The Perception and Production of Interdental Fricatives in Second Language Acquisition

Kathleen J. Brannen

Department of Linguistics

McGill University, Montréal

June 2011

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

This thesis is dedicated to my mother, Mary Cécile Brannen

and to the memory of my father, Edward Vernon Brannen

TABLE OF CONTENTS

ABSTRACT

This dissertation investigates the differential substitution of interdental fricatives ([θ,ð]) by learners of English as a second language. Differential substitution (or transfer) occurs when learners whose native language does not include the "th" sounds, substitute different segments in their place, depending on the particular L1: [t,d], [s,z], or [f,v]. Throughout the past 50 years, various proposals have been forwarded to explain this phenomenon. The majority of these approaches have focussed on structural differences within the contrastive phonemic systems of various languages. This thesis examines two languages and two dialects of the same language: Japanese, Russian, European French, and Québec French. Japanese and European French are known to substitute [s,z] in place of [θ,ð], while for Québec French and Russian, [t,d] are reported. Since European and Québec French arguably have the same phonemic inventory of obstruents, this thesis explores the function of non-contrastive phonetic information in interdental substitution, in both perception and production. It is hypothesized that perception underlies production, such that those errors observed in production will be the sounds that are apt to be perceptually associated with the target. Furthermore, it is considered that non-contrastive features play a determining role in segmental transfer. In particular, the feature STRIDENT is hypothesized to be key in the choice of interdental substitute.

To account for how second language learners perceptually map target sounds to their internal representations, the Auditory Distance Model is developed. This model is based upon the following hypotheses. In the initial state of acquisition, learners potentially have access to all phonetic features provided by Universal Grammar. However, availability

of this universal set of features is limited by the language-specific *phonetic* inventory, such that the absence of features or particular combinations of features in the L1 grammar forces the L2 learner to choose from among the phonetically closest L1 sounds. The selection process is implemented via the Auditory Distance Algorithm which compares the target intake features with those encoded on L1 segments. The Algorithm evaluates whether the intake and L1 features constitute matches or mis-matches and additionally calculates their relative weight. A feature's weight can be augmented if it stands in an enhancement relation with another feature. Enhanced features are preferred, but not mandatory, in linguistic systems. Thus cross-linguistic phonetic variation and the resulting diversity in feature weight is what determines differential substitution in perception and hence in production.

These hypotheses are empirically verified in five studies. The first two, the AXB-1 and AXB-2 perception tasks, were designed to tap phonetic and phonemic processing in separate conditions to demonstrate that the observed patterns of differential substitution emerge in phonetic, but not phonemic processing. The third perception experiment, Picture Identification, examines phonemic processing. The final studies analyze production. The results of one, a Word Production task, are compared with the perception findings. The other involves a Spectrographic Analysis of the L1 coronal fricative [s] to determine the degree to which the feature STRIDENT is acoustically manifested for each of the languages.

The results from these studies largely support the hypotheses outlined above. To account for discrepancies between predictions and results, the role of visual information in lexical representations and the possibility of task-induced bias are discussed.

ABRÉGÉ

Cette dissertation explore la substitution différentielle des fricatives interdentales par les apprenants de l'anglais langue seconde. La substitution différentielle se produit quand les apprenants dont la langue maternelle (L1) ne contient pas ces sons, substituent ces phonèmes par un autre. Dépendamment de la L1, la substitution se fait avec [t,d], [s,z] ou [f,v]. Au cours des années, plusieurs propositions ont été émises pour expliquer ce phénomène. Ces hypothèses reposent principalement sur les différences structurelles des systèmes phonémiques de diverses langues. Cette thèse examine la substitution de [θ,ð] par des locuteurs de différentes langues ou dialectes, à savoir: le japonais, le russe, le français européen, et le français québécois. Le japonais et le français européen sont caractérisés par les substituts [s,z], tandis que pour le français québécois et le russe, par [t,d]. Puisque le français européen et le français québécois ont le même inventaire contrastif de segments obstruants, cette dissertation explore la contribution de l'information non-contrastive à la substitution des interdentales, de la perspective de la perception et de la production. Il est proposé que la perception sous-tend la production de telle manière que les erreurs observées dans la production correspondent aux sons substitués au niveau de la perception. De plus, il est considéré que les traits non-contrastifs jouent un rôle déterminant dans le transfert. Spécifiquement, il est proposé que le trait STRIDENT est central au choix de substituts pour les interdentales.

Pour expliquer comment les apprenants des langues secondes projettent les sons cibles sur leur représentation interne dans la perception, le modèle de distance auditoire est développé. Ce modèle est basé sur les hypothèses suivantes: à la phase initiale de l'acquisition, les apprenants ont potentiellement accès à tous les traits phonétiques fournis par la grammaire

universelle. Cependant, cet ensemble de traits est limité à l'inventaire phonétique de la langue spécifique. Ainsi, l'absence de traits ou de combinaisons spécifiques de traits dans la grammaire de la langue maternelle force l'apprenant à choisir parmi les sons de sa L1 ceux qui sont phonétiquement les plus proches. Ceci est accompli via un algorithme qui évalue la correspondance entre les traits cibles et ceux de la L1 ainsi que leur poids relatifs. Le poids d'un trait peut être augmenté s'il se trouve dans une relation de rehaussement avec un autre trait. Les traits rehaussés sont préférables, mais pas obligatoires, dans les systèmes linguistiques. Alors la variation phonétique et inter-linguistique de même que la diversité dans le poids des traits qui résulte, sont les facteurs qui déterminent la substitution différentielle dans la perception et de là, dans la production.

Ces hypothèses sont vérifiées empiriquement dans cinq études. Les deux premières, AXB-1 et AXB-2, ont été construites pour capter le traitement phonétique et phonémique dans des conditions distinctes pour démontrer que les réflexes de la substitution différentielle émergent dans le traitement phonétique et non phonémique. La troisième, l'identification d'image, examine le traitement phonémique. Les deux dernières études analysent la production. Les résultats d'une de ces dernières, une tâche sur la production de mots, sont comparés avec les résultats des analyses de la perception. L'autre consiste en une analyse spectrographique de la fricative coronale [s] de la L1 pour déterminer à quel degré le trait STRIDENT est manifesté acoustiquement pour chacune des langues.

En général, les résultats de ces études appuient les hypothèses décrites ci-haut. Pour rendre compte des quelques divergences entre les prédictions et les résultats, la discussion aborde le rôle de l'information visuelle dans les représentations lexicales et la possibilité de la distorsion induite par la tâche.

ACKNOWLEDGEMENTS

Here I am, at the end of my long journey. I have had the pleasure to meet so many wonderful people along the way.

First and foremost, I owe so much gratitude to my thesis advisor, Heather Goad. When I first came to McGill, her infectious love for phonology was what instigated me in this pursuit. Throughout my career at McGill, she has gone way beyond what is required of a professor and thesis supervisor. I have benefitted enormously from our long discussions and from her critical reading and fastidious corrections of many draft copies of this thesis. Her keen insight and assiduity have been most appreciated. Without her encouragement, I would never have gotten to this point. Heather is a champion of her students: she has a deep respect for them and faith in their abilities. She inspires others through her passion, generosity, and ethics. I consider myself very fortunate to have her as a mentor and a friend. Thank you Heather!

I would also like to give special thanks to Lydia White for her esteemed involvement in my work and continued support. I will always aspire to Lydia's level of efficiency and ability to "cut to the chase", but sadly realize that this is probably an unattainable goal.

My other professors at McGill have made my overextended stay there stimulating and rewarding. In particular, I especially appreciated the phonology seminars with Glyne Piggott. He instantiates my idea of what it is to be a true scholar. I am also grateful to Lisa Travis for her encouragement, Mark Baker, Charles Boberg, Meghan Clayards, Brendan Gillon, Eva Kehayia, Michel Paradis, and Linda Polka.

Thanks are also extended to the administrative staff at the McGill Department of Linguistics: Lise Vinette, Andria de Luca, Connie Di Guiseppe, and Lauren Mak for their indispensable work.

My career began at the University of Ottawa. It started with a French class I took with Julien Couture. His class was my springboard into the study of language. I never had the chance to thank him.

I credit my professors at UofO for kindling my excitement for linguistics. In particular, I would like express my gratitude to my MA Supervisor, Gerry Neufeld.

The time I spent at McGill will never be forgotten, in large part because of the company of my fellow “McGillians”: Walcir Cardoso, Lotus Goldberg, Ayse Gürel, Mordecai Lieberman, Jennifer Mah, Evan Mellander, Yvan Rose, Jeff Steele, Elena Valenzuela, Mikael Vinka, and in fond memory, Lara Riente.

I would also like to thank my friends and colleagues at UQAM: Josée Blanchet, Imad Buali, Jaime Demperio (thanks for dragging me, not so unwillingly, out on the town!), Peter Enns, Francisco Goire, Paul John, Émile Khordoc, Kerwin Myler, Carey Nelson, Jane Nicholls (thanks for our wonderful conversations about anything and everything), Marie Ploquin, Emily Rosales (I did it Emily!), Gerald Rosenau, and Beau Zuercher.

And, in a class of their own, there are the "McGillians/UQAMiens": Ernie Bauer, Martyna Kozlowska, Heather Newell, and Tobin Skinner.

A thesis that involves experimental work requires the aid of many people. I am grateful to all those who contributed to this work:

- Yvan Rose for his generous help in the experimental design (et pour la chance de pratiquer mon français!);
- Charles Boberg, Heather Goad, and Gerald Rosenau for help with task development;

- Charles Boberg for technical support;

- Erika Lawrance, Annie Lord, and Inna Tolskaya for coding and verification;

- Lucy Cumyn and Mehga Sundara for help with statistics;

- Aysa Pereltsvaig and Yuhko Kayama for translation;

- Marie Ploquin for proofreading;

- Ruslan Buyalo, Tomokazu Takehisa, Naoko Tamioka, Tohru Uchiumi, and Mari Umeda for help in subject recruitment.

- all the participants in the studies, for their interest and patience.

My PhD studies would not have been possible without financial support from the following sources: FCAR Doctoral Fellowship, McGill FGSR Fellowship, FCAR to Heather Goad, FCAR to Lydia White, SSCHRC to Eva Kehayia, and SSHRC to Lydia White.

Commuting to Montréal was not easy, so I extend special thanks to Walcir Cardoso, Heather Goad, and Martyna Kozlowska for their hospitality.

I also want to show my appreciation to all my students, especially my English pronunciation students for keeping me grounded.

Last, but far from least, my biggest thanks goes to the most important men in my life: Daniel, Julien, and Xavier Habets for, well....EVERYTHING!!

CHAPTER 1: Introduction

One predominant characteristic of second language (L2) speech is the presence of foreign accent. A major goal of researchers working in the domain of second language sound systems is to account for accented speech. The factors influencing second language production are many, and they interact in complex ways.

Variables such as age of acquisition, aptitude, motivation, quantity and quality of L2 input, and social attitudes influence second language acquisition, and in particular the L2 sound system (e.g. Lenneberg 1967, Gardner & Lambert 1972, Lamendella 1977, Schumann et al. 1978, Krashen 1985, Scovel 1988, Long 1990, Carroll 1993). For nearly a century now, researchers have invoked one or more of these factors in an attempt to explain the phenomenon of foreign accent. Although the aforementioned variables constitute possible contributions to non-native sound productions, the one major underlying factor that we have not mentioned is the pre-existence of the first language.

It is commonly believed that one's second language is strongly influenced by one's native language (for phonetics and phonology see e.g. Weinreich 1953, Trubetzkoy 1939/1969, Eckman 1977, Robinett & Schachter 1983, Wode 1983, Ioup & Weinberger 1987, Flege 1995, James 1994). This influence of native language (L1) structures and features on the second language (L2) has variably been termed *transfer, interference*, or *substitution.* Transfer can involve all components of the grammar. This dissertation is concerned with the phonetic and phonological components.

Accent is frequently attributable to the substitution of L1 sounds in the place of L2 sounds which have no native counterpart. An overview of the substitution facts reveals that the segment that is transferred often varies depending on the native language of the speaker. The same target sound may be realized differently depending on the L1. This phenomenon has been termed "differential substitution" (Weinberger 1988:118).

This dissertation addresses cross-linguistic differential substitution as it applies to the L2 English target interdental fricatives, [θ] (theta) and [ð] (eth).[1] The cross-linguistic perspective is investigated by examining European French (EF), Québec French (QF), Japanese (JA), and Russian (RU). These languages all have the sounds /t,d,s,z/ in their phonemic inventories, but lack the sounds /θ,ð/. European French and Japanese speakers are known to substitute [s,z] in place of [θ,ð] (e.g. Berger 1951, Hancin-Bhatt 1994a,b) while Québec French and Russian speakers are reported to use [t,d] (e.g. Gatbonton 1978, Michaels 1974, Weinberger 1988).

