	ŋ	
	liquids		l			ʀ[11]

[9] Dalbor (1997) does not use International Phonetic Association (IPA) notation in his transcriptions; for expository reasons, we have adapted his chart to IPA notation.
[10] The term 'apical' is used to refer to sounds that are articulated using the tongue tip, i.e., coronal sounds (Rogers 1991).
[11] Walker (2001) uses the symbol [ʁ], which according to the IPA is a voiced uvular fricative, and not a liquid; uvular segments are all assumed to have a [PHARYNGEAL] node.

Table 3.4. Consonant phoneme inventory of Spanish

		Labial	*Interdental*	*Dental*	*Alveolar*	*Palatal*	*Velar*
Stops	voiceless	p		t			k
	voiced	b		d			g
Fricatives	voiceless	f	(θ)		s		x
	voiced						
Affricates	voiceless					tʃ	
Sonorants	nasals	m			n	ɲ	
	rhotics				r, ɾ		
	approximants				l	j	w

The data presented here suggest that adult L2 learners will assign novel rhotic sounds to an existing rhotic category; that is, all r-sounds are grouped together perceptually. The L2 French learner learns to make a two-way within category distinction with respect to rhotic sounds: within the phonemic category of /ɹ/, he must learn to reliably distinguish between [ɹ] and [ʀ]. The L2 Spanish learner, however, must learn to make a three-way within category distinction among rhotics: within the phonemic category of /ɹ/, he must learn to reliably distinguish between [ɹ], [r], and [ɾ]. This added dimension of complexity may be responsible for the observed delay in improvement on /r/ among L1 English speakers.

On the whole, the electrophysiological data suggest a view of L2 phonological acquisition in keeping with the strong L1 feature hypothesis rather than the weak L1 feature hypothesis. The evidence indicates that adult L2 learners do not create new categories for new segments, regardless of the amount of L2 exposure or level of proficiency. This in turn suggests that these learners are unable to build appropriate phonological representations if the L1 grammar fails to supply all the required features; that is, these missing phonological features do not become available by some other means as acquisition proceeds. Instead, learners appear to be improving their abilities over time on within-category discrimination. The observed patterns of improvement, however, are not consistent with the predictions of the weak L1 feature hypothesis with respect to the degree of difficulty associated with the type of feature that is missing from the L1 grammar.

Chapter 4

The effects of additional language experience

4.1. The 'other' cases

The experiments reported in previous chapters examined the neural responses obtained from L1 English speakers of varying proficiency in either French or Spanish. The results from both sets of experiments suggest that with respect to the question of acquiring new phonological structure, the strong L1 feature hypothesis finds more empirical support, in that regardless of the type of feature required to appropriately represent a given novel segment, if that feature is not supplied by the L1 grammar, the learner will be unable to build an appropriate phonological representation for it, and will thus categorize it with an existing L1 segment.

It was noted, however, that in recruiting participants for these experiments, a number of individuals were identified to have had additional experience with language that did not permit them to be included in the previous analyses. Their data, however, provide us with the opportunity to further address other issues in the study of second language acquisition. In particular, we are able to explore issues regarding the effects of the age of initial exposure, as well as gather additional insight into the role of the length of L2 exposure. A substantial portion of the existing literature in the field of second language acquisition is concerned with the role of the age of initial L2 exposure and the length of L2 exposure (Yamada 1995, Werker 1995, Pallier et al. 1997, Long 1990, Newport 1990). The general finding is that the earlier the initial L2 exposure, the better (Long 1990), although Pallier et al.'s (1997) results of an examination of vowel perception among L1 Spanish L2 Catalan speakers suggest that even very early L2 exposure (beginning in preschool) is insufficient to trigger formation of separate categories for a vowel contrast present in Catalan but absent in Spanish (Werker (1995) notes, however, that it is unknown whether experience affects the perception of vowels in the same way as the perception of consonants). Work by Yamada (1995) on L1 Japanese speaker's perceptual abilities with respect to English /l/ and /ɹ/ supports the 'earlier is better' approach, but Yamada argues that her data do not allow us to tease apart the role of age of initial L2 exposure from that of the length of L2 exposure, as those subject groups who had earlier initial exposure also had longer exposure. The data discussed in the present chapter may afford us the opportunity to examine the effects of each of these factors; we have already seen that for adult L2 learners, a greater period of L2 exposure does not allow formation of new categories. What remains to be seen is how the length of L2 exposure affects acquisition which begins in early childhood.

Consider the following scenarios: one subject reported that her L1 is not English, but Polish, yet claimed that she had not used her L1 in several years and now considers English to be her primary language. Unlike English, Polish does contain an alveolar trill /r/ like the one found in Spanish (Kopczyński 1977). This subject's L1 grammar, then, has all the necessary features to build an appropriate representation for /r/ in Spanish. As her L2 acquisition of English began at the age of 5, it is of interest to see whether she has successfully established a novel category, and in turn a separate phonological representation, for /ɹ/. Indeed, we expect that her neural responses will reflect the existence of separate categories for /r/ and /ɹ/: not only did acquisition begin very early in life, but the representation of /ɹ/ is identical to that of /r/ with the notable exception of /r/ requiring the feature [vibrant].

Similarly, three additional subjects reported having completed at least three years of French immersion in early elementary school with little or no classroom exposure to French thereafter, and one subject reported having studied L2 German for nine years beginning at the age of 7. These individuals, then, afford us the opportunity to further assess the role of age of initial exposure as well as the effect of the length of L2 exposure. We have seen in the previous chapter that adult L2 learners are unable to establish distinct phonological representations for novel L2 segments for which they lack the appropriate features in their L1, regardless of the length of L2 exposure. In other words, a lengthy period of L2 exposure alone does not permit acquisition of new phonological features. The data examined here will allow us to make inferences about the effects of age of initial L2 exposure alone; that is, does L2 exposure that begins before the age of 9 (early childhood) enable learners to acquire features the L1 lacks and thus construct appropriate phonological representations for L2 segments even when the length of L2 exposure is shortened? The current chapter seeks to find answers to this question.

4.2. Data Collection

4.2.1. Stimuli

The stimuli used in this experiment were the syllables /ɹi/, /ʀi/, and /ri/, previously used in the experiments described in Chapters 2 and 3. The syllables were elicited in isolation from a male native speaker of English who is also a near-native speaker of both French and Spanish. Each syllable was repeated three times, and the second repetition was used in preparing the task items. The syllables were recorded using a Sony TCD-D100 DAT recorder and Sony ECM-MS908C electret condenser microphone. These were then edited using Peak LE 2.62 digital audio editing software on a Macintosh desktop computer to eliminate empty portions of the sound files. The syllables were then transferred to a desktop PC computer, which was used to present the stimuli during the experiment using Neuroscan Inc.'s STIM stimulus presentation software package.

4.2.2. Subjects

The data discussed here was obtained from subjects originally recruited to participate in the experiments described in previous chapters. All subjects gave informed consent for their participation in this research, and were paid $20 for their participation. As previously discussed, these individuals were found to not fit the desired L2 learner profile, either due to previous acquisition of second languages other than French and Spanish (E5, BF1, BS5), due to childhood exposure to French (E7, BF3, BF4), or due to having a language other than English (BS2 – Polish) as a first language.

As was the case in Chapter 3, the discussion here will be based on a comparison of elicited neural responses from individual participants. Therefore, it should not be assumed that these are robust group responses. Again, they do serve a purpose here in allowing us to gain a glimpse of the interplay of age of initial L2 exposure and length of L2 exposure.

4.2.3. Procedure

The experiment was conducted at the Language Research Centre at the University of Calgary. Surface electrodes were placed using a Neuroscan Quik Cap electrode cap, which arranges the electrode array according to the International 10-20 system (Jasper 1958, cited in Tremblay et al. 2001), along with a forehead ground and right mastoid reference (A2). Data recording was done using a Neuroscan NuAmps digital EEG amplifier and Neuroscan's Scan 4.2 and 4.3 Acquisition software package[12], run on a Dell Inspiron laptop computer. Stimuli were presented using Neuroscan's Stim stimulus presentation software package on a desktop PC. Analysis of raw data was performed off-line using Neuroscan's Scan 4.3 Edit software package, run on a Dell Inspiron laptop computer.

Subjects watched a video while being fitted with the electrode cap. Once all the equipment had been properly set up and configured, the experiment began. A mismatch paradigm was used, in which a stream of identical syllables is occasionally interrupted by a different syllable, which is referred to as the deviant stimulus. The stimulus syllables were paired as described in Table 3.1 below; for all conditions, 85% of the tokens consisted of the syllable labelled as the standard, and 15% of the tokens consisted of the syllable labelled as the deviant. Syllables were presented with an onset-to-onset interstimulus interval of 850 milliseconds. Two different stimuli presentation methods were used, reflecting that these subjects were recruited for different experiments. Subjects E5 and E7, recruited as monolingual English speakers, listened to eight blocks

[12] Recording for subjects E5 and E7 was done using Scan 4.2, an upgrade became available shortly after, allowing us to record all additional data from other subjects using Scan 4.3.

of 250 syllables for each of the stimulus pairings given in Table 3.1, for a total of 2000 tokens for each syllable[13]. Subjects BF1, BF3, and BF4 were recruited as beginner French learners; subjects BS2 and BS5 were recruited as beginner Spanish learners. These subjects, for all conditions, listened to four blocks of 250 syllables, for a total of 1000 tokens for each syllable. Stimuli were presented to the subject's right ear via earphone at a volume of 75 dB, and subjects were asked to indicate the occurrence of the deviant stimulus by pressing a button on a response pad held in their left hand using their left thumb. Those subjects who were recruited as monolingual English speakers listened to all contrasts listed in Table 4.1 below; those subjects recruited as beginner L2 learners only listened to pairs containing the rhotic trill used in their L2; that is, French learners only heard pairs consisting of /ɹi/ and /ʀi/, while Spanish learners only heard pairs consisting of /ɹi/ and /ri/.

Table 4.1. Stimulus pairings

Standard	Deviant	
/ɹi/	/ʀi/	French learners
/ʀi/	/ɹi/	
/ɹi/	/ri/	Spanish learners
/ri/	/ɹi/	

A number of methodological standards were followed in recording the data: electroencephalogram (EEG) channels were analog bandpass filtered on-line from 0.5 Hz to 70 Hz, amplified with a gain X 1338 and converted using an Analog-to-Digital (A/D)

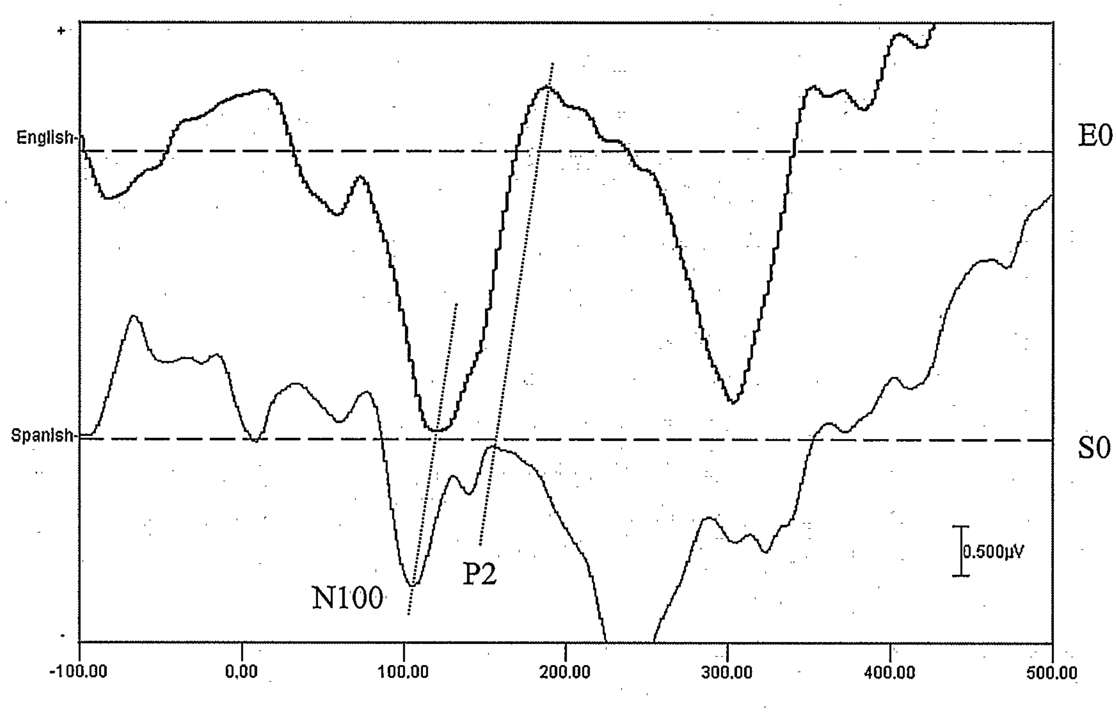
[13] Due to a technical problem encountered during recording, not all data was available for analysis.

rate of 500 Hz. Eyeblink artifact was removed off-line, and the continuous data files were segmented into recorded sweeps consisting of 100 millisecond pre-stimulus, 500 millisecond post-stimulus epochs. Epochs containing voltage deflections greater than 100 μV were deemed to be contaminated by artifact and were rejected off-line. Remaining epochs were baseline corrected and filtered off-line using a 30 Hz low pass filter (12 dB/octave). All data and resulting waveforms were recorded from the electrode Cz, following Sharma and Dorman (1999) and Tremblay et al. (2001).

4.3. Results

Figure 4.1 presents the results obtained from BS2, whose L1 is Polish. She began acquiring L2 English in school in kindergarten at the age of 5.

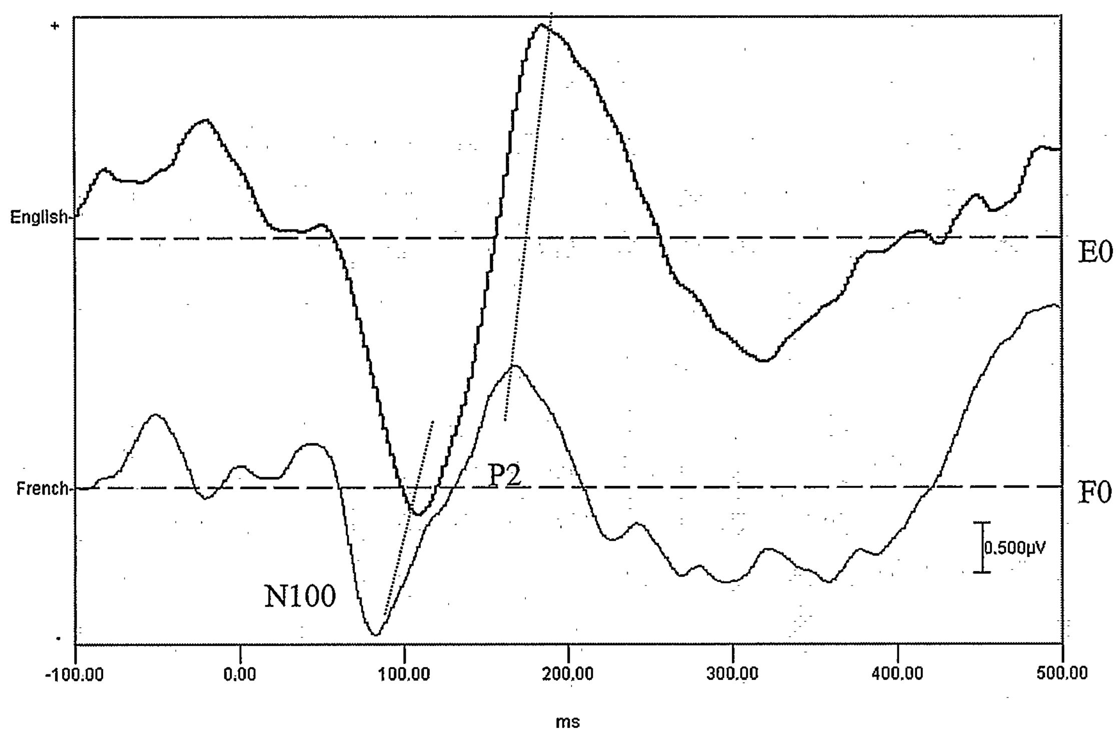
Figure 4.1. Averaged responses to /ɹ/ and /r/ from an L1 Polish speaker



Recall that we are assuming that waveforms with different N100 peak latencies or distinct N100 waveform morphologies were indicative of separate phonological representations for the segments being examined, following Sharma and Dorman (1999). That is, a difference in response along either dimension with respect to the N100 would be viewed as evidence of separate categories. Our previously discussed results showed very similar N100 responses for the rhotics being examined for any given participant in that peak latencies were nearly identical for all three rhotics, suggesting a single underlying representation. The data presented above illustrate a very different scenario, as the N100 deflections differ in overall appearance: the N100 elicited by /r/ shows a second negativity occurring 140 msec post-stimulus. Additionally, the main N100 peak elicited by /r/ occurs earlier than that elicited by /ɹ/. This suggests that separate categories with separate underlying phonological representations are available for the two rhotic segments. With respect to the N1 – P2 complex, we find here a picture similar to that we have already seen, in that the amplitude of the N1 – P2 complex elicited by the L2 segment is smaller when compared with that elicited by /ɹ/. Considering the fact that for this particular participant, /r/ is actually an L1 segment while /ɹ/ is not, this finding is particularly interesting: all data considered thus far has suggested that the N1 – P2 complex is largest for the L1 rhotic. The observed reversal for this participant may be related to her claim that she is more comfortable using English (her L2) than Polish (her L1).