Differential substitution presents a challenge for phonological theory. Two languages can have very similar segmental inventories, yet the speakers of these languages will choose different substitutes for the target interdental fricatives. And of particular importance in this dissertation, sometimes speakers from two dialects of the same language, with the same segmental inventories, demonstrate differential substitution. In the absence of any clear evidence pointing to these languages or dialects having different underlying representations, it seems that we must look beyond the contrastive features of phonemic inventories for an explanation.

In the recent literature, there has been much discussion as to the respective role of phonetics versus phonology involved in the learning of a second language sound system (e.g. Hancin-Bhatt 1994a,b; Weinberger 1988; Brown 1997,1998; Flege 1995; Flege 2003; Johnson 2003;

[1] Henceforth, the term *differential substitution* will exclusively refer to differential substitution of the interdental fricatives.

Lombardi 2003; Muñoz-Sánchez 2003; Escudero 2005; Grenon 2010). Concerning the interdental fricatives of English, some previous work on sound transfer has argued that its explanation is to be found at the phonemic level (e.g. Weinberger 1988; Hancin-Bhatt 1994a,b; Brown 1997,1998; Lombardi 2003).

In his 1988 dissertation, Weinberger adopts the theory of Radical Underspecification (e.g. Kiparsky 1985, Archangeli 1988, Archangeli & Pulleyblank 1989), whereby phonemic inventories are minimally specified for phonological features. For a given inventory of sounds, Radical Underspecification allows for a variety of representations. Weinberger argues that the differential substitution data constitute evidence for the way sounds are represented in a given language as detailed in its underlying feature matrix and redundancy rules.

Weinberger exemplifies this through an investigation of how Japanese and Russian speakers treat English /θ/. He states that Japanese learners of English substitute /s/ for /θ/ and that Russian learners substitute /t/ for /θ/ despite them both having /s/ and /t/ in their phonemic inventories. He considers that in the underlying Japanese inventory, /s/ is the default obstruent, being the minimally specified segment marked as [-sonorant, -voice], while in Russian, it is /t/ that is the default obstruent, with these same minimal specifications. Other features are unmarked and filled in later in the derivation by redundancy rules. In the case of Japanese, the redundancy rules fill in the unmarked features as [+anterior], [+coronal], and [+continuant]; whereas, in Russian, redundancy rules fill them in as [+anterior], [+coronal], and [-continuant]. He proposes that interdental substitution involves inserting the default obstruent in place of the target. Weinberger states that neither Japanese nor Russian offer internal phonological rules to justify the choice of default segment. This leads him to appeal to external cues: "In the absence of native language evidence, L2 data serves as a portal to the otherwise obscured native language grammar" (Weinberger 1988:205). This claim is problematic in terms of the issue of learnability. It is unclear how the child learning her L1 could arrive at the correct underlying specifications without native language internal evidence.

A few years after Weinberger's work, another dissertation investigated the differential substitution phenomena through the framework of Radical Underspecification. Hancin-Bhatt (1994a,b) supplements the Radical Underspecification analysis with the notion of "functional load". The functional load of a given feature value is equivalent to the number of times it is specified in the inventory divided by the total number of phonemes. The feature value with the higher functional load will be underspecified in the phonemic inventory; thus, this feature will be later inserted via a redundancy rule, resulting in either a continuant or a stop substitute. She investigates the functional load of the binary distinctive feature [±continuant] in the inventories of German, Hindi, Japanese, and Turkish to predict which of these languages will substitute [s,z] for the interdental fricatives, and which will substitute [t,d].

Based on her calculations of functional load, she predicts that German speakers of English will substitute [s,z] in place of [θ,ð] and that Hindi, Japanese, and Turkish speakers will substitute [t,d]. Hancin-Bhatt's predictions were borne out for the Turkish listeners and largely for the German group. However, the results of her perception tests for Hindi and Japanese did not support her predictions. Unfortunately as well, Hancin-Bhatt's model cannot capture the difference between two dialects of the same language which share the same underlying consonant inventory. There are also some theoretical problems with her use of Radical Underspecification theory. This theory requires that language-internal phonological processes motivate the configuration of an underlying inventory; and in the absence of such processes, it should be markedness which dictates the inventory specifications (Archangeli 1988). In Hancin-Bhatt's thesis, these requirements of Radical Underspecification are not met as she provides no language-internal evidence, nor does she rely on markedness. However, Radical Underspecification is but one theory of underspecification in the phonological arena.

Another theory that states that certain features are absent from the underlying inventory is Minimally Contrastive Underspecification (Avery & Rice 1989). According to this theory, the only features specified are those needed to distinguish one segment from all other segments in the inventory. Working within this framework and that of Feature Geometry, Brown (1997, 1998) investigates how Japanese speakers of L2 English perceive several obstruent contrasts, including /s-θ/. Brown hypothesizes that second language learners cannot acquire new phonological features, that is, features that do not function contrastively in their native language. However, she claims that L2 learners are able to recombine features that are contrastive in their native language in order to form representations for new segments.

Brown’s work on Japanese seems to support her hypothesis; however, as we shall see later, the results from some of the research presented in this dissertation indicate that Brown’s hypothesis may be too strong, and that some groups of L2 learners do appear able to distinguish contrasts based on novel features (see also Ingram & Park 1998, Hazan et al. 2006, Lieberman 2010). One of the merits of Brown’s work is that it presented a strong and falsifiable hypothesis. As a result, it piqued the interest of researchers and generated several studies aimed at testing her hypothesis.

One of these studies is that carried out by LaCharité & Prévost (1999). While Brown proposes that features that are non-contrastive in the L1 phonemic inventory cannot be acquired, LaCharité & Prévost hypothesize the some non-contrastive features can be acquired, although not all are equally acquirable. As did Brown, they adopt the theory of Feature Geometry, in which features are organized into a hierarchical tree structure. The Feature Geometry tree is composed of organizing nodes, which can dominate other nodes or features. Terminal features are those which are dependent on superior nodes or features but do not themselves have any dependents. Some organizing nodes are also called articulator nodes.

LaCharité & Prévost suggest that the acquisition of a feature that is absent from the L1 phonemic inventory is easier if it requires the generation of a new terminal feature on an already existing articulator node in the feature geometry than if its acquisition requires the generation of a new articulator node. LaCharité & Prévost use both an AX and ABX task to determine how well advanced Québec French learners of English perceive the contrasts [h] vs. ∅ compared to [θ] vs. [t].[2] They predict that the contrast [h] vs. ∅ should be difficult to discriminate because it supposedly involves generating a new articulator node, Pharyngeal.[3] On the other hand, [θ] vs. [t] is predicted to be more easily discriminated since this contrast allegedly involves the generation of a new terminal feature, [distributed]. Results of the AX task support these

[2] In an AX task, participants are presented with pairs of stimuli. Within a pair, the stimuli are either the same (AA) or different (AB). The participant's task is to determine whether the pair consists of two stimuli which are the same or two stimuli which are different. In an ABX task, A and B are always different; for each item in the test, X either corresponds to A or X corresponds to B. The participant's task is to select whether X = A or whether X = B.

[3] See Mah (2011) for arguments against this representation of /h/ for English.

predictions; however, the ABX task did not reveal a significant difference between the two contrasts, contrary to their hypothesis.

LaCharité & Prévost's experiments show that results may vary depending on the task. Different tasks demand distinct types of processing strategies. This dissertation directly investigates disparities elicited by different types of tasks, and we shall later see that results can indeed be task-dependent.

Missing from LaCharité & Prévost's design is the contrast [θ] vs. [s]. This contrast arguably involves the same terminal node as postulated for [θ] vs. [t]. Cross-linguistic differences in the perception and production of these two contrasts are at the heart of the differential substitution phenomenon.

Lombardi (2003) examines cross-linguistic variation in the production of [θ], looking specifically at learners from two languages, Japanese and Thai, that replace it with [s] and [t] respectively. Lombardi provides an analysis of the differential substitution problem using Optimality Theory (OT) (Prince & Smolensky 1993/2004).

In OT, the phonological grammar consists of a series of ranked and violable constraints. For each input (underlying representation), a (theoretically infinite) series of potential outputs (surface representations) is generated. These outputs are evaluated by the constraint ranking. The main classes of constraints are markedness constraints, which militate against cross-linguistically marked structures, and faithfulness constraints, which preserve identity between the output and the input form. The optimal output, the form that ultimately surfaces, is that which has the fewest violations of higher ranked constraints. Language variation principally occurs due to the different ranking of constraints in the L1 grammar.

One advantage of Lombardi's analysis is that it incorporates internal evidence to motivate differential substitution. She suggests that languages whose speakers replace [θ] with [t] are showing the universally unmarked substitute. On the other hand, when the change is from [θ] to [s], this is the result of a marked ranking, instigated by language-internal evidence.

This is exemplified with Japanese and Thai. She suggests that Thai represents the unmarked situation, where the relevant markedness constraints outrank faithfulness constraints. Specifically, she assumes that continuants are universally marked as compared to stops. The following ranking represents the unmarked situation, resulting in [t] substitution in languages with no /θ/: *[θ] »*cont » *stop » IdentManner.

Lombardi hypothesizes that the affrication process in Japanese whereby [t] becomes [ts] before [ɯ] constitutes the language-internal evidence which causes the child to change her constraint ranking from the initial unmarked state. Specifically, Lombardi proposes that affrication requires an "explosion" of the general faithfulness constraint on manner of articulation into its specific components of stop and continuant faithfulness. The following ranking shows the exploded constraint which yields an [s] in place of [θ]: *[θ] » IdentStop » IdentCont » *cont.

Lombardi's analysis captures the Japanese and Thai facts, but her prediction that languages with affrication will substitute [θ] with [s] does not hold for French. In fact, the reported substitutes for European French and Québec French are opposite to Lombardi's predictions. European French does not have affrication of /t/, yet the most widely reported substitute is [s]. On the other hand, Québec French has a well-known process of affrication before high front vowels and yet speakers substitute [t]. This dissertation disputes the role of affrication in differential substitution.

The work by Weinberger, Hancin-Bhatt, Brown, LaCharité & Prévost, and Lombardi largely focuses on phonemic inventories and phonologically contrastive features. However, there is evidence that L2 learners pay attention to non-contrastive phonetic information. For example, although Japanese speakers of L2 English have trouble with the /r-l/ distinction, they are better at hearing the difference in word-final position (Sheldon & Strange 1982; Brown 1997, 1998). This may be because word-finally, English /l/ is velarized; also, the liquid has significant effects on the preceding vowel. This is especially evident when the vowel is tense and long, which creates a [ʊ]-like off-glide before [ɫ]. These phonetic but non-contrastive cues make the distinction between /r/ and /l/ more salient. This result would not be expected if listeners were basing their evaluation on a phonological representation, underspecified for non-contrastive features.

The phenomenon of phonetic approximation also shows that L2 learners take into consideration non-contrastive phonetic information in the target language.[4] For example, Flege (1987) shows that L2 learners of English and French can adjust voice onset times in word-initial onset position in order to more closely approximate those of the target sound. This would not be possible if learners only had access to the categorical information of contrastive features such as [±voice].[5]

Recent work on loanword adaptations has also emphasized the importance of phonetic information. In Cantonese loanwords from English, [v] becomes [w] and not [f]. Peperkamp & Dupoux (2003) consider that this is because [w] more closely approximates the acoustic

[4] The difference between contrastive phonemic features and non-contrastive phonetic features will be elaborated upon in Chapter 2.

[5] Not everyone agrees that the feature in question here is [±voice]. Some argue that the relevant feature in English is [spread glottis] (Iverson & Salmons 1995; Avery 1996).

properties of English [v]. They view this adaptation as being defined by phonetic distance rather than phonological distance (for a similar view, see Silverman 1995).

This dissertation adopts the hypothesis that phonetic distance plays a key role in the language learner's comparison of the target sound and choice of sound substitute. It will be argued that phonetic information is necessary to account for certain facts involving interdental substitution.