Figure 4.2 presents the results obtained from BF3, who was enrolled in a French immersion program throughout elementary school beginning in grade 1 at the age of 6. He continued studying French in a classroom setting for two years in junior high and one year in high school.

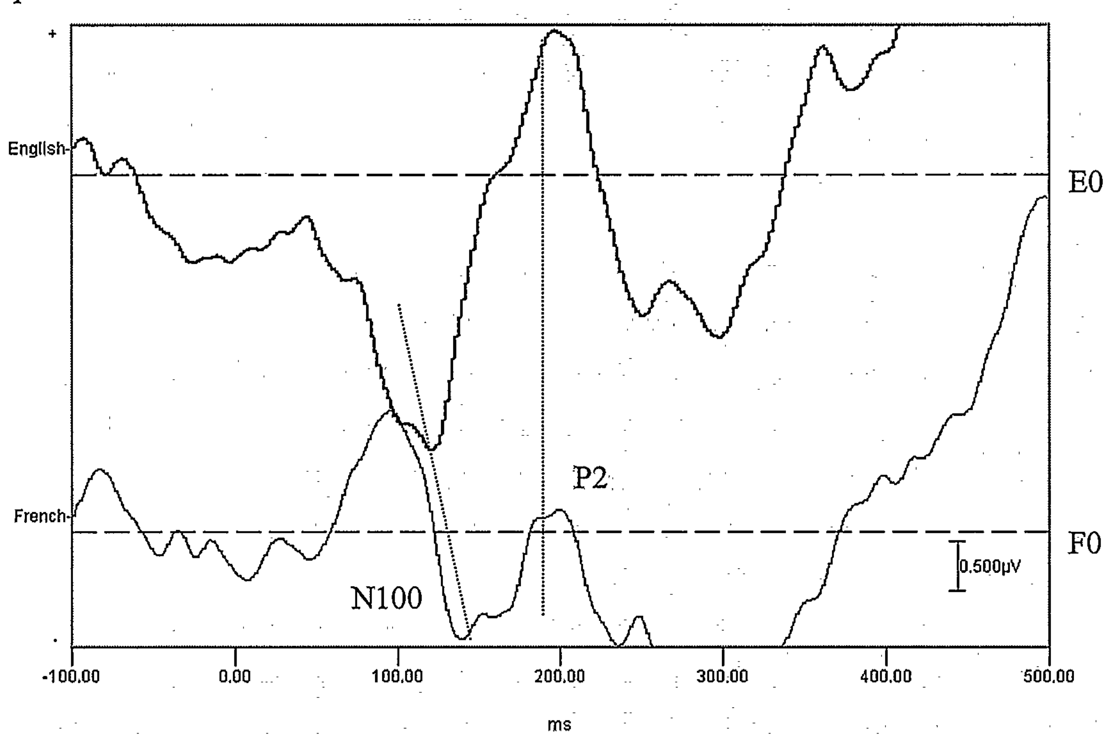
Figure 4.2. Averaged responses to /ɹ/ and /ʀ/ from an L1 English speaker with childhood exposure to French



Again, as was seen in Figure 4.1 above, we can make note of a difference in N100 overall shape and peak latency in the neural responses to the two rhotics: the N100 peak elicited by /ʀ/ occurs earlier than that elicited by /ɹ/, an even greater difference than that observed in Figure 4.1 above. This again suggests separate categories with appropriate phonological representations for the two rhotic segments being examined. The N1 – P2 complex is also larger for /ɹ/ than for /ʀ/, which is consistent with the pattern emerging in all data considered thus far.

Figure 4.3 below presents the results obtained from BF1, who began studying German in a non-immersion setting at the age of 7 and had continued to do so up until the time of testing at age 19, save for a period of three years (junior high school). Like French, German has /ʀ/ in its phonemic inventory (Prowe 2001).

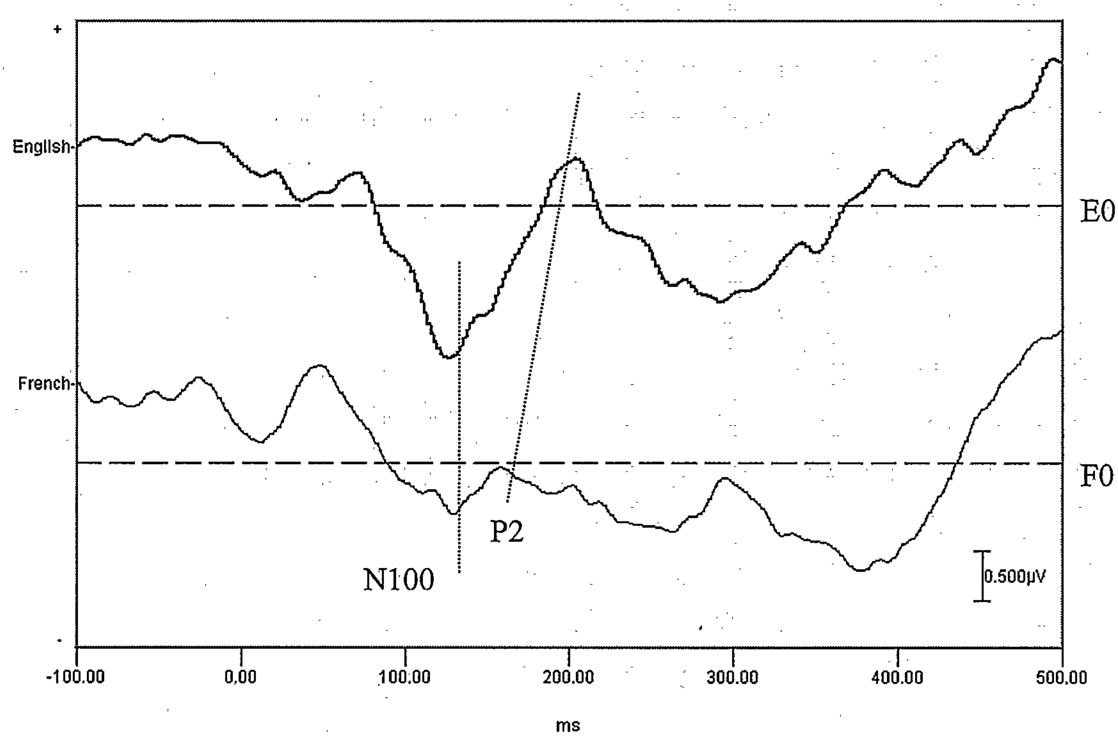
Figure 4.3. Averaged responses to /ɹ/ and /ʀ/ from an L1 English speaker with childhood exposure to German



Here, while the N100 overall shape and peak latencies obtained for the two rhotics differ, it is interesting to note that the direction of difference is the opposite of what was observed for BF3: the N100 response for /ʀ/ occurs later than that for /ɹ/. In spite of this difference, overall these results fall in line with the pattern that is beginning to emerge – exposure to the L2 rhotic during childhood results in different neural responses, suggesting distinct phonological processes.

Figure 4.4 below presents the results obtained from Subject BF4, who was enrolled in a French immersion program in elementary school from kindergarten (age 5) up until grade 2. He continued studying French in a non-immersion classroom setting for four additional years (up until grade 6).

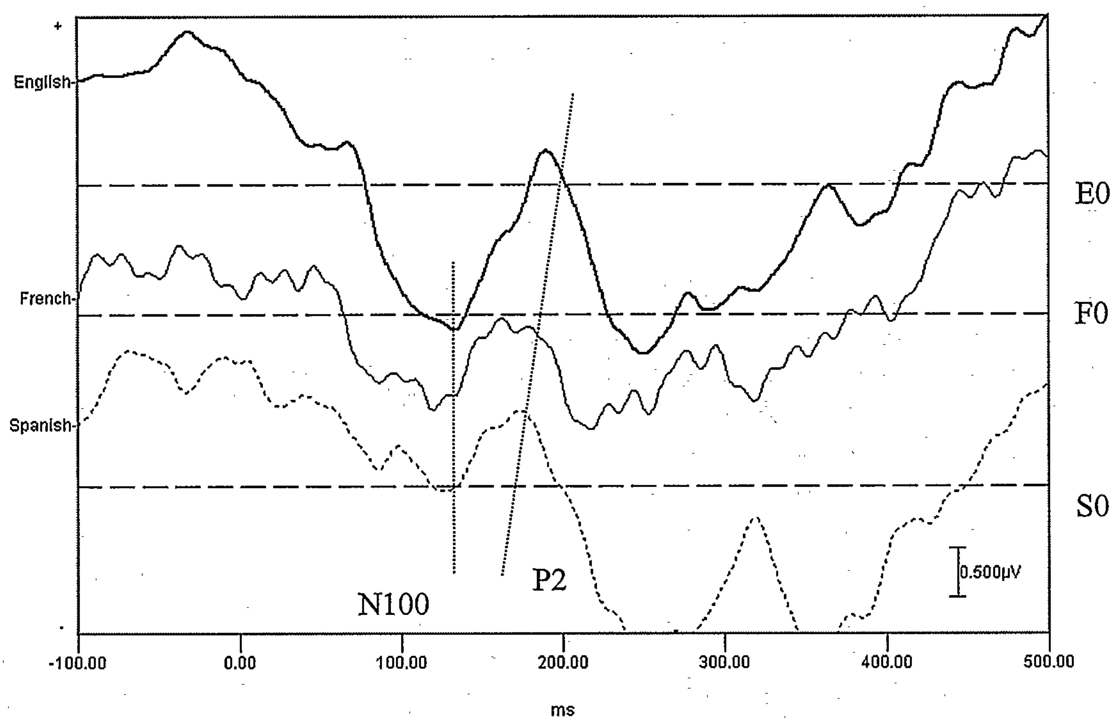
Figure 4.4. Averaged responses to /ɹ/ and /ʀ/ from an L1 English speaker with some childhood exposure to French



These results are quite different from those obtained from subject BF3: for both rhotics, the N100 responses look remarkably similar, and the N100 peak latency occurs within the same time frame. That is, this individual appears to be categorizing the two rhotic segments within a single category – he does not have distinct phonological representations for these. Further, the N1 – P2 complex elicited by /ʀ/ is much smaller than that observed for BF3, reflecting a lesser amount of synchronous neural activity occurring in response to this segment.

Figure 4.5 below presents the results from E7, who was recruited as a monolingual English speaker but was found to have had three years exposure to French through an immersion program in elementary school beginning in kindergarten (age 5).

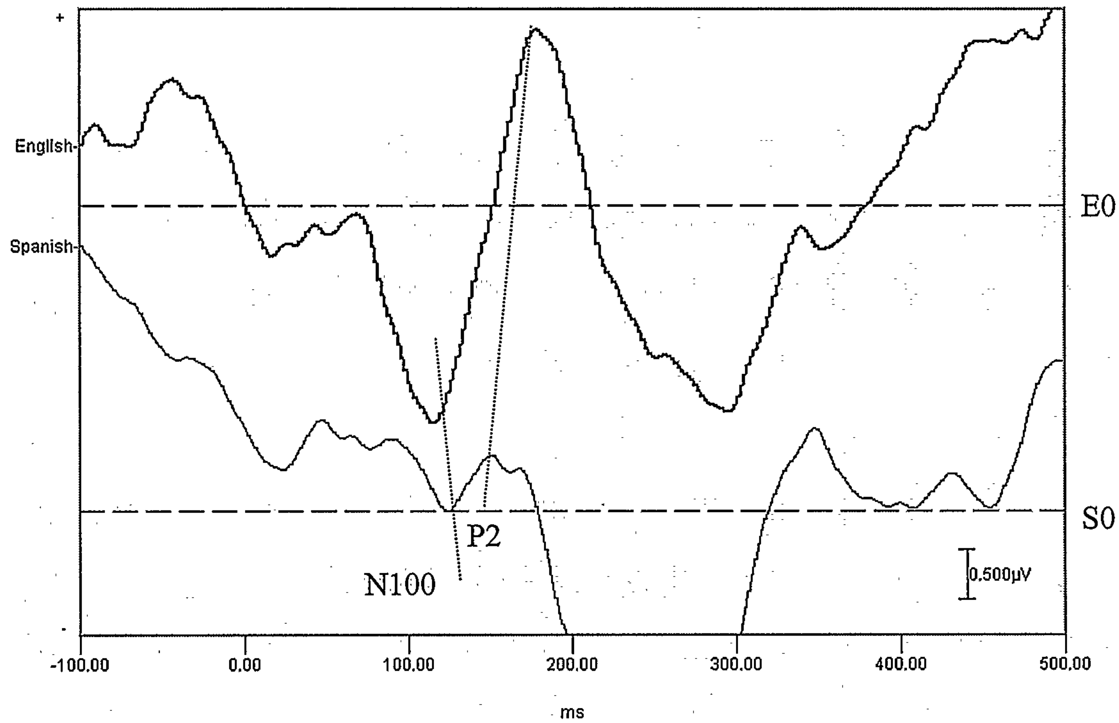
Figure 4.5. Averaged responses to /ɹ/, /ʀ/, and /r/ from an L1 English speaker with some childhood exposure to French



Much like what was observed with BF4, this participant seems to follow the pattern that was established among monolinguals (Chapter 2) and adult L2 learners (Chapter 3), in that the N100 overall shape and peak latency for all three rhotics is very similar, indicating that all three segments are being assigned to a single category. Further, N1 – P2 complex amplitude seems to set a distinction between the native /ɹ/ and the non-native /ʀ/ and /r/, with the native segment eliciting greater synchronous neural activity. Again, this is the pattern that was observed to occur in monolingual English speakers in Chapter 2.

Two additional participants were not included in earlier analyses owing to the fact that they had acquired a second language other than French or Spanish as adults. Figure 4.6 below presents the results obtained from BS5, who was an intermediate speaker of Finnish. Finnish, like Spanish, contains an alveolar trill /r/ in its phonemic inventory (Karlsson 1999). This individual began acquiring Finnish in an immersion setting at the age of 18.

Figure 4.6. Averaged responses to /ɹ/ and /r/ from an L1 English speaker with adult exposure to Finnish

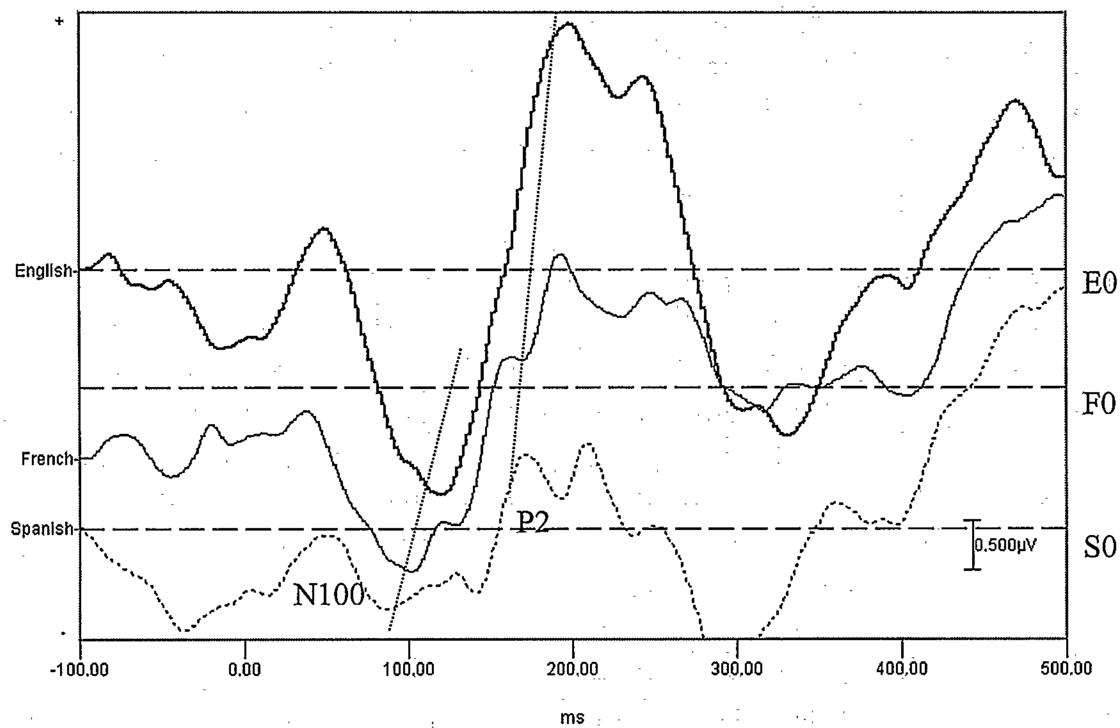


These results also follow the general pattern observed among monolingual English speakers and adult L2 learners of varying proficiency: both rhotics elicit N100 responses with similar shape and peak latency, and the native /ɹ/ elicits a much larger N1 – P2 complex than the non-native /r/. That is, both /ɹ/ and /r/ are assigned to a single (L1) category.