As mentioned earlier in this chapter, reports based on production data show that European French and Japanese speakers usually replace English [θ,ð] with [s,z]; while Québec French and Russian speakers are reported to substitute [t,d].[6] The discrepancy in substitution patterns for European versus Québec French is a particularly interesting case. Although the surface vowel system of these two dialects is different enough to allow the possibility of different underlying vowel specifications, the phonemic consonant inventories of both dialects are the same. Most traditional views of phonology consider that European and Québec French have the same underlying consonant inventory (Walker 1984, Hannahs 2007). If we assume this point of view, as I do in this thesis, then it is difficult to understand the differential substitution facts if the basis of this phenomenon is considered to stem from the phonological system of contrastive features. With an identical underlying consonant system, there is no impetus for differential transfer when speakers from each of these dialects are confronted with the novel English segments [θ,ð]. On the other hand, if one considers phonetic differences in the consonants of European and Québec French, then differential

[6] Actually, the Russian facts are not clear-cut in the literature. While several researchers have reported [t,d] as the preferred substitutes for these speakers (Weinreich 1953, Weinberger 1988), others have reported [s,z] (Teasdale 1997). Some of the studies in this dissertation will shed some light on this issue.

substitution may find an explanation. This dissertation investigates this latter possibility not only with respect to European and Québec French, but also regarding Japanese and Russian. To my knowledge, there are no perception studies which compare perceptual patterns for European French and Québec French. This thesis fills this gap by investigating the perceptual patterns for these language groups.

Although much of the literature on differential substitution is based on production data, there are some investigations of the role of perception in substitution (Hancin-Bhatt 1994a,b; Hanulíková & Weber 2010). For example, Hancin-Bhatt (1994a,b) studied the perception and production of English interdental fricatives by Japanese, German, and Turkish speakers. The results from the Japanese group showed that they commonly produce and perceive [θ,ð] as [s,z]. This study indicated that perceptual confusion patterns seem to be carried over into production for Japanese speakers. In other words, when a target sound is associated with a native sound, that native sound will be produced in place of the target. This leads us to another major hypothesis underpinning the work presented in this dissertation: that transfer in production is due to perceptual factors.

This dissertation begins with the assumption that all features are theoretically available to the L2 learner; in other words, learners have access to Universal Grammar. However, the creation of new segmental representations requires exposure to and practice with the L2; therefore, less advanced learners may not yet be able to represent new features or recombine existing features into a new representation. Like Brown (1997, 1998), I assume that if a particular feature is not active in the L1, then it will not be parsed, i.e. processed, at least for less proficient learners. Unlike Brown, I assume that new features can eventually be acquired.

In the current work, a distinction is drawn between three processing levels: a phonological level, which defines language-specific phonemic contrast; a phonetic level, which defines language-specific implementation; and a non-language-specific acoustic level. Previous research (e.g. Werker & Logan 1985) has indicated that different experimental tasks can tap different processing levels.

In this thesis, it is proposed that during the perception process, the acoustic signal is fed into the phonetic component, where all features on the target segment are evaluated: both those that function contrastively as well as non-contrastive phonetic features. The intake form — a fully specified surface representation of the L2 target — is compared with fully specified phonetic representations of the L1. This is accomplished via a mechanism which assesses the auditory salience of intake features against those of native representations. However, the salience of features is not computed in isolation. One feature may be influenced by another feature with which it co-occurs; that is, one feature may enhance or augment the prominence of another feature; conversely, a feature may mute or diminish another feature.

As concerns differential substitution, this mechanism selects the native phonetic representation which is closest to the target segment through an algorithmic evaluation of auditory distance. Since the phonetic inventories of the native and second language are usually different, one or more of a substitute's features may not match that of the target segment. Featural mismatches which are auditorily muted or diminished are preferred over those which are auditorily salient because the discrepancy is less evident.

This thesis presents three experiments aimed at testing how learners of English from different language backgrounds perceive the interdental fricatives of English. The fourth experiment examines the production of target interdentals by English as a second language (ESL) speakers to determine whether production can be correlated with perception. The final study is an acoustic investigation of the native /s/ of European French, Québec French, Russian, and Japanese speakers that measures differences in the production of this coronal fricative. The remainder of this introduction outlines the structure of the thesis.

In Chapter 2, I introduce a model of speech perception and production as well as an algorithm which are intended to represent how learners process L2 sounds, particularly when no such sound exists in their native inventory. The model and algorithm strive to explain why L2 learners match the target segment with different native segments depending on their L1 background.

Chapters 3 and 4 present experiments conducted using an AXB forced choice paradigm. These will be referred to as AXB-1 and AXB-2 respectively. Chapter 5 presents a Picture Identification Task.

The hypotheses underpinning these studies are as follows:

1. Transfer has a perceptual basis; consequently, the substitutes observed in production are the same as those observed in perception.
2. Transfer is based on an assessment of phonetic features, not solely phonological features. As detailed in Chapter 2, the phonetic features of the input are compared with the phonetic features of L1 representations. The L1 representation which is phonetically closest to the input is selected as the best substitute.
    a. Although the phonological representations of consonants may be similar across languages, their phonetic specifications may differ.
    b. The phonetic quality of the native [s,z] predicts the choice of interdental substitute. More specifically, a native [s,z] that is non-strident or weakly strident, due to lack of enhancement, predicts that speakers of that language will perceptually substitute their native [s,z] for the English target non-strident [θ,ð].
3. Different substitutes emerge with distinct experimental tasks.

These hypotheses were first tested in a pilot study (AXB-1, Chapter 3) which investigated advanced learners of English from Japanese (JA), Québec French (QF), and European French (EF) native language backgrounds as well as including a native English control group. Several contrasts were examined under differing vowel and voicing conditions. As mentioned above, previous studies have suggested that different experimental tasks may target different processing levels (e.g. Werker & Logan 1985). Therefore, this task was conducted under two conditions: the first is the Long Interstimulus Interval (ISI) condition, intended to evoke phonological processing, in which contrastive features alone are available, and the second is the Short ISI, intended to evoke phonetic processing, in which both contrastive and non-contrastive features are accessible. We will see that results from AXB-1 fail to evoke differences between the two ISI conditions. Nonetheless, this study does reveal cross-linguistic differences in perception as well as influences from vowel and voicing contexts.

In a second larger experiment (AXB-2, Chapter 4), two proficiency levels were tested, and a Russian (RU) group of learners was added. The AXB-2 task tapped phonetic processing, and yielded results that are consistent with the hypothesis that production errors are based upon perceptual errors.

In Chapter 5, these same language groups and proficiency levels are tested again, but this time on a Picture Identification Task. This type of task must invoke phonological processing because learners must access their internal lexical representations of pictured objects and determine which phonological representation corresponds to the auditory stimulus. Therefore, with the Picture Identification task, we can get a clearer idea of what learners can perceive in this type of processing. Results from this task show that learners are able to make distinctions that are not part of their native phonology, even at lower proficiency levels, indicating that phonological representations include non-contrastive information. And again, results are consistent with the hypothesis that perception underlies production.

Chapter 6 investigates production from two angles. First, speakers from EF, QF, JA, and RU native language backgrounds completed a Word Production task involving target "th". The results from this production task confirm other production reports and establish [s] as the preferred production substitution for Russian ESL learners. It also reveals correlations between perceptual errors and production errors.

The final study, reported on in Chapter 6, directly addresses the hypothesis that L1s with a non-strident or less strident [s] will substitute this sound in place of [θ]. Production data are subjected to an acoustic analysis to investigate whether there are differences in the articulation of the coronal fricative [s] in each of the different test languages. The results from this acoustic analysis show differences between languages on the dimension of spectral mean, relative amplitude, and relative length. These results suggest that stridency may be manifested in different ways at the acoustic surface.

Chapter 7 concludes with a summary and discussion of the results from the various experiments conducted and puts them in perspective within the wider question of how second language sound systems are acquired.

CHAPTER 2: THE AUDITORY DISTANCE MODEL

2.1. Introduction

Why do some second language (L2) learners use [t,d] when attempting to produce the English interdental fricatives, while others use [s,z], and still others use [f,v]? In this chapter, I address this question by presenting a speech processing model as well as an algorithm which together are intended to simulate the process that takes place when learners encounter an L2 sound that does not figure among the sounds in their native language (L1) inventory.

Before proceeding, let us recall the hypotheses behind the research in this thesis:

1. Differential substitution is due to transfer from the L1;
2. Transfer in production is caused by transfer in perception;
3. All features are potentially available in phonetic processing; therefore, the choice of substitute depends on a comparison of the phonetic properties of the target segment with phonetic properties of segments in the L1 sound system;
4. Transfer of non-contrastive, phonetic features is involved in substitution.

As mentioned in Chapter 1, there is an abundance of research in Second Language Acquisition (SLA) that has shown that many learner errors are due to transfer from the L1 (e.g. Weinreich 1953, Trubetzkoy 1939/1969, Eckman 1977, Robinett & Schachter 1983, Wode 1983, Ioup & Weinberger 1987, Flege 1995, James 1996). In fact, transfer is what leads native speakers to distinguish between different foreign accents, e.g. to identify a Japanese speaker of English versus a French speaker of English. Although researchers have also shown that transfer cannot explain all of foreign accent, and that some role must be given to universal phonological and phonetic factors (e.g. Major 1987, 1995, 2001; Hurch 1986, Nathan 1989), transfer remains perhaps the most important component of foreign accent. In this thesis, I hypothesize that the differential nature of interdental substitution is due to L1 transfer.

The interplay between perception and production is a long-standing issue in SLA research. Polivanov (1931) and Trubetzkoy (1939/1969) considered that L2 sounds are (mis)perceived by the L1 system, and concordantly (mis)produced. Intuitively, it makes sense that accurate production implies accurate perception, and conversely, inaccurate production is due to inaccurate perception. More recently, many researchers have argued for the primacy of perception over production (e.g. Neufeld 1988, Flege 1991 and other work, Brannen 1996, Escudero 2005). Indeed, many studies have shown that transfer in production has a perceptual basis. In other words, errors that are observed in production are due to the learner misperceiving the target sound and mapping it onto an L1 sound (e.g. Best & Strange 1992; Hancin-Bhatt 1994a,b; Rochet 1995; Brown 1997,1998; Flege et al. 1999).

Not all studies have shown this effect however; some have shown differences between what is perceived and what is produced (e.g. Goto 1971, Nemser 1971, Sheldon & Strange 1982, Flege & Eefting 1987). Some of these findings, however, can be attributed to other factors. For example, as Sheldon & Strange (1982:254) themselves point out, the relatively accurate production of /r/ and /l/ by Japanese ESL learners, despite their difficulty in perceiving this contrast, may be due to explicit instruction on their articulation. This would be supported by orthography as well. In some cases, methodological problems in task design may have biased the results (Escudero 2005).

Thus, arguments and evidence favour the view that perception precedes production, and this is the hypothesis I adopt in this thesis:

production errors involving interdental fricatives are due to perceptual errors.

Some researchers have argued that choice of sound substitute is sometimes not determined by the structure of the phonological or phonetic system (either L1 or L2), but that factors such as the nature of L2 input from other members of the learner's social group or overt choices based on social conventions may be involved (Roy 1992, Paradis & LaCharité 1997). Others propose that substitution need not be based on L1 internal evidence, i.e. that the L1 may not display any phonological or phonetic properties or processes that would point to one or another transfer variant being selected as a replacement for the target interdental segment, but rather that substitution behaviour in itself can constitute evidence for the structure of the L1 phonological or phonetic system (Weinberger 1988).

As mentioned in Chapter 1, I argue that these ideas are untenable because they pose a problem for learnability. How can structure that differs cross-linguistically be learned in the absence of evidence? If the L1 gives no evidence to lead the learner in the direction of one substitute or another, then one should expect either intra-individual (free) variation or a unitary universal substitute, but not cross-linguistic differential substitution. Further, if the transfer variant depends on social conventions, one would predict different behaviour from those exposed to different conventions, or from those who reject such conventions. What would we expect from the learner who has never heard L2 input from a member of her social group? I hypothesize that the learner relies on evidence from her L1 sound system in order to guide her choice of substitute.

As discussed in Chapter 1, some researchers propose that segmental transfer occurs at the phonological level of representation (e.g. Weinberger 1988; Hancin-Bhatt 1994a,b; Brown 1997,1998; Lombardi 2003). In other words, substitution is based on an evaluation of L1 *contrastive* features. The problem with this idea is that there are cases of differential substitution in which two or more of the L1s involved have the same or similar phonemic inventories (e.g. Rochet 1995). For example, in the case under focus, European French speakers produce [s,z] in place of the target interdentals (e.g. Weinreich 1953) whereas Québec French speakers produce [t,d] (e.g. Jamieson & Morosan 1986, Teasdale 1997) in place of these targets despite the fact that they have the same phonemic consonants and thus arguably the same contrastive feature inventory. If transfer is based on phonemic representations containing only contrastive features, then L1 evidence cannot determine the different choice of substitute that European and Québec French learners of English make with regards to the target interdental fricative. However, although the phonemic consonant inventories of these two dialects are the same, their phonetic inventories differ (Caramazza & Yeni-Komshian 1974, Teasdale 1997, Hannahs 2007). Thus, I hypothesize that differential substitution is based on L1 evidence from the phonetic level of representation, which embodies both contrastive *and* non-contrastive information.

2.2. Substitution has a Perceptual Basis

As stated in §2.1, I hypothesize that production is based on perception. For example, if a learner substitutes [t,d] for [θ,ð] in production, it is because target [θ,ð] has been associated with a perceptual representation for [t,d].

I propose that transfer is based on the perception of intake features. The raw acoustic signal is processed by the brain and converted into a form that can be used by the linguistic system, i.e. a set of features, both contrastive and non-contrastive. This form is called the *intake*. At the lowest level of linguistic processing, the intake is scanned, then only those features which serve to define phones in the language are retained or processed; those with no phonetic status are stripped from the intake. As discussed in detail in the next section, if, for example, a language has laminal obstruents but no apical obstruents, then the feature LAMINAL will be processed and represented, while APICAL will not. Features which are retained from the intake I will call *intake features*. These intake features serve to define the language-specific phonetic system.

As proposed by Stevens et al. (1986) and others, I consider that the saliency of certain features is enhanced by other features, making their presence more prominent to the listener. For example, the degree to which the difference between a stop and a fricative is perceived depends on the phonetic and phonological representations of the obstruents in the L1 as well as their enhancement relationships. Thus the features that distinguish stop consonants from fricatives (STOP and CONTINUANT) can be subject to language-specific influences.

The relation between enhancement and phonological prominence is related to markedness and typological frequency (e.g. Stevens et al. 1986). Some feature combinations occur more frequently in languages of the world, and it is reasonable to assume that one reason for this is because of the auditory salience of such combinations.

It is argued that some features are inherently salient, as all languages make use of them. This is the case for the Major Place features, viz. LABIAL, CORONAL, and DORSAL. However, other features require enhancement to boost their auditory salience. Not all languages make use of enhancement for a given feature. Thus, in addition to inherently salient features, which I consider to be "unmarked" in the sense that they are found in all languages, non-salient features also exist cross-linguistically. In some languages a feature is salient due to enhancement by another feature that co-occurs with it; whereas, in others, it is not enhanced and hence is attenuated or muted. This will be elaborated on in §2.5. I suggest that non-contrastive features and/or the effects of enhancement play a role in cross-linguistic differences in perception that have been found by various researchers.

As discussed above, I hypothesize that both contrastive and non-contrastive features play a role in transfer. Transfer takes place when the L2 intake is associated with an L1 auditory phonetic representation. This representation is (indirectly) mapped to a phonetic articulatory representation, resulting in production transfer.[7]

2.3. Features

While there is a close relationship between phonetic and phonological features and representations, there are some fundamental differences between the phonetic and phonological components of the grammar. I consider that phonology differs from phonetics in three principal ways. First, I assume that phonological representations are relatively abstract and interact with information from all modalities, for example, auditory, articulatory, and visual (see e.g. Hardison 1999 for discussion). Phonetic features and representations on the other hand, while still abstract in the sense that they form part of the linguistic system, are more closely tied to anatomical systems, either articulatory or auditory. The acoustic signal is transduced into auditory features, and articulatory features are precursors to motor instructions to the articulators.

Second, I consider that features that function phonologically are categorical: a feature is either present or absent in the system; there is no variation in degree of specification. On the contrary, phonetic features are gradiently specified, due to the effects of enhancement or lack thereof.

---

[7] The mapping is "indirect" because it first passes through a phonological representation; this is shown in the model in Figure 1 below.

Third, I adopt the position that phonological features function contrastively within the sound system, serving to distinguish one segment from another. This means that non-contrastive phonetic features are not active in the phonology, i.e. they are underspecified (see e.g. Steriade 1995, Dresher 2009 for discussion). In this thesis, I adopt a version of the theory of Contrastive Underspecification (e.g. Steriade 1987; Calabrese 1988; Mester & Itô 1989; Rice & Avery 1995). In Contrastive Underspecification, a segment is specified for a feature only if that feature serves to distinguish that segment from another in the inventory of a particular language. The general idea behind underspecification theory comes from empirical evidence which shows that non-contrastive features tend to be inactive in phonological processes. For example, Steriade (1987), in discussing Lamda height assimilation, argues that the feature [–high] is contrastive for mid vowels, but not for low vowels. Mid vowels are triggers in a process which spreads [–high]; however, low vowels do not trigger this process. Being redundant on low vowels, [–high] is underspecified.

In the phonetic component, however, I consider that non-contrastive features are specified. Features found in the phonetic component of a given language form a subset of the universal set of features. For example, the features APICAL and LAMINAL would be specified both in the *phonological and phonetic* components of Australian languages which contrast laminal and apical stops (Hamilton 1996). On the other hand, in English, APICAL is arguably absent from the *phonological* module: since English has only one series of coronal stops, APICAL is non-contrastive. However, given that English coronal stops are articulatorily described as being APICAL (Dart 1991), the feature APICAL specifies how the coronal stops are phonetically implemented and thus would be present in the *phonetic* component. Nonetheless, even if absent both phonologically and phonetically from the grammar, APICAL and LAMINAL are features which remain part of the universal repertoire of features, and hence a learner could potentially acquire and incorporate hitherto dormant features such as these. In other words, learners are hypothesized to have full access to the complete universal array of phonetic and phonological features.

While I hypothesize that L2 learners have access to all features, i.e. that Universal Grammar is available in SLA, this does not necessarily mean that in L2 acquisition, they are able to instantaneously incorporate new features into their L2 sound system or recombine existing features into a new perceptual representation. If their interlanguage grammar has not yet developed to the point where it allows the appropriate new representation, learners must attempt to map the perceived features onto those of existing stored phonetic representations in their native inventory. This results in perceptual transfer.

2.3.1. Articulatory and Auditory Features

In the spirit of Jakobson, Fant, & Halle (1963/1969), Keating (1990), Hamilton (1996), Flemming (1995/2002), and Boersma (1998), I adopt the view that there exist auditory features in addition to articulatory features. I assume that speech is perceived in terms of auditory features and produced in terms of articulatory features.

Equivalencies between articulatory and auditory features under discussion are shown in (1) below. The first column shows types or classes of features. Articulatory features are listed in the second column, and their auditory counterparts, in the third column.[8]

---

[8] Except for those in the Major Articulator and Location class, these features are binary, but each value is given a different name. I use small capital letters when referring to features.

(1) Articulatory and Auditory Features

Type of Feature	Articulatory Features	Auditory Features
Major Articulator (Place)[9]	LABIAL, CORONAL, DORSAL	F2 RISES, F2 STEADY, F2-F3 CONVERGE[10]/MID, HI, LO SPECTRAL MEAN
Location	LIP, DENTAL, ALVEOLAR, POST-ALVEOLAR	LOW F2, LOW-MID F2, MID F2, HI F2/ LO-MID, HI, MID-HI, MID-LO SPECTRAL MEAN
Minor Articulator	LAMINAL, APICAL	SLOW, FAST TRANSITION
Airflow	CONTINUANT, STOP	CONTINUANT, INTERRUPTED
Turbulence	CHANNEL, SPREAD	STRIDENT, MELLOW

Major Articulator features reflect the major active articulators involved, for example, the lips (LABIAL), the tongue tip or blade (CORONAL), and the tongue body (DORSAL). LOCATION refers to the passive articulator, for example, LIP, DENTAL, ALVEOLAR, POST-ALVEOLAR (e.g. Gnanadesikan 1993). Major Articulator (Place of Articulation) and Location features, which represent a finer gradation on the place dimension, are auditorily cued by formant transitions for obstruents (e.g. Wilde 1995, Benki 1998), or by the relation between formants for stop consonants (Jakobson et al. 1963/1969,

[9] Although technically Major Articulators, LABIAL and CORONAL as well as DORSAL are often referred to as Place features (e.g. Sagey 1986).

[10] Although these values refer to formant transitions from stop consonants into the following vowel, Wilde (1995) indicates that F2 onset in a post-fricative vowel is a reliable identifier of place, with the lowest F2 onset following labial fricatives, and F2 progressively rising the farther back the articulation.

Keyser & Stevens 2006) or by spectral energy for fricatives, e.g. SPECTRAL MEAN (e.g. Strevens 1960; Jongman et al. 2000).[11]

The Minor Articulator refers to the configuration of the tongue blade and whether the tongue tip is raised (APICAL) or whether the tip is down and/or the constriction wide and flat (LAMINAL). The auditory features that correspond to APICAL and LAMINAL are SLOW and FAST TRANSITION (Stevens 1998).

The Airflow features (analogous to Manner) relate to the degree of constriction of obstruents: total constriction defines a stop, and narrow constriction characterizes a continuant. Acoustically, a stop corresponds to a brief period of silence followed by a burst of noise, INTERRUPTED (Jakobson, Fant, & Halle (1963/1969); for a continuant, there is steady noise with no period of silence, and consequently, no burst.

Turbulence features involve a degree of noise: STRIDENT fricatives are noisy because the airflow is directed through a narrow channel at high velocity; the airstream may also be deflected off an obstacle such as the teeth. Acoustically, these sibilant sounds have a spectral peak of intensity below 10kHz (Hughes & Halle 1956). On the other hand, mellow fricatives are produced with a long narrow slit between the active and passive articulators, thus their synonym, slit fricatives. Because of the wide channel and lack of obstacle, these non-sibilant sounds generate little noise. They have a flat smooth spectrum below 10kHz with very low intensity. It has also been found that formant transitions into a following vowel may serve to distinguish between strident and non-strident sounds (Heinz & Stevens 1961, Wagner et al. 2006). Although stridency is usually attributed to fricatives, affricates too are often considered to be "strident stops" due to

---

[11] There is debate as to whether it is locus equations between the F2 onset and the F2 in the mid portion of the vowel which determine place of articulation (Fowler 1994).

the fricative portion following the stop (Jakobson, Fant, & Halle 1963/1969; Rubach 1994). The auditory features INTERRUPTED and MELLOW/STRIDENT are borrowed from Jakobson, Fant, & Halle (1963/1969).

Though this dissertation is primarily concerned with auditory features, for the sake of clarity, I will usually refer to the more familiar articulatory labels such as CORONAL and STOP, rather than the perhaps less familiar terms such as F2 STEADY or INTERRUPTED. The latter labels may be implied. Exceptions to this convention are the features STRIDENT and MELLOW.

2.4. The Auditory Distance Model

In this section, I elaborate a model of speech perception and production which demonstrates the mechanisms behind transfer at both the phonemic and phonetic levels. The Auditory Distance Model (ADM) strives to explain how language learners process L2 phonetic input via an algorithm that assesses the auditory distance between the target and L1 representations by incorporating the role of features and feature salience through enhancement.

In the speech processing model in Figure 2.1 below, we see that the acoustic signal enters the language module, and this linguistic form is labeled *intake* (Brown 1993). (Intake features are enclosed in curly brackets.) The features which are present in the intake form depend on the type of processing involved. In phonetic processing, surface features are present in the intake; and as will be discussed below, enhancement is active at this level (Keyser & Stevens 2006).[12] In phonemic processing however, only contrastive features are retained (as shown by the strikethrough of non-contrastive features), and enhancement does not apply. The intake form is mapped to the closest L1 representation (see §3.5.4). In a phonemic assessment, the intake form is equivalent to an underlying representation; in a phonetic assessment, it is an auditory surface representation. During production, the underlying representation feeds into an articulatory representation; the latter determines the spoken output.

[12] Keyser & Stevens present a production model only. In their model, phonological features in the production planning stage are flagged if of low saliency, then in the phonetic component, features that serve to enhance a contrast (enhancement features) are added. I assume that in perception, enhancement features are present in the phonetic component, but are shed in the phonological component.