Figure 4.7 below presents the results obtained from E5, a native speaker of English who was also an intermediate speaker of Italian. This individual began studying

Italian in a non-immersion classroom set in an immersion environment[14] at the age of 15. Like Spanish, Italian has /r/ in its phonemic inventory (Messora 1983).

Figure 4.7. Averaged responses for /ɹ/, /ʀ/, and /r/ from an L1 English speaker with adult exposure to Italian



These results are noticeably different from those obtained thus far: this individual seems to have distinct neural responses for the three rhotics with respect to both the N100 as well as the N1 – P2 complex amplitude. The N100 response elicited by /r/ occurs with an earlier peak latency, coupled with a noticeable second negativity occurring 140 msec post-stimulus. Further, the N100 elicited by /ʀ/ does not describe a smooth curve as it approaches P2, whereas the N100 elicited by /ɹ/ does. This suggests that this speaker has

[14] This individual reports attending a boarding school in Italy where all material was taught in English, save two Italian language courses.

established separate categories with appropriate phonological representations for all three rhotic segments under consideration. The data considered in previous chapters led us to predict the precise opposite outcome for this individual: we would have expected to see evidence of a single representation for all rhotics. Given the relatively small number of participants examined throughout this research, we cannot say with certainty whether E5's data should be considered a genuine counterexample to the analysis being formulated thus far, or if we should consider her an exceptional language learner, in that her electrophysiological data suggest that she is able to acquire novel phonological features and thus construct appropriate underlying representations to form new categories for classifying rhotic segments. Only further expansion of the data pool will allow us to decisively select one of these possibilities as being correct.

4.4. Discussion

The results of the present analysis suggest a pattern that is in keeping with the one emerging from the previous chapters: adult L2 learners cannot acquire new phonological features. The overall shape and peak latency of N100 responses elicited by L1 and L2 rhotics indicate that these segments are assigned to a single category, and thus are assigned a single phonological representation: that appropriate to the L1 rhotic.

We can note, however, that a different pattern emerges for those speakers who have had childhood exposure to a second language (BF1, BF3). These individuals have neural responses that indicate that they have acquired some L2 phonological feature that would otherwise be absent from the L1; here, both individuals have acquired both the [PHARYNGEAL] node and the [vibrant] node required for an appropriate representation for /R/. The presence of these features allows them to establish separate categories for the two rhotics. Of course, this raises the question of how phonological features are extracted from an acoustic input; reasons of space prevent us from giving this topic a full and adequate treatment here (see Carroll (2001) for discussion of these issues). Many distinguishing characteristics of speech segments, such as place and manner of articulation, can be recovered from the acoustic signal (Borden et al. (1994): place of articulation is revealed through formant transitions, stops have a characteristic noise burst, vowel height and backness are correlated with formant frequencies, to name only a few examples. Brown (1997) proposes that this acoustic input is funneled into phonetic categories, which are in turn funneled into phonemic categories by available phonological structure. Acquisition of a particular feature, then, would arise as the result of two phonetic categories being used contrastively in the input, prompting the learner to establish some way to differentiate between the two phonologically. That is, learners would need to detect contrastive usage of two sets of acoustic signals in order to retrieve an appropriate phonological feature from UG that would allow phonological discrimination of these. It seems likely that both contrastive usage within one language and contrastive usage across two distinct languages would both serve to drive phonological acquisition in this view. Given that BF1 and BF3 differ from the participants examined in previous chapters with respect to the age of initial L2 exposure, we can attribute the observed differences in neural patterns to this difference in exposure with some confidence: BF1 and BF3 received the necessary input to drive phonological acquisition at a stage when they were still able to retrieve novel phonological features from UG. The adult L2 learners discussed in Chapter 3, however, were unable to do so.

The results further indicate that early childhood L2 exposure is necessary but not sufficient for the acquisition of L2 phonological features; in other words, there is some threshold level on the amount of exposure required. It appears that exposing a young child to a second language for a period of two or three years is not sufficient to trigger acquisition of new phonological features. While both E7 and BF4 attended a French immersion program in elementary school beginning at 5 or 6 years of age, much like BF1 and BF3, neither shows evidence of having acquired the phonological features appropriate to /ʀ/; N100 responses from both individuals indicate that they, like the adult L2 learners, are assigning both rhotics to a single category. Both rhotics are assigned a phonological structure that is appropriate to the native /ɹ/ only. It appears that what distinguishes individuals like E7 and BF4 from individuals such as BF1 and BF3 is the length of the period of L2 exposure, summarized in Table 4.2 below.

Table 4.2. Summary of early exposure to /ʀ/ and /ɹ/

	Age of initial L2 exposure	**Length of L2 exposure – immersion**	**Length of L2 exposure – non-immersion**	**Distinct representations?**
BF1	7	3 months	9 years	yes
BF3	6	6 years	3 years	yes
BF4	5	3 years	3 years	no
E7	5	3 years	--	no

Both E7 and BF4 experienced a marked decrease in L2 exposure after three years: E7 had no further exposure to French after three years of immersion at school; BF4 had

French immersion at school up through grade 2, followed by a single French class each year through to grade 6. Both BF1 and BF3 experienced a longer period of L2 exposure: BF1 had classroom exposure to L2 German from the age of 7 onward, except for a period of three years while attending junior high school, thus giving her classroom exposure over a period that at the time of testing spanned approximately 9 years; BF3 had French immersion at school throughout elementary school, and continued with classroom exposure to French for two years of junior high and one year of high school. It appears to be the case that it is not enough to simply expose children to a second language, they must be exposed over a sufficient period of time before they are able to acquire phonological features that are absent from the L1. An open question that remains for future research is whether children such as E7 and BF4 ever show any evidence of having acquired these features – that is, are the features acquired, then lost due to lack of use? Or are they not acquired at all with this amount of exposure?

The data obtained from BS2 also makes an interesting contribution to our discussion of these issues, as this individual reported that although Polish was her first language, English had become her dominant language, and she indicated that she considered her proficiency in Polish to be at an intermediate level. As Polish has an alveolar trill in its phonemic inventory, much like Spanish, we would expect that this individual would have the feature [vibrant] in her L1 feature geometry. It was of interest, then, to see whether this feature was still available for constructing a representation for /r/ in Spanish given that English had become her dominant language. The N100 responses indicated that [vibrant] was indeed present and available for the L2 phoneme representation. This suggests a kind of permanence for the L1 grammar: although for this speaker L2 English was dominant, the features of her L1 remained available for constructing phonological representations for segments found in additional languages.

It appears to be the case, then, that phonological features can be acquired and added to the grammar in L1 acquisition, or in early L2 acquisition, but there appear to be constraints on the latter. That is, both age of initial L2 exposure and length of L2 exposure play a role in determining whether new L2 phonological features can be acquired, such that the age of initial L2 exposure alone is sufficient to prevent adults from acquiring L2 phonological features, but insufficient in allowing young children to acquire these. In this case, the length of L2 exposure seems to play an important role. The results obtained from BS5 serve to further bolster the findings of Chapter 3: adults, regardless of amount of L2 exposure and proficiency, are unable to acquire new phonological features in a second language.