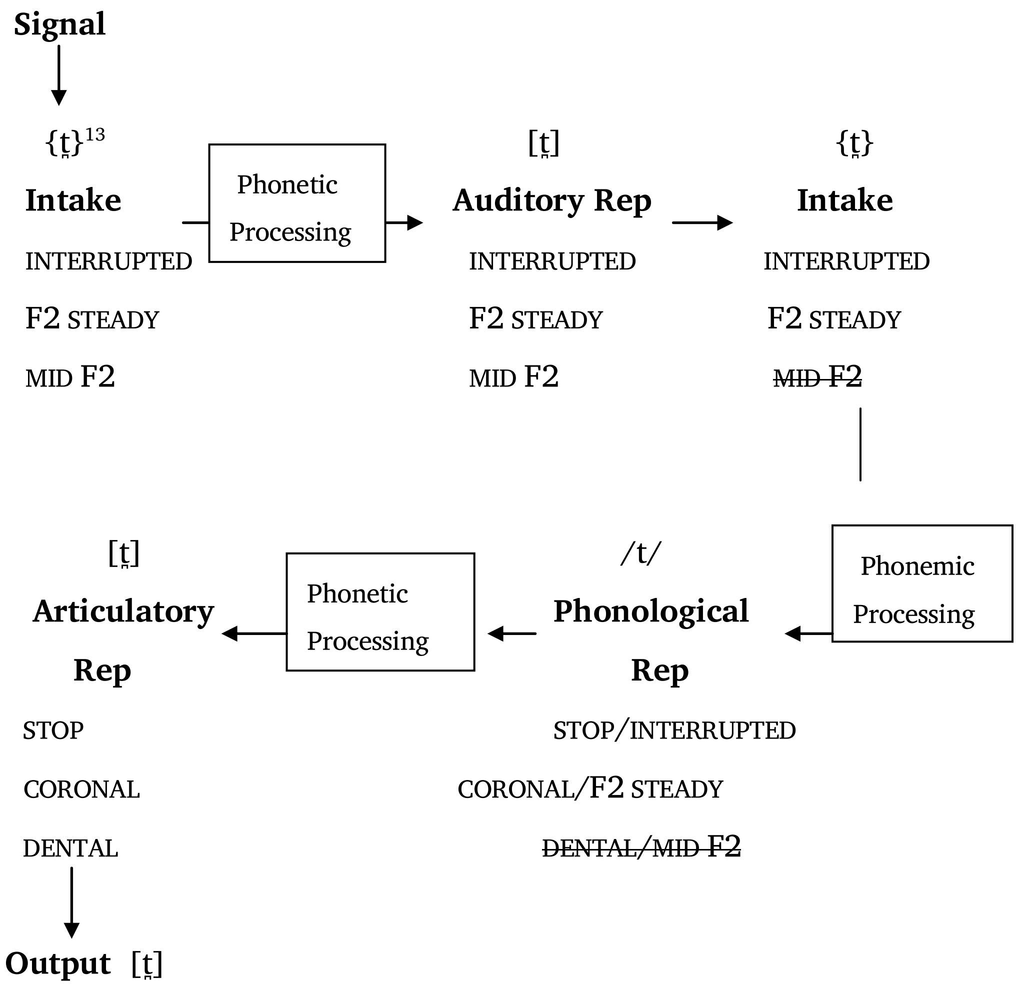

Figure 2.1. The Perception-Production/Phonetics-Phonology Link

This view of the grammar provides a starting point within which to situate issues discussed in this study. We begin with discussion of the role of features and saliency in perception.

[13] Processing likely proceeds in syllable-sized chunks (Peperkamp & Dupoux 2003). Although descriptions and argument in this thesis focus on featural and segmental units, some of the results showing consonant-vowel interactions are best explained if we consider the perceptual window to encompass a syllable.

2.5. Features and Saliency

There is a growing literature on the role of auditory salience in determining the shape of phonological systems (e.g. Hamilton 1996, Flemming 1995/2002, Hume & Johnson 2001, Wright 2004). All else being equal, salient features are preferred over less salient features.

Many of these perceptual theories of phonology and phonetics have drawn on the construct of auditory cue enhancement. Much has been written on the role of perceptual salience and enhancement in optimizing language-specific phonological contrasts (e.g. Flemming 1995/2002, Boersma 1998, Keyser & Stevens 2006, Boersma & Hamann 2007).

The work in this thesis was inspired by the theory of feature enhancement developed by Stevens, Keyser, and others (Stevens et al. 1986, Stevens & Keyser 1989, Keyser & Stevens 2001, 2006; Clements 2005). Keyser & Stevens describe enhancement as occurring when one feature serves to boost the prominence of an otherwise perceptually weak feature. It has been shown that certain features are poorly discerned in the speech signal, often depending on adjacent context or position in the syllable. This is exemplified with the [ʃ] of English. They hypothesize that in the lexicon, this segment is contrastively specified as [-anterior], in order to distinguish it from [s] which is [+anterior]. They point out that the acoustic correlate of [-anterior] is a low spectral prominence in the F3 range. This acoustic property may be only weakly present for a post-alveolar fricative. However, lip rounding can increase the F3 peak. Thus, Keyser & Stevens propose that in this example, [+round] is a non-contrastive enhancement feature implemented in the phonetic component, which serves to increase the salience of the [-anterior] feature of [ʃ] so it will be more clearly distinguished from [+anterior] on [s].

The focus of this thesis is not so much on perceptual robustness with regard to optimal systemic contrasts, but on the role of perceptual strength in matching a non-native input with a native representation during L2 processing. While I agree that enhancement can serve to augment the distinction between two features that function contrastively in a language, I suggest that in some cases enhancement may also apply to features that do not function contrastively in a language. Furthermore, independent of whether or not they function contrastively, it is contended that enhanced features are preferred over "muted", i.e. non-enhanced features.

It has been proposed that perceptual robustness may account for the cross-linguistic frequency of some types of segments over others (e.g. Wright 2004). Thus, a consequence of enhancement is typological frequency. Based on evidence from cross-linguistic data, perceptual confusion studies, and/or auditory analysis, I propose the following feature enhancements of interest in this thesis:

ROUND ENHANCES POST-ALVEOLAR: The feature POST-ALVEOLAR is equivalent to [-anterior]; thus this is the same enhancement relation as discussed above regarding [s] and [ʃ]. Keyser & Stevens (2006) state that the defining characteristics of [ʃ] is energy in the F3 region; however, this energy is rather weak. They argue that lip rounding serves to augment the energy in the area of the third formant. Keyser & Stevens (2001) suggest that rounding on [ʃ] is frequent cross-linguistically, occurring in several languages such as English, French, and Polish.

STRIDENT ENHANCES CONTINUANT (Stevens et al. 1986): Stevens et al. consider that this enhancement serves to make a clearer perceptual distinction between stops, which are [-strident] (a.k.a. MELLOW), and fricatives: The most salient continuant segment is STRIDENT. In comparison to mellow continuants, e.g. [f] and [θ], strident continuants such as [s] and [ʃ] are more common cross-linguistically (Crothers et al. 1979, Maddieson 1984).

MELLOW ENHANCES STOP (Stevens & Keyser 1989): Affricates (strident stops) are less frequent cross-linguistically than are non-affricated stops (Maddieson 1984). This is perhaps due to the potential for perceptual confusion between affricates and continuants. Stevens & Keyser use the negative value [-strident] (MELLOW) to enhance [-continuant] (STOP) to account for the observation that /t/ vs. /s/ is a better contrast than /t^{s}/ vs. /s/.

ALVEOLAR, POST-ALVEOLAR ENHANCE STRIDENT: In her 1985 thesis, Shadle elaborates mechanical models of fricative consonants and compares them with acoustic analyses of real speech. In one component of her work, she manipulates the distance between place of constriction (articulation) and obstacle (teeth). Modeling fricatives corresponding to [s], [s̡], and [ʃ] (where [s] is alveolar, [ʃ] is post-alveolar, and [s̡] is intermediate between these two), she finds that the amplitude in decibels (≈stridency) increases as the distance between constriction and obstacle increases. Extrapolating from this data, I suggest that dental [s̪], with the shortest distance between constriction and obstacle, has a lower degree of stridency than does alveolar [s], hence ALVEOLAR ENHANCES STRIDENT; whereas, DENTAL *does not* enhance STRIDENT. This latter relationship is key to the arguments developed in this thesis; therefore, I will often refer to it as DENTAL MUTES STRIDENT.

The transitive relation between ALVEOLAR ENHANCES STRIDENT and STRIDENT ENHANCES CONTINUANT means that we should expect to find a higher frequency of alveolar [s] cross-linguistically as compared to dental [s̪]. This is confirmed in Maddieson's 1984 database; notably in inventories that contrast /s/ with /θ/, further supporting the idea that alveolar [s] is more strident than dental [s̪]. Thus, I propose that alveolar [s] is perceptually farther from [θ] than is dental [s̪], which has implications for L2 perception and transfer.

2.6. Calculating Auditory Distance

Transfer occurs when the intake form is compared to stored representations, either phonetic or phonological.[14] Features in the intake which do not match those in a stored representation are noted. The "distance" of these features from stored correspondents is evaluated, and the intake is associated with the L1 representation that is closest.

The comparison of the target intake with native representations involves assessing the number of featural mismatches. Moreover, goodness-of-fit is assessed within featural matches. The latter is akin to Best's (1995) Category Goodness classification, wherein an L2 sound is equated with a given native sound, but may be perceived as being a good or poor exemplar of that native sound.

Featural mismatches and the degree of similarity of featural matches are assessed as to their cost. For example, it is more costly to substitute a highly salient feature in place of the intake feature than it is to substitute a recessive or less salient feature (see also e.g. Steriade 2001, Kenstowicz 2003). This is because the "violation" is more easily perceived in the case of a salient feature; in other words, it is more evident that one is not being faithful to the auditory intake if one substitutes a salient feature. On the other hand, if one substitutes a less perceptible feature, it is less costly. So, if

[14] More precisely, I assume that what is stored are features and their combinatorial possibilities. In other words, segments are constructed "on the fly" during speech production and perception.

presented with a choice, it is better to substitute a “wrong” (non-target appropriate) feature which does not readily stand out, than a wrong feature which is blatantly wrong. When native representations are compared with the intake, features which do not match are tagged as mismatches, and the degree of similarity in featural matches is also earmarked. Thus featural mismatches and category goodness enter into the calculation of auditory distance. Auditory distance is measured by how far the mismatched feature in the native representation is from the corresponding intake feature.

Let us now examine the mechanism that assesses the auditory distance between target and L1 representation. During a comparison between target and L1 features, if it is found that the feature on the L1 representation does not match that of the target, it falls in the mismatch region and receives a negative value, either -1 or -2. In some cases, the target and L1 feature may be the same, but their salience differs, due to degree of inherent saliency or enhancement/muting; in these situations the L1 feature falls in the target region. The evaluation of distance between target and the L1 representation takes place along a scale:

$$\underbrace{\overset{-2}{\bullet}\text{----------}\overset{-1}{\bullet}\text{----------}\overset{0}{\bullet}}_{\text{Mismatch region}}\underbrace{\text{----------}\overset{+1}{\bullet}\text{----------}\overset{+2}{\bullet}}_{\text{Target region}}$$

Figure 2.2. Auditory Distance Scale

In Figure 2.2 we see that the target feature has a positive value: if it is a strong (i.e. salient) feature, its value is +2; if it is a weak (i.e. non-salient) feature and not enhanced by another feature, its value is +1; if enhanced, its value is $+2$. For the features of interest in this thesis, I argue that the major place features are strong ($+2$); other features are weak ($+1$) unless enhanced ($+2$). This is shown in (2) below.

(2) Feature weighting

Feature	Inherent Weight	Enhanced Weight
LABIAL, CORONAL, DORSAL	2	n/a
STOP, CONTINUANT	1	MELLOW ENHANCES STOP $\Rightarrow$ 2 STRIDENT ENHANCES CONTINUANT $\Rightarrow$ 2
STRIDENT, MELLOW	1	ALVEOLAR, POST-ALVEOLAR ENHANCE STRIDENT $\Rightarrow$ 2[15]
LIP, DENTAL, ALVEOLAR, POST-ALVEOLAR	1	ROUND ENHANCES POST-ALVEOLAR $\Rightarrow$ 2
LAMINAL, APICAL	1	n/a
ROUND, UNROUND	1	n/a

The only features that have an inherent weight of 2 are the articulator-bound major class features LABIAL, CORONAL, and DORSAL. Major Place features are considered strong since virtually all languages have a distinction between LABIAL, CORONAL, and DORSAL segments. As well these features have prominent status in the phonological component (e.g. Sagey 1986/1991).

As discussed in the previous section, the acoustic counterparts of these features are F2 values and spectral moments. The acoustic differences are of course greater between the major places of articulation (LABIAL, CORONAL, and DORSAL) than they are between location (LIP, DENTAL,

[15] Note that DENTAL MUTES STRIDENT does not reduce the inherent weight, as this is a short-hand notation for DENTAL DOES NOT ENHANCE STRIDENT.

ALVEOLAR, POST-ALVEOLAR), thus the different inherent weighting for each of these types of place features.

The Airflow and Turbulence features, STOP/CONTINUANT and STRIDENT/MELLOW, as well as the features tied to coronality, are intrinsically weak (a.k.a. recessive), but can have their saliency enhanced by other features as shown above.

Let us look at how this works with a hypothetical example. In (3) below, the feature values for the fully specified intake form are in the upper left corner. All target intake features by definition have a positive value, and thus are positioned on the right side of the scale (see Figure 2.2 above). The feature CORONAL is inherently salient and so has a weight of +2; the feature STOP of the intake is not salient (recessive) and thus has a weight of +1. In this example, the target feature STRIDENT is enhanced by ALVEOLAR, and thus it has a weight of +2.[16]

This intake form is compared with stored representations, whose features are listed in the middle column. Features of the same type are compared, viz. Major Articulator, Airflow, Turbulence, Location, and Minor Articulator (see (1) in §2.3.1). For Candidate representation #1, the Major Articulator feature CORONAL is a perfect match to the intake feature CORONAL, since both have a value of 2; thus, there is no distance between the intake feature and the candidate feature. The intake feature STOP is recessive (weak) and therefore has a value of 1.[17] But Candidate #1's STOP feature is enhanced by MELLOW (as shown by the arrow), so it has a value of 2. Although both intake and candidate agree on the feature STOP, as shown by the plus sign, they differ in degree, so there is a distance of 1 between

[16] See §2.5 for the motivation behind these weightings.

[17] This is because the fricative portion of the affricate makes it less "stop-like"; as well, interruption of airflow may be less for an affricate than it is for a stop (Miller-Ockhuizen & Zec (2003)).

them. Candidate #1 is specified as MELLOW, so this constitutes a mismatch with the (enhanced) intake feature STRIDENT; therefore, there is a negative value incurred in the evaluation: in this case, MELLOW falls on the left side of the scale (-1), compared to the intake (+2), which results in a distance of 3. The last feature, non-salient location ALVEOLAR, is shared by both the intake and the candidate, so there is no distance between these. The total distance is the sum of distances incurred for each type of feature: for Candidate #1, the total distance is 4.

Candidate #2 matches the intake exactly for the features CORONAL, STRIDENT, and ALVEOLAR. They disagree on Airflow however: the intake is a non-salient STOP, while Candidate #2 is a salient (enhanced) CONTINUANT, which results in a total distance of 3. Thus Candidate #2, with a distance of 3, is closer to the intake than is Candidate #1, with a distance of 4, and is therefore selected as the substitute. Note that if either the intake or candidate is not specified for a particular type of feature, this type will be ignored, and no distance will be calculated.

(3) Selection of Candidates

	Evaluation of mismatches			Distance on Scale
Intake Target Form [ts]:	CORONAL	(salient)	+2	
	STOP	(recessive)	+1	
	STRIDENT	(enhanced)	+2	
	ALVEOLAR	(recessive)	+1	
Candidate #1 [t]	CORONAL	(salient)	+2	0
	STOP	(enhanced)	+2	1
	MELLOW	(recessive)	-1	3
	ALVEOLAR	(recessive)	+1	0
				Total = 4
☞ **Candidate #2 [s]** [18]	CORONAL	(salient)	+2	0
	CONTINUANT	(enhanced)	-2	3
	STRIDENT	(enhanced)	+2	0
	ALVEOLAR	(recessive)	+1	0
				Total = 3

2.7. Auditory Representations for Language Groups

We will soon turn to the auditory distance calculations for each language to be examined in this thesis, i.e. European French, Québec French, Japanese (and Russian in Chapter 4). First however, let us examine the phonetic representations assumed for these languages.

[18] This evaluation is similar to Optimality Theory, but there is no constraint ranking. Instead, what is important is the distance between the target feature and the candidate feature.

As discussed in Chapter 1, two languages can have the same underlying representations, with different phonetic manifestations of these representations. Below I give the phonetic specifications for the segments of relevance for European French, Québec French, and Japanese.[19] The features which function contrastively in each language are in bold and underlined.

2.7.1. European French

The phonetic inventory for European French (EF) contains a labiodental fricative, an apical dental stop and a laminal dental fricative (Dart 1991; Teasdale 1997). Dart describes the EF stop as apico-laminal, articulated with the tongue tip up and touching the upper teeth. She argues that it can be considered to be APICAL alone, since it is difficult to make an apical dental stop without also involving the blade. Regarding dentality, although Dart finds some variation in the production of [s,z] in EF, she notes that they "tend to be articulated with closure contacting some part of the upper incisors" (p.22).

These segments function phonemically and can occur before any vowel. They are represented in (4) below.

---

[19] The post-alveolar fricatives [ʃ,ʒ] are not included here as they were not tested in the first experiment (AXB-1). Refer to Appendix A for the calculations for these segments.

(4) European French Phonetic Representations[20]

[f,v]	[t̺,d̺]	[s̻,z̻]
<u>**CONTINUANT**</u>	<u>**STOP**</u>	<u>**CONTINUANT**</u>
<u>**LABIAL**</u>	<u>**CORONAL**</u>	<u>**CORONAL**</u>
MELLOW	MELLOW	STRIDENT
DENTAL	DENTAL	DENTAL
	APICAL	LAMINAL

2.7.2. Québec French

The Québec French (QF) phonetic inventory contains a labiodental fricative, a laminal dental stop, and a laminal alveolar fricative (Teasdale 1997). Teasdale gives evidence that the QF fricative is alveolar, but not that it is laminal. I have chosen to represent it as LAMINAL in accordance with EF. Dart (1991) notes that Charbonneau & Jacques' (1972) spectrograms show that QF [t̻,d̻] is also apicolaminal, along with EF. I argue however that QF [t̻,d̻] is specified as LAMINAL while EF [t̺,d̺] is specified as APICAL. I suggest that this difference accounts for why QF has affrication of stops before high front vocoids while EF does not. Annie Brasseur (p.c.) states that the coarticulatory period between the occlusion of the QF stop and following vocoid is longer than it is in other languages and dialects of French. This coincides with Stevens' description that a stop with a flattened blade (LAMINAL) results in a much slower release than an apical stop. Nevertheless, whether or not QF /t,d/ are specified as LAMINAL or APICAL, they will be selected over [s,z] as substitutes for [θ,ð].

All of these segments function phonemically. The labiodental and sibilant can occur before any vowel; however, [t̻,d̻] are affricated before

[20] The diacritics ̺ and ̻ represent apico-dental and lamino-dental respectively.

high front vowels.

(5) Québec French Phonetic Representations[21]

[f,v]	[t̪,d̪]	[s,z]
<u>**CONTINUANT**</u>	<u>**STOP**</u>	<u>**CONTINUANT**</u>
<u>**LABIAL**</u>	<u>**CORONAL**</u>	<u>**CORONAL**</u>
MELLOW	MELLOW	STRIDENT
DENTAL	DENTAL	ALVEOLAR
	LAMINAL	LAMINAL

2.7.3. Japanese

The Japanese (JA) phonetic inventory includes a voiceless bilabial fricative, an apical alveolar stop (Someda 1966, cited in Vance 1987:18; Akamatsu 1997), and a laminal alveolar (MELLOW) fricative (Vance 1987; Li 2008; Teasdale 1997; Toda & Honda 2003). Li describes JA [s] as less strident than English [s]. Teasdale's spectrographic analysis of JA [s] shows it to be MELLOW. Toda & Honda describe it as alveolar. The bilabial fricative [ɸ] has traditionally been analyzed as an allophone of /h/ before [ɯ]. However, [ɸ] does appear to be contrastive in recent loanwords, appearing before every vowel. Vance (2008) thus argues that Japanese has undergone a phonemic split in the last 50 years, and that [ɸ] now has contrastive status in JA. Whatever its status, [ɸ] is included here; since it is hypothesized that transfer occurs at the phonetic level of representation, allophones are admissible as possible candidates. The other two JA segments clearly function phonemically: [s̠,z̠] occurs before any vowel

[21] The affricates [t̪s̪,d̪z̪] of QF are not included as they were not contrasted with [θ,ð] in this study. For auditory distance calculations for these affricates, see Appendix A.

except [i], and [t,d] occurs before non-high vowels.

(6) Japanese Phonetic Representations[22]

[ɸ]	[t,d]	[s̠,z̠][24]
CONTINUANT	**STOP**	**CONTINUANT**
LABIAL	**CORONAL**[23]	**CORONAL**
MELLOW	MELLOW	MELLOW
LIP	ALVEOLAR	ALVEOLAR
	APICAL	LAMINAL

2.8. Theta-Eth

In (7) below, I give a phonetic representation of the English target interdental fricatives.

[22] In Japanese, [ts,dz] occur as allophones of /t,d/ before [ɯ]; also [ɕ,ʑ] occur as allophones of /s,z/ before [i]. These allophones are not included here because they were not contrasted with [θ,ð] in this study. For auditory distance calculations for these allophones, see Appendix A.

[23] Brown (1997, 1998) argues that JA does not have the feature CORONAL. However, Mester & Itô (1989) argue that the presence of CORONAL is necessary to account for palatalization in mimetic forms. Even if Brown's account is correct, at the phonetic level, I assume CORONAL is present in JA.

[24] The diacritic _ represents a lamino-alveolar articulation.

(7) Phonetic Representation of Theta and Eth

Intake Target {θ-ð}
MELLOW
CONTINUANT
CORONAL
DENTAL
LAMINAL

The above representations in (4)-(6) are fully specified phonetic representations. The main hypothesis of this study is that learners use phonetic, rather than strictly phonemic (underspecified) information when assessing potential interdental substitutes in naturalistic learning situations. Let us now turn to the predictions that this proposal makes regarding interdental substitution for the learner groups investigated in this study. First I will show how transfer based on contrastive features alone fails to account for differential substitution. Next, we will see that if we consider that non-contrastive phonetic features are also assessed in transfer, the correct results are generated.

2.9. Predictions

In this section, predictions are made for interdental substitutes using the auditory distance algorithm first applied to phonemic and then to phonetic representations.

2.9.1. Phonemic Predictions

Here I give the predictions based on a phonemic assessment, where only contrastive features are accessible for comparison. Recall that many accounts of segmental transfer assume that substitution occurs at the phonemic level of representation (e.g. Weinberger 1988; Hancin-Bhatt 1994a,b; Brown 1997,1998; Lombardi 2003).

The reader is also reminded that I consider phonological feature representations to be categorical. That is, feature weighting does not enter into phonological calculations; thus only the number of featural conflicts is assessed.

(8) Predictions for EF, QF, and JA Based on a Phonemic Assessment

Intake {θ,ð}	CONTINUANT CORONAL	
Potential Substitute	*Mismatches*	*# of Featural Conflicts*
☞ /s,z/		0
/t,d/	STOP	1
/f,v/	LAB	1

In (8), we see that the two features of /s,z/ that have phonemic status in EF, QF, and JA, i.e. CONTINUANT and CORONAL, both match the features on target [θ,ð]. Given that there are no featural mismatches, the distance of /s,z/ from [θ,ð] is zero. For /t,d/, there is a match on CORONAL and there is one featural conflict, STOP. For /f,v/, there is also one featural conflict: in this case, on the place feature, LABIAL. This evaluation, based on contrastive features, predicts /s,z/ to be the preferred substitute for EF, QF, and JA, and in fact, for the majority of languages. To my knowledge, /s,z/ has *never* been reported as a substitute for theta-eth in Québec French.

Even if we allowed enhancement to play a role at the phonemic level, the absence of non-contrastive features such as STRIDENT would lead to the same result, with all three languages showing a preference for the /s,z/ substitute. Thus, it is obvious that an assessment based on contrastive features alone fails to capture differential substitution. Let us now see how the algorithm applies to a phonetic level of representation for each of the learner languages investigated in this chapter.

2.9.2. Phonetic Predictions

Recall from §2.3.2 my hypothesis that features which serve to phonetically define L1 segments, even if non-contrastive, are perceived in the intake of a phonetic assessment. Only those features that are absent from both the phonology and phonetics of the L1 will not be perceived; thus the intake varies from language to language. The intake is compared with phonetic representations from the L1, which are fully specified for language-specific phonetic properties.

Since all the features of {θ,ð} serve to define segments of EF and QF, the surface intake for EF and QF is as represented in (9).