Chapter 5

Summary and Implications

5.1. General findings and implications

The discussion of the first chapter of this thesis established four questions to be addressed through experimental research:

1) Do monolingual English speakers differentiate between English /ɹ/ and other rhotics?
2) Do monolingual English speakers make distinctions among non-native speech segments based on absent phonological features?
3) Can adult language learners trigger new structure (here, features) and build appropriate phonological representations for new segments?
4) How does experience with a new segment in a second language affect the acquisition of this same segment in a third language?

Investigation was carried out through analysis of event-related potential (ERP) data, which provide information about the perception and neural processing and encoding of speech segments. The particular neural responses examined were the N100, which is a major negative deflection in the waveform occurring between 80 and 140 milliseconds following the presentation of some auditory stimulus, and the N1 – P2 complex, which is the portion of the waveform consisting of the N100 and the following positive deflection, the P2. The N100 has been shown to reflect the categorical perception of speech segments (Sharma and Dorman 1999), such that speech segments which are assigned to separate phonological categories will elicit N100 responses with distinct peak latencies and waveform morphologies. The N1 – P2 complex has been shown to reflect the amount of synchronous neural activity associated with the perception of an auditory stimulus (Tremblay et al. 2001), such that an increase in neural activity is indicated by an increase in N1 – P2 complex amplitude. The N1 – P2 complex has also been shown to reflect experience-related changes in neural processing, in that increases in accuracy on the perception of a novel sound contrast brought about by training were correlated with increases in N1 – P2 complex amplitude (Tremblay et al. 2001). Thus, distinct categories (and consequently distinct phonological representations) are evidenced by distinct N100 response patterns, whereas degree of difficulty can be inferred by examining the rate of change in neural responses, with respect to both the N100 and the N1 – P2 complex: those segments that are easier to acquire due to their feature content will show changes in their neural processing at earlier stages of acquisition: either the N100 responses would be expected to diverge, indicating that separate categories have been established for the segments being examined, or the N1 – P2 complex amplitude would be expected to increase, indicating greater neural activity for those segments. In other words, if L2 learners must acquire either some feature x or some feature y and x is easier to acquire than y, then we would expect to see evidence of feature x in responses from both beginner and intermediate L2 learners, but evidence of feature y should only be found in responses from the intermediate learners. We are now in a position to summarize the results of the experimental research that was carried out in order to find answers to these questions, and discuss the broader implications of our findings.

Recall that the first three questions were formulated upon consideration of two models of L1 interference in L2 phonological acquisition: the strong L1 feature hypothesis (Brown 2000) and the weak L1 feature hypothesis (LaCharité and Prévost 1999). The strong version claimed that only those phonological features that were active in the L1 grammar were available for building phonological representations for novel L2 segments, and if a given segment required a feature that was absent in the L1, that segment could not be acquired: the learner would never be able to build an appropriate underlying representation and establish a distinct category for this segment. This strong claim about the permanence of this effect was found to be supported by work carried out by Matthews (1997), as no amount of instruction or training brought about any improvement on contrasts involving such segments. In other words, L2 learners cannot acquire novel phonological features. The weak version, on the other hand, claimed that there were further distinctions involved for those segments that require an absent phonological feature: it was argued that those phonological features that serve as terminal nodes in the feature geometry tree (such as [vibrant]) were more easily acquired than those that serve as articulator nodes (such as [PHARYNGEAL]). This model of L1 interference, then, differs from the strong version with respect to two claims: first, it predicts that among those new segments identified as being difficult to acquire, we can identify additional degrees of difficulty; second, it predicts that novel phonological features can be acquired in L2 acquisition, where success is a function of the segment's featural degree of difficulty. It is of interest, then, to evaluate the claims of both these models in order to assess which one is more appropriate to the patterns that can be observed among L2 learners.

Both of these models of L1 interference were formulated on the basis of data gathered from L2 learners at various stages of acquisition. We felt that an extrapolation of their claims could be made in order to generate testable predictions about the initial state of L2 acquisition, before exposure to the second language. These were formulated as the first two questions listed above: for the first question, both versions predict that monolingual English speakers should differentiate between their native rhotic /ɹ/ and the new rhotics /ʀ/, found in French, and /r/, found in Spanish, for which the necessary phonological features are inactive in the grammar. The two versions diverge with respect to the second question, as only the weak version predicts that monolingual English speakers should maintain an additional distinction between /ʀ/ and /r/ as these segments differ with respect to the type of feature required for their representation that is missing from the L1 grammar.

The results detailed in Chapter 2 indicate that for monolingual English speakers, all three rhotics examined (/ɹ/, /ʀ/, and /r/) are assigned to a single category, suggesting that these individuals construct a single segmental structure for the phonological representation of all three segments: all three segments elicited nearly identical N100 responses, any differences among these were not significant. There was evidence to suggest, however, that some distinction was being maintained within this category between the native rhotic /ɹ/ and the two non-native trills, /ʀ/ and /r/: the amplitude of the

N1 – P2 complex was significantly greater in response to /ɹ/ than in response to either /ʀ/ or /r/. Furthermore, no significant difference in N1 – P2 complex amplitude was found between the responses elicited by the non-native segments. This data is then interpreted as empirical support for the strong L1 feature hypothesis, in that while subjects do differentiate between native and non-native segments in the initial state, there do not appear to be any initial distinctions between new segments that are made on the basis of the type of absent phonological features required for appropriate representations.

The third question guided the investigation described in Chapter 3 and sought to evaluate the two versions under consideration with respect to rate of acquisition of predicted problematic L2 segments in L2 learners whose initial L2 exposure had occurred in adolescence or adulthood. We compared the neural responses elicited by rhotic segments in L2 French learners and L2 Spanish learners at varying levels of proficiency. We examined beginner speakers with eight months of classroom L2 exposure, advanced speakers with a minimum 3 years of classroom exposure, and very advanced speakers with several years' classroom exposure and extended periods of adult L2 immersion. The results suggest that these learners were unable to establish a new category for the L2 rhotic segment, regardless of the type of feature that is absent in the L1 grammar or the length of exposure to the L2. For all groups, the N100 peak elicited in response to both native and non-native rhotics were remarkably similar. Much like the finding among our monolingual English speakers, the L2 learners appear to be constructing a single segmental representation for all rhotic sounds. Changes were observed in the amplitude of the N1 – P2 complex elicited in response to the non-native segment, with amplitude increasing with L2 proficiency. That is, increased L2 proficiency led to an increase in subjects' ability to make within-category distinctions between native and non-native rhotic segments. To couch this finding in Brown's (1997) model of the mapping of acoustic signals to phonemic categories, this could be interpreted as learners' non-native phonetic categories becoming increasingly robust and well-defined (within the native phonemic category). That is, while [ʀ] is assigned to the phonemic category /ɹ/ regardless of L2 proficiency, the phonetic category becomes increasingly well-defined such that at some level, the learner is aware of the differences between [ɹ] and [ʀ].

It should be noted, however, that our interpretation of the weak L1 feature hypothesis predicted that we should observe changes in the neural responses elicited by /r/ before those elicited by /ʀ/, owing to the fact that while /r/ requires a terminal node absent from English, /ʀ/ crucially requires an articulator node absent from English. The data suggest a pattern that is the reverse of the one predicted: the N1 – P2 complex amplitude was shown to increase at an earlier stage of acquisition for /ʀ/; beginner L2 Spanish learners showed an N1 – P2 complex for /r/ that was smaller than that obtained from beginner L2 French learners for /ʀ/. It was pointed out, however, that this pattern may be due to differences in the phonemic inventories of French and Spanish, the result being that while L2 French learners were learning to make a two-way within-category distinction among rhotics, L2 Spanish learners were learning to make a more complex three-way within-category distinction. Our observed pattern, then, is not to be considered a counterexample to the weak L1 feature hypothesis.

In spite of our best efforts in recruiting participants, a number of individuals volunteered for the study whose experience with language did not fit the desired profiles. The presence of these subjects in our data sample allowed us to formulate our fourth question. This investigation demonstrated that it is possible to acquire novel phonological features in a second language, as indicated by distinct N100 response patterns; there are, however, conditions that must be met in order for this to be achieved. The general pattern that emerges in the data indicates that new phonological features may be triggered in a second language if acquisition begins in early childhood, rather than adolescence or adulthood. That is, early childhood L2 exposure is a necessary condition for the acquisition of novel phonological features. Furthermore, there appears to be a minimum threshold level with respect to the duration of exposure, such that childhood L2 exposure that does not meet this minimum requirement will not result in acquisition of non-native phonological features. Adult third language (L3) exposure to new segments that are also found in a previously acquired L2 does not appear to have any effect on the acquisition of the L3 segment.