(9) Surface Intake for EF and QF

Intake			
{θ,ð}	*salient*	COR	+2
		CONT	+1
		MELL	+1
		DENT	+1
		LAM	+1

In (10) below we see that it is predicted that EF speakers will choose [s̪,z̪] as the substitute for the interdental. Recall from §2.5 that the feature STRIDENT is not enhanced by DENTAL (DENTAL MUTES STRIDENT). Since EF [s̪,z̪] are dental fricatives, no enhancement occurs, and these sounds are relatively close to the interdentals on the stridency dimension. The [s̪,z̪] reflex is what is most commonly reported in production studies of European French (Berger 1951, Wenk 1979, Brannen 1998). The next closest substitute is predicted to be [f,v], which also has been reported, but not as frequently substituted (Wenk 1979, Brannen 1998).

(10) Predictions for EF Based on a Phonetic Assessment

Intake			
{θ,ð}	*salient*	COR	+2
		CONT	+1
		MELL	+1
		DENT	+1
		LAM	+1

Potential Substitute	**Mismatches**			**Distance**	**Total Distance**
☞ [s̪,z̪]	*enhanced*	CONT	+2	1	
		STRID	-1	2	3
[f,v]	*salient*	LAB	-2	4	4
	enhanced	STOP	-2	3	
[t̪,d̪]		AP	-1	2	5

For QF, [t̪,d̪] is selected as the first candidate. This situation conforms to reports in the literature on production substitutes for Québec French (LaCharité & Prévost 1999, Brannen 2002). The next choice is [f,v].

(11) Predictions for QF Based on a Phonetic Assessment

Intake			
{θ,ð}	*salient*	COR	+2
		CONT	+1
		MELL	+1
		DENT	+1
		LAM	+1

Potential Substitute	**Mismatches**			**Distance**	**Total Distance**
☞ [t̪,d̪]	*enhanced*	STOP	-2	3	3
[f,v]	*salient*	LAB	-2	4	4
[s,z]	*enhanced*	CONT	+2	1	
		ALV	-1	2	
	enhanced	STRID	-2	3	6

In (12) below, I give the intake form for {θ,ð} that JA listeners are hypothesized to perceive. Since JA only has alveolar obstruents and no dental obstruents (Vance 1987), DENTAL is absent from even the phonetic component of the JA grammar. Thus the surface intake for {θ,ð} differs in JA compared to EF and QF. This is shown by the strikethrough of DENTAL. It is predicted that JA speakers will choose their lamino-alveolar, MELLOW [s̠,z̠] as a substitute. This is what is reported in the literature (e.g. Hancin-Bhatt 1994a,b; Brown 1997,1998). The bilabial fricative is predicted to come in second.[25]

[25] Actually [ɕ,ʑ] is predicted to be the second choice, but since this was not tested, it is not included here. However, see Appendix A for complete calculations.

(12) Predictions for JA Based on a Phonetic Assessment

Intake {θ-ð}	*salient*		
		COR	+2
		CONT	+1
		MELL	+1
		~~DENT +1~~	
		LAM	+1

Potential Substitute	**Mismatches**			**Distance**	**Total Distance**
☞ [s̱,ẕ]				0	0
[ɸ]	*salient*	LAB	-2	4	4
[t,d]	*enhanced*	STOP	-2	3	
		AP	-1	2	5

In summary, we have seen that in EF, the dental fricative [s̪,z̪] is the best perceptual substitute for theta-eth since it matches the target in all respects except on the stridency dimension; but even there, the discrepancy is minimal because stridency is muted at the dental location. For QF, it is the dental stop which is predicted to be the closest substitute; while in JA, it is the alveolar fricative [s̱,ẕ] by virtue of it being specified as a mellow continuant in Japanese.

2.10. Summary

In this chapter, I have proposed that language-specific articulatory variations in the production of obstruents and their related language-specific phonetic representations govern cross-language differences in the way ESL learners perceive and produce the interdental fricatives. In particular, I have argued that certain featural combinations can result in a particular feature being enhanced; thereby making it more perceptually salient. Thus, the solution to the differential substitution problem relies not only on an evaluation of phonologically contrastive features, but crucially implicates non-contrastive phonetic features as well. It is thereby predicted that the differential substitution patterns which correspond to what has been observed for various L1s will emerge in phonetic processing, but not in phonological processing.

I have presented the Auditory Distance Model (ADM), which depicts speech processing both from a perceptual and articulatory point of view. The ADM shows the inter-relation between the acoustic signal, phonetic representations, and phonological representations, and delineates between phonetic and phonological processing. The model includes an algorithm which explains how a non-native sound is mapped to a native representation. It was shown that an assessment based on contrastive phonological features alone failed to account for differential substitution; on the other hand, if non-contrastive phonetic features along with feature enhancement are considered, the correct results are obtained: European French listeners associate the interdental fricatives of English with their native [s̪,z̪]; Québec French auditors associate them with [t̪,d̪]; and Japanese, with [s̠,z̠].

The ADM now needs to be tested. It is necessary to experimentally investigate the issue of whether substitutions are phonetically or phonologically based. Chapter 3 presents the first in a series of three perception experiments which empirically test the hypotheses presented in the current chapter. This first study tests how English learners from European French, Québec French, and Japanese L1 backgrounds perceive pairs of sounds, some of which differ in contrastive features, and others which differ in non-contrastive features. I investigate cross-linguistic differences in auditory distance with a task whose parameters are modified first to tap phonetic processing, then to elicit phonemic processing. It is predicted that the language-specific associations noted above will emerge during phonetic processing, but not during phonemic processing, thus supporting the hypothesis that differential substitution occurs because listeners access non-contrastive and language-specific phonetic detail.

CHAPTER 3: AXB-1

In this chapter, I report on two experimental tasks designed to test how learners of English perceive various consonantal contrasts.[26] Recall the major hypotheses behind this research:

1. Differential substitution is due to transfer from the L1;
2. Transfer in production is caused by transfer in perception;
3. Choice of substitute depends on a comparison of the phonetic properties of the target segment with phonetic properties of segments in the L1 sound system;
4. Transfer of non-contrastive, phonetic features is involved in substitution.

The learners investigated in these studies are from European French, Québec French, and Japanese L1 backgrounds, and all are at an advanced level of English proficiency. One task is intended to induce phonemic processing, and the other phonetic processing. In the phonemic test, when only contrastive feature categories are available, it is hypothesized that all three learner languages investigated will have a tendency to misperceive [θ,ð] as [s,z]. On the other hand, in the phonetic test, non-contrastive phonetic features are also available, and it is hypothesized that in this task, European French and Japanese listeners will misperceive [θ,ð] as [s,z], but Québec French listeners will misperceive [θ,ð] as [t,d] (see Chapter 2, §2.9). It is proposed that the latter reflects naturalistic learning situations: learners perceive non-contrastive phonetic information in the input; hence differential substitution emerges due to phonetic differences in the various L1 consonant inventories.

---

[26] AXB-1 was conducted as a pilot study, and a previous version was published as: Brannen, Kathleen (2002) The role of perception in differential substitution. *Canadian Journal of Linguistics* 47(1):1-46.

If it can be found that learners use non-contrastive information when attending to their L2, this will have implications for how learners treat other L2 sounds and may help explain some substitutions and approximations that do not have any phonological solution.

3.1. Experimental Design

This section describes the design of the experimental portion of this study.

3.1.1. Phonetic versus Phonological Processing

Research has shown that different components of the grammar can be accessed depending on the experimental task employed. Werker & Logan (1985) propose that speech can be processed at three different levels, depending on the interstimulus interval (ISI) used in an experimental task:

(1) Three Levels in Speech Perception (Werker & Logan 1985)

Acoustic:	Processing of fine, non-linguistic distinctions; i.e. listeners assess physical identity, for example, fundamental frequency, amplitude.
Phonetic:	Processing of linguistically relevant information only, both contrastive and non-contrastive; normalization of non-linguistic differences.
Phonemic:	Processing of contrastive information only; normalization of non-linguistic and non-distinctive information.

The levels which are of relevance to this chapter are the phonetic and phonemic levels. Werker & Tees (1984) have demonstrated processing at both these levels. Using a category change procedure, adult monolingual English speakers failed to discriminate contrasts which are not phonemic in English.[27] Specifically, they failed to hear the difference between Hindi dental and retroflex stops and between Thompson Salish velar and uvular stops. However, in an AX procedure, English participants *were* able to hear the difference between these contrasts. Werker and Tees suggest that these results present evidence of phonemic vs. phonetic processing respectively.[28]

In a within-subjects design, Werker & Logan (1985) found phonetic processing at a 250msec ISI and, in a between-subjects design, evidence of phonemic processing at 1500msec ISI.[29] In accordance with Werker & Logan, the present study tests phonetic processing with a Short ISI and phonemic processing with a Long ISI. With a long interval between two stimuli, by the time a listener hears the second stimulus, the information present in the acoustic signal of the first stimulus has faded. It is commonly held that what fades are the non-distinctive, phonetic features. On the contrary, in the Short ISI condition, the interval between stimuli is short enough for information on both contrastive features and non-contrastive, phonetic features to be preserved. These are the intervals I use to evoke phonetic and phonological processing respectively.

---

[27] In the category-change procedure, a given stimulus is repeatedly presented at fixed intervals. At a specific point during the presentation, a different stimulus is introduced. Afterwards, presentation of the original stimulus is resumed. The participant's task is to press a button whenever they detect the change in stimulus.

[28] Note that Werker & Tees do not address the issue of cross-linguistic differences within the phonetic task. Their experiments compared Native English speakers with Hindi speakers on Hindi contrasts, but did not include any other language with which to make a comparison.

[29] In a within-subjects design, all experimental conditions are experienced by all participants. This contrasts with a between-subjects design, where, for example, one group of participants experiences Condition 1; whereas, another group of participants experiences Condition 2.

3.1.2. Participants

The present experiment involved four language groups. There were three groups of learners of English as a second language: European French (EF), Québec French (QF), and Japanese (JA). In addition, there was a control group of Native North American English speakers: Native English (NE). The non-native speakers were students from Montréal universities. They were all in their 20s or 30s. Native English speakers were all students from McGill University; three were in their 20s and one in her 40s. Each group consisted of five participants, for a total of 20 subjects.[30]

All learners in this study began learning English after the age of seven, the purported critical period for phonological acquisition (e.g. Scovel 1988); thus they are not considered to have acquired both languages simultaneously in early childhood and, as such, we are in fact examining L2 acquisition. All non-native participants were classed as advanced learners of English based on the aural comprehension component of the standardized Michigan English Placement Test.[31] All participants had normal hearing according to self-report.

[30] One Native English participant was left out of the analysis because she had a significantly higher error rate than the other participants. In the final analysis then, there were only four participants in the NE group, for a total of 19 participants.

[31] One JA speaker scored as high intermediate on the proficiency test, but she was otherwise very fluent. Her results did not differ from the other JA speakers. Although all speakers were classed as advanced, with the exception of one QF speaker, none could be considered as indistinguishable from a native speaker of English. Their scores on the Michigan test were not perfect, and their production was accented, with instances of interdental substitution, and included syntactic and lexical errors as well.

3.1.3. Test Design

3.1.3.1. Tasks

The experimental paradigm used was an AXB forced choice task (e.g. Best and Strange 1992). Each item was a triad consisting of three stimuli: two non-identical tokens of the same type and one token of another type, as exemplified in (2):

(2) Item Example

thigh$_1$ thigh$_2$ tie
A X B

The participant's task was to determine whether A=X or whether B=X. If A=X, the subjects were to press a key labelled A on the left side of the keyboard. If B=X, the subjects were to press a key labelled B on the right side of the keyboard.

The AXB task was chosen because it is thought to be less cognitively demanding than the more common ABX paradigm; for the latter, one must retain A in memory in order to compare it with X (Beddor & Gottfried 1995). In the AXB task, this does not occur because the target is in the middle. Beddor & Gottfried also claim that the AXB task has a lower sensitivity to response bias as compared to an AX task. In other words, in the AX task, participants might tend to respond with either all the same or all different. Also, with an AX task, if a participant responds that the two stimuli are different, we have no way of knowing on what basis this judgment was made. For example, it could be based on non-linguistic cues such as differences in timbre, not a difference in phonetic or phonological features. With an AXB task, this pitfall is largely avoided.

Finally, the AXB task is shorter than the 4IAX task. In the 4IAX paradigm, listeners hear four pairs of stimuli for each test item, viz. AB-AA, BA-AA, AA-AB, or AA-BA, where the hyphen represents a period of silence. The listener has to decide which pair contain identical stimuli, e.g. in AB-AA, either AB or AA. The AXB task only requires three stimuli per item, while retaining the advantages of the 4IAX (e.g. low response bias). Considering the large number of contrasts examined in the present study, this was a necessary consideration.