Generally speaking, then, the results of the present investigation best support the strong L1 feature hypothesis: triggering of phonological features that are absent from the L1 seems possible only with sufficient childhood L2 exposure, and the predictions made by the weak version with respect to degree of difficulty in acquisition are not borne out in the electrophysiological data. While adult L2 learners' neural responses to novel segments do show a change over time that has been argued to be associated with improved perception (Tremblay et al. 2001), crucially these changes do not result in the creation of a new category for the L2 segment. If, however, initial L2 exposure occurs in early childhood (i.e., before the age of 9), and satisfies some minimum threshold criteria for length of L2 exposure, then the required phonological features will be acquired and appropriate representations for L2 segments can be built.

Returning to our four research questions, the present electrophysiological results seem to indicate that:

5) Monolingual English speakers assign both native and non-native rhotic segments to a single category; they do, however, maintain some distinction within this category between those segments which are native and those which are not.

6) Monolingual English speakers do not make distinctions among novel L2 segments on the basis of phonological features that are absent in the L1 grammar.

7) Adult language learners are unable to trigger new structure and thus cannot construct appropriate phonological representations for novel L2 segments.

8) Previous exposure to a non-native segment in a second language affects the perception and processing of this segment in a third language only if a sufficient amount of childhood L2 exposure has enabled the learner to acquire the necessary features, thereby allowing him to construct an appropriate phonological representation.

More broadly, the present findings suggest that adult L2 learners cannot acquire new phonological features in a second language. Further, we find no evidence that notions of degree of difficulty should be encoded by making reference to the feature content of a given segment.

This, in turn, raises the question of why L2 learners cannot acquire novel structure. This is frequently attributed to a loss (Birdsong 1999, Long 1990) or decline (Yamada 1995) in neural plasticity that occurs progressively as learners age. This is the idea that neural structures are somehow "set" to accommodate the L1 and cannot be "re-set" (loss) or are increasingly difficult to "re-set" (decline) as we age; one promising proposal to account for this claims that as neuron axons develop their myelin sheath, neural pathways for various types of information become firmly established in the brain (see Long 1990). Our chosen research methodology allows us to comment on this issue, as ERPs directly measure electrical activity resulting from neural activity in the brain. The finding that the N1 – P2 complex elicited by novel L2 segments increases in amplitude with increasing L2 proficiency reveals that we cannot maintain an approach that incorporates a notion of loss of neural plasticity. We may, however, be able to support the notion of decline of neural plasticity, as we noted that although the amplitude of the N1 – P2 complex elicited by L2 segments increases with L2 proficiency, it is never of equal or greater amplitude than that elicited by L1 segments. Additional work with expanded data pools is required in order to determine whether this notion will find robust empirical support, but at present the data discussed here leave us with an intriguing suggestion for further examination.

5.2. Future directions

As discussed in previous chapters, the time-consuming nature of the experimental work described in this thesis prevented us from recruiting large numbers of participants for study, and it may therefore be more appropriate to view the current research as a pilot study. Nonetheless, the research achieves its purpose in that the experimental results present a clear picture of the neural processing of novel L2 segments, from the initial state before L2 acquisition begins through to a final state of near-native L2 proficiency. Due to the relatively small number of participants were involved in this research, it is of considerable interest to repeat the testing procedure described here with a larger group of participants; in a number of the groups examined, only one participant provided the data for analysis. Doing so would allow us to determine how robust the observed pattern is, as well as provide valuable insight as to how we should interpret any data that does not fit well with the emerging pattern.

As with all scientific enquiry, as potential answers to existing questions are found, new questions are raised. The current results leave us with the question of how degrees of difficulty among L2 segments should be explained. While our findings suggest that these cannot be encoded into phonological representations through features, a potential confound with our chosen segments was identified: it was pointed out in Chapter 3 that native speakers of English must acquire a two-way within-category distinction for rhotic segments when acquiring French and a three-way distinction for rhotic segments in Spanish. We would therefore want to investigate this further with a larger number of participants and different L2 segments before completely abandoning the possibility that phonological features account for observed degrees of difficulty. Additionally, the finding that not all individuals with childhood exposure to a second language succeed in acquiring novel L2 phonological features raises questions about how much exposure is required for successful acquisition. Additionally, the finding raises questions concerning whether all individuals with childhood L2 exposure acquire L2 features, which later become unavailable due to lack of use, or are the features never acquired if L2 exposure ends within a certain time frame? Questions such as these prompt us to carry out tests such as those described in this thesis with children.

In sum, the experimental work described in this thesis explored a new and innovative technique for assessing L2 phonological acquisition. Examination of the data gathered suggests that the role of age of initial L2 exposure is absolute: adult L2 learners are not able to trigger novel phonological structures. Their knowledge of L2 segments does, however, undergo improvements with length of exposure and increasing proficiency, but these improvements do not result in a native-like representation for L2 segments. Furthermore, our results suggest that if initial L2 exposure occurs in early childhood, the length of L2 exposure plays a role in that some minimum threshold criteria for exposure must be met, otherwise no evidence of distinct representations for new L2 segments can be found when these individuals are tested later in life. With respect to the issue of whether observed degrees of difficulty in acquisition of L2 segments can be encoded using phonological features, our results indicate that this is not the case. Future research is needed, however, in order to identify plausible alternatives.

References

Archibald, J. 2000. Parsing procedures and the question of full access in L2 phonology. In J. Leather and A. James (eds.). *New Sounds 2000.* Klagenfurt: University of Klagenfurt Press.

Baković, E. J. 1994. Strong onsets and Spanish fortition. *MIT Working Papers in Linguistics* 23, 21 – 39.

Best, C. T. 1995. A direct realist view of cross-language speech perception. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research.* Baltimore: York.

Birdsong, D. 1999. Introduction: whys and why nots of the Critical Period Hypothesis in second language acquisition. In D. Birdsong (ed.). *Second Language Acquisition and the CPH.* Mahwah: Erlbaum.

Blevins, J. 1994. A place for lateral in the feature geometry. *Journal of Linguistics* 30, 301 – 348.

Bley-Vroman, R. 1989. What is the logical problem of foreign language learning? In S. Gass and J. Schachter (eds.). *Linguistic Perspectives on Second Language Acquisition.* Cambridge: Cambridge University Press.

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

Bradley, T. G. 2001. A typology of rhotic duration contrast and neutralization. In M.-J. Kim and U. Strauss (eds.). *Proceedings of the North East Linguistic Society 31*. Amherst: GLSA Publications.

Brown, C. A. 2000. The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (ed.). *Second Language Acquisition and Linguistic Theory*. Malden: Blackwell.

Brown, C. A. 1997. Acquisition of segmental structure: consequences for speech perception and second language acquisition. PhD dissertation, McGill University.

Brown, C. M., and P. Hagoort. 1999. The cognitive neuroscience of language: challenges and future directions. In C. M. Brown and P. Hagoort (eds.). *The Neurocognition of Language*. New York: Oxford.

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

Campbell, N. A., L. G. Mitchell, and J. B. Reece. 1997. *Biology: Concepts and Connections*. Menlo Park: Benjamin/Cummings.

Carroll, S. 2001. *Input and Evidence: the Raw Material of Second Language Acquisition*. Amsterdam: John Benjamins.

Chomsky, N. 1965. *Cartesian Linguistics*. New York: Harper and Row.

Colantoni, L. M. 2001. Mergers, chain shifts, and dissimilatory processes: palatals and rhotics in Argentine Spanish. PhD dissertation, University of Minnesota.

Coles, M. G. H., and M. D. Rugg. 1995. Event-related brain potentials: an introduction. In M. D. Rugg and M. G. H. Coles (eds.). *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*. New York: Oxford.

D'Arcy, R. C. N., J. F. Connolly, and S. F. Crocker. 2000. Latency shifts in the N2b component track phonological deviations in spoken words. *Clinical Neurophysiology* 111, 40 – 44.

Dalbor, J. B. 1997. *Spanish Pronunciation: Theory and Practice*. Orlando: Harcourt.

Dehaene-Lambertz, G. 1997. Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport* 8, 919 – 924.

Dehaene-Lambertz, G., E. Dupoux, and A. Gout. 2000. Electrophysiological correlates of phonological processing: a cross-linguistic study. *Journal of Cognitive Neuroscience* 12 (4), 635 – 647.

Eckman, F. R. 1987. Markedness and the Contrastive Analysis Hypothesis. In G. Ioup and S. H. Weinberger (eds.). *Interlanguage Phonology: the Acquisition of a Second Language Sound System*. Cambridge: Newbury House.

Flege, J. E. 1995. Second language speech learning: theory, findings, and problems. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York.

Flege, J. E. 1987. The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics* 15, 47 – 65.

Halle, M., B. Vaux, and A. Wolfe. 2000. On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31, 387 – 444.

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

Harris, J. 2002. Flaps, trills, and syllable structure in Spanish. *MIT Working Papers in Linguistics* 42, 81 – 108.

Hohne, E. A., and P. W. Jusczyk. 1994. Two-month-old infants' sensitivity to allophonic differences. *Perception and Psychophysics* 56 (6), 613 – 623.

Howe, D. 2003. Segmental phonology. Ms., University of Calgary.

Jusczyk, P. W. 1997. *The Discovery of Spoken Language*. Cambridge: MIT Press.

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

Karlsson, F. 1999. *Finnish: an Essential Grammar*. New York: Routledge.

Killingley, S.-Y. 1993. *Cantonese*. München: Lincom Europa.

Kopczyński, Andrzej. 1977. *Polish and American English Consonant Phonemes*. Warsaw: Państwowe Wydavnictwo Naukowe.

Kraus, N., T. McGee, T. D. Carrell, and A. Sharma. 1995. Neurophysiologic bases of speech discrimination. *Ear and Hearing* 16, 19 – 37.

Kuhl, P. K., and P. Iverson. 1995. Linguistic experience and the "perceptual magnet effect." In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York.

Kutas, M., K. D. Federmeier, and M. I. Sereno. 1999. Current approaches to mapping language in electromagnetic space. In C. M. Brown and P. Hagoort (eds.). *The Neurocognition of Language*. New York: Oxford.

LaCharité, D., and P. Prévost. 1999. The role of L1 and of teaching in the acquisition of English sounds by francophones. In A. Greenhill, H. Littlefield, and C. Tano (eds.). *Proceedings of BUCLD 23*. Somerville: Cascadilla.

Lado, R. 1957. *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press.

Long, M. H. 1990. Maturational constraints on language development. *Studies in Second Language Acquisition* 12, 251 – 285.

Mah, J. 2002. Knowledge and performance: an examination of the role of explicit linguistic knowledge in L2 phonological acquisition. *Calgary Working Papers in Linguistics* 24, 76 – 128.

Mah, J. 2001. Turning Japanese: Optimality Theory and loanword phonology. Ms., University of Calgary.

Maiste, A. C., A. S. Wiens, M. J. Hunt, M. Scherg, and T. W. Picton. 1995. Event-related potentials and the categorical perception of speech sounds. *Ear and Hearing* 16 (1), 68 – 90.

Major, R. C. 2001. *Foreign Accent: the Ontogeny and Phylogeny of Second Language Phonology*. Mahwah: Erlbaum.

Major, R. C., and E. Kim. 1996. The Similarity Differential Rate Hypothesis. *Language Learning* 46 (3), 465 – 496.

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.* Klagenfurt: University of Klagenfurt Press.

Messora, N. 1983. *Mastering Italian.* London: MacMillan.

Newport, E. 1990. Maturational constraints on language learning. *Cognitive Science* 14, 11 – 28.

Pallier, C., L. Bosch, and N. Sebastian-Gallés. 1997. A limit on behavioural plasticity in speech perception. *Cognition* 64 (3), B9 – B17.

Phillips, C. 2001. Levels of representation in the electrophysiology of speech perception. *Cognitive Science* 25 (5), 711 – 731.

Phillips, C., T. Pellathy, A. Marantz, E. Yellin, K. Wexler, D. Poeppel, M. McGinnis, and T. Roberts. 2000. Auditory cortex accesses phonological categories: an MEG mismatch study. *Journal of Cognitive Neuroscience* 12 (6), 1038 – 1055.

Poeppel, D., and A. Marantz. 2000. Cognitive neuroscience of speech processing. In A. Marantz, Y. Miyashita, and W. O'Neil (eds.). *Image, Language, Brain: Papers from the First Mind Articulation Project Symposium.* Cambridge: MIT Press.

Polka, L., and J. F. Werker. 1994. Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance* 20 (2), 421 – 435.

Prowe, G. 2001. *The Oxford German Dictionary and Grammar.* New York: Oxford.

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*. Mahwah: Erlbaum.

Rochet, B. L. 1995. Perception and production of second-language speech sounds by adults. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York.

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

Rugg, M. D. 1999. Functional neuroimaging in cognitive neuroscience. In C. M. Brown and P. Hagoort (eds.). *The Neurocognition of Language*. New York: Oxford.

Rugg, M. D., and M. G. H. Coles. 1995. The ERP and cognitive psychology: conceptual issues. In M. D. Rugg and M. G. H. Coles (eds.). *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*. New York: Oxford.

Sams, M., R. Aulanko, O. Aaltonen, and R. Näätänen. 1990. Event-related potentials to infrequent changes in synthesized phonetic stimuli. *Journal of Cognitive Neuroscience* 2 (4), 344 – 357.

Sharma, A., and M. F. Dorman. 1999. Cortical auditory evoked potential correlates of categorical perception of voice-onset time. *Journal of the Acoustical Society of America* 106 (2), 1078 – 1083.

Sharma, A., N. Kraus, T. McGee, T. Carrell, and T. Nicol. 1993. Acoustic versus phonetic representation of speech as reflected by the mismatch negativity event-related potential. *Electroencephalography and Clinical Neurophysiology* 88, 64 – 71.

Shirai, S. 2001. Gemination in loans from English to Japanese. M. A. thesis, University of Washington.

Steriade, D. 1993. Closure, release, and nasal contours. *Phonetics and Phonology* 5, 401 – 470.

Tremblay, K., N. Kraus, T. McGee, C. Ponton, and B. Otis. 2001. Central auditory plasticity: changes in the N1 – P2 complex after speech-sound training. *Ear and Hearing* 22 (2), 79 – 90.

van Turennout, M., P. Hagoort, and C. M. Brown. 1998. Brain activity during speaking: from syntax to phonology in 40 milliseconds. *Science* 280 (5363), 572 – 574.

van Turennout, M., P. Hagoort, and C. M. Brown. 1997. Electrophysiological evidence on the time course of semantic and phonological processes in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23 (4), 787 – 806.

Walker, D. C. 2001. *French Sound Structure*. Calgary: University of Calgary Press.

Walsh Dickey, L. 1997. The phonology of liquids. PhD dissertation, University of Massachusetts Amherst.

Werker, J. F. 1995. Age-related changes in cross-language speech perception: standing at the crossroads. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York.

Werker, J. F., and R. C. Tees. 1984. Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75 (6), 1866 – 1878.

White, L. 2003. *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.

White, L. 1989. *Universal Grammar and Second Language Acquisition*. Amsterdam: John Benjamins.

Yamada, R. A. 1995. Age and acquisition of second language speech sounds: perception of American English /ɹ/ and /l/ by native speakers of Japanese. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York.

Appendix A: Ethics Approval

UNIVERSITY OF CALGARY

MO SHÙILE TOGAM SÙAS

CERTIFICATION OF INSTITUTIONAL ETHICS REVIEW

This is to certify that the Conjoint Faculties Research Ethics Board at the University of Calgary has examined the following research proposal and found the proposed research involving human subjects to be in accordance with University of Calgary Guidelines and the Tri-Council Policy Statement on *Ethical Conduct in Research Using Human Subjects*". This form and accompanying letter constitute the Certification of Institutional Ethics Review.

Applicant(s): Jennifer Mah

Department/Faculty: Department of Linguistics

Project Title: Acquisition of Segmental Structure

Sponsor (if applicable):

Restrictions:

This Certification is subject to the following conditions:

1. Approval is granted only for the project and purposes described in the application.
2. Any modifications to the authorized protocol must be submitted to the Chair, Conjoint Faculties Research Ethics Board for approval.
3. A progress report must be submitted 12 months from the date of this Certification, and should provide the expected completion date for the project.
4. Written notification must be sent to the Board when the project is complete or terminated.

Chair
Conjoint Faculties Research Ethics Board

Date: 21 February 2003

Distribution: (1) Applicant, (2) Supervisor (if applicable), (3) Chair, Department/Faculty Research Ethics Committee, (4) Sponsor, (5) Conjoint Faculties Research Ethics Board (6) Research Services

09/00