Participants heard a total of 2430 triads over five (non-consecutive) days. The same test was administered twice, once with a Long ISI between stimuli and again with a Short ISI.

(3) Task Outline

Long ISI	*Short ISI*
Spanned three non-consecutive days	Spanned two non-consecutive days
405 items per day	675 items on the first day; 540 items on the second day
1500msec interstimulus interval	250msec interstimulus interval
Intended to evoke phonological processing	Intended to evoke phonetic processing

The Long ISI condition was presented in its entirety before the Short ISI condition. The reason for this was that I wanted to ensure phonological processing in the Long ISI condition. Werker & Logan (1985) found that in an AX task, participants could not easily switch from one processing mode to another. Thus I reasoned that if 250msec promotes a phonetic mode of processing, then I would not observe phonological processing in the Long ISI if it followed the Short ISI condition. Based on Werker & Logan's findings, I further reasoned that it would be easier to switch from a phonological to phonetic strategy, than from a phonetic to phonological strategy. Werker & Logan had difficulty tapping phonological processing. In a within-subjects design, they report that they did not tap phonological processing at all in the 1500 msec condition. In a between-subjects design, they did manage to tap phonological processing, but only in the first two blocks of the 1500msec condition. In the remaining three blocks, participants switched to a phonetic mode. These results suggest that phonological processing is best induced when the participant is given a Long ISI condition from the outset. In other words, practice may promote phonetic processing.

In the present study, participants were tested on a Macintosh PowerBook using Grado SR60 headphones. Each day, they began with four practice items. Then they were presented blocks of 135 items. Items were randomly ordered within each block and across participants. For the Long ISI condition, these blocks were about 15 minutes long, three blocks each day for Days 1, 2, and 3. For the Short ISI, they were about 10 minutes long, five blocks on Day 4 and four blocks on the last day. Between blocks, participants had a five-minute break in order to compensate for fatigue and adaptation effects. In sum, participants were tested for approximately one hour each day.

Participants were told that they would hear three English words. One reason that I told them they were listening to English words was to bias them towards using their L2 English grammar rather than their L1 grammar. This assumes that grammars are separate/modular (e.g. Bialystok & Cummins 1991, Paradis 2004, Sharwood-Smith 2004). Another reason for telling them they were listening to English *words* was to prime them to listen in a linguistic mode rather than general auditory mode. Again, this assumes modularity — in this case, of cognitive functions.

Participants were instructed to respond as quickly and as accurately as possible. Reaction times and accuracy were recorded. However, only accuracy is reported in this study.[32]

3.1.3.2. Stimuli

In (4) below, I give a list of test contrasts presented in the experiment. The first column gives the feature or features which serve to contrast the two segments. The segments are listed in the second column. The next three columns show the number of tokens per contrast for each vowel context (A = low vowels; I = front (non-low) vowels; U = back (non-low) vowels). The final column gives the total number of tokens for each contrast.

Filler items were included in the experiment at a ratio of approximately 1:2 with the test items; the Onsets of the fillers did not include any target segments. Fillers were added to provide distracters in order that the participants not become cognizant of the purpose of the experiment. Only test items were analyzed. All target segments appeared in simple Onset position of monosyllables.

The table in (4) shows that five of the contrasts tested differed by a single feature (#1-4, #7). By manipulating one feature, we avoid introducing confounding factors. This enables us to have a clearer picture of just what featural differences learners can perceive. The other two contrasts compare more than one feature (#5 and #6). These were included for a couple of reasons. First, #5 and #6 both investigate whether listeners can perceive apico-alveolars from the lamino-dentals. Item #5 is

[32] This is because of a fault in the methodology. Participants had the option of responding as soon as they heard X or waiting until they heard B. As a result, it was found that different subjects employed different strategies: some were "impulsive", responding during X; others were "conservative", waiting until all three stimuli had been presented. This inconsistency in response strategy made it impossible to compare reaction times across (and even within) subjects. Van Hessen & Schouten (1999) also encountered this problem with the AXB design (see Gerrits 2001:42).

intended to be compared with Item #3. In #5, the comparison is essentially between an "English" [t,d] and [θ,ð], and in #3 between a "French" [t̪,d̪] and [θ,ð].[33] A comparison of the results on #3 versus #5 should show whether listeners perceive [θ,ð] as being closer to lamino-dental [t̪,d̪] than to apico-alveolar [t,d]. It is predicted that this will be the case for the French participants, since they have a dental stop; however, the opposite order is predicted for Japanese and English participants, since they have an alveolar stop. This will be elaborated upon later in this chapter in the context of Best’s Perceptual Assimilation Model (Best 1995). Contrast #6 also tests the apico-alveolar versus lamino-dental dimension, but this time the comparison is between two stops. This tests whether listeners can perceive the difference between an “English” apico-alveolar stop and a “French” lamino-dental stop. This contrast is predicted to be difficult for all language groups; however, it may be more difficult for the EF, because both stimuli are equi-distant from their apico-dental [t̺,d̺]. Again this will be elaborated upon later on.

[33] Apico-dental stops (as in EF see Chapter 2, §2.7.1) were not contrasted with the interdental fricatives. These sounds are intermediate between lamino-dental and apico-alveolar stops. It was decided that if results showed differences between the latter two, then the apico-dental stop vs. interdental contrast would be included in a subsequent experiment.

(4) Test Contrasts[34]

TEST ITEMS					
Distinguishing Feature(s)	Segments	Number of tokens per ISI by vowel			
		A	I	U	Total tokens
1. LIP vs. DENTAL	[f,v] vs. [ɸ,β]	9	9	6	24
2. LABIAL vs. CORONAL	[f,v] vs. [θ,ð]	9	6	15	30
3. STOP vs. CONTINUANT	[t̪,d̪] vs. [θ,ð]	18	12	6	36
4. STRIDENT vs. MELLOW	[s̪,z̪] vs. [θ,ð]	21	18	9	48
5. STOP vs. CONTINUANT and ALVEOLAR vs. DENTAL APICAL vs. LAMINAL	[t,d] vs. [θ,ð]	27	30	6	63
6. ALVEOLAR vs. DENTAL and APICAL vs. LAMINAL	[t,d] vs. [t̪,d̪]	30	27	21	78
7. MELLOW vs. STRIDENT stop	[t̪,d̪] vs. [t̪s̪,d̪z̪]	12	15	6	33

3.1.3.3. Stimuli Quality

All stimuli respect English syllable structure and phonotactics. Clearly, however, some of the test contrasts involve non-English segments. This was necessary in order to test whether participants were sensitive to certain featural distinctions that might play a role in differential substitution. Both

[34] The token number is unequal across test items because of the need to use only real words of high-frequency.

words and non-words were used in the experiment. The real words incorporated in the study are all high frequency as rated by seven native English speakers on a 5-point scale.[35] Non-words were checked by three linguists. They were rejected if they were homonyms, or close to homonyms, with a real English word or if they were similar to a socially unacceptable real word.[36]

Because of the limited number of participants in this study, several tokens of each type were constructed in order to increase the power of the experiment. Three factors were analyzed: Contrast, Voicing, and Vowel.

Contrast

As seen in (4), seven contrasts were tested in this experiment: [ɸ,β f,v θ,ð t̪,d̪ s̪,z̪ t,d t̪s̪,d̪z̪]. All targets, including fillers, were located in onset position. All onsets were simple.

The rhyme was always heavy, respecting word minimality. Thus the syllable was either open with a diphthong (vowel + glide); or closed by a consonant or consonant cluster. The codas used were [p k m n]. The consonants [m n] were selected because nasals are favoured cross-linguistically over other segments in coda position due to their relatively high sonority (e.g. Venneman 1988, Clements 1992). The consonants [p k] were chosen because they have little distorting effect on the quality of the preceding vowel and are relatively unmarked in terms of structural

[35] Word frequency survey participants were asked how often they thought they had heard the word spoken and seen it written. I did not consult standard written frequency counts (e.g. Kučera & Francis 1967) because I am primarily interested in aural frequency.

[36] Cross-language homonyms were not controlled for, e.g. *bow* and *beau*. One participant remarked on hearing words that sounded French; thus future research should take these into account.

complexity and frequency in this position.[37] For non-words, only simple codas were used; however, in word contrasts, it was necessary to allow for branching codas. Words beginning with [θ-ð] are relatively rare in English. If I did not allow for complex codas in real words, the number of real word stimuli would have been very limited.[38]

Voicing

Both voiceless and voiced stimuli were examined. Voicing was included as a factor because some researchers have found differences in the perception of interdental fricatives on this measure. For example, Hancin-Bhatt (1994a,b) found that in onset position, the voiceless interdental was better perceived than the voiced interdental. As well, she found that the voiced interdental was more likely to be perceived as a stop. Therefore, based on Hancin-Bhatt's results, it is predicted that there will be fewer errors for the voiceless contrasts involving the interdentals than for the voiced contrasts. In addition, we might expect more errors on [d-ð],[d̪-ð] compared to [z-ð],[z̪-ð].

Vowel

This factor was included because several studies on fricatives have suggested that F2 formant transitions on the vowel may affect the perception of obstruents, including fricative perception (e.g. Johnson &

---

[37] The alveolar stop was not used because it could introduce an Obligatory Contour Principle (OCP) effect in some of the test items. The OCP is a constraint which militates against (near) identical segments within a given domain (McCarthy 1988).

[38] Statistical analysis for wordhood is not presented in this chapter. Refer to Chapter 4 (AXB-2) for analysis and discussion of the effect of wordhood on perception.

Babel 2007). Of particular interest, Shadle et al. (1996, cited in Tabain 1998:109) showed that the quality of non-strident fricatives is affected by the backness/roundness of the following vowel. Thus, vowel quality was tested in order to see if it has an influence on the perception of word-initial contrasts.

The target stimuli occurred before one of three types of vowels: FRONT/NON-LOW (I) -- [i j ɪ ej ɛ]; BACK/NON-LOW (U) -- [uw ow]; LOW (A) -- [æ ʌ ɑ aj aw].[39] The mid central vowel [ʌ] was included with the low vowels because many L2 learners confuse it with [ɑ]; also it is unrounded, while other back non-low vowels are round. The diphthongs [aj aw] were classified as LOW according to the category of the head. As much as possible, true diphthongs were avoided. However, they were used in a handful of real words to increase the number of stimuli. Diphthongs introduce a potential confounding factor, given that they involve several features, for example, LOW plus FRONT and HIGH in [aj].

3.1.3.4. Quantity of Stimuli

Two stimuli are considered to be of the same type if they agree on all factors named above. For example, *teak, team, teen, tip, tick, Tim, tin, tape* are all of the same type: they are all real words that share voiceless [t] in onset, occur before a front vowel (I), and are in a closed syllable (C). Each type had 6–8 tokens. These tokens were either the same ‘word’ (existing or nonce) or different ‘words’. Note that multiple tokens of the same word were not physically identical (see §3.1.3.5). The number of tokens constructed depended on the contrast in which they appeared.

---

[39] The high back round lax vowel, [ʊ], was not used because it does not have a high type frequency in English.

Different tokens of the same type. The set [*teak, team, teen, tip, tick, Tim, tin, tape*] consists of eight non-identical tokens.

Same tokens of the same type. The construction of different tokens was often inhibited by the non-existence or low frequency of a real word for one of the members of the pair. For example, for the contrast [θ] vs. [t] in open syllable before a LOW vowel, only one token could be constructed, *thigh* vs. *tie*. The word *thaw* has a counterpart, *taw*, but it is of low frequency and probably unfamiliar to participants. So *thigh* vs. *tie* was repeated six times. However, the same token was not repeated more than six times. This was done in order to avoid repetition effects. Where more tokens could be constructed, these were repeated to a maximum of eight tokens per type.

3.1.3.5. Recording of Stimuli

Stimuli from three trained phonologists (talkers) were recorded in a professional sound studio using a Tascan DA 30 DAT recorder and AKG 414 EB microphone with cardiod pattern setting and pop filter. [40]Natural rather than synthetic stimuli were chosen in order to promote a linguistic rather than general auditory mode of listening. Items were recorded in two carrier phrases: *I learn* ___ and *You hear* ___. This was done in order to ensure that the target seem as natural as possible for the talkers; that is, to avoid a "list effect".

All talkers produced all stimuli. Each talker produced each stimulus twice. This was done in order to have two versions of the same token which were not physically identical; triads were composed of two different

[40] Les productions DNA inc., 4200 Boul. Saint-Laurent, Bureau 409. ((514) 842-5491). Montréal, Québec.

tokens of the same type and another token of a different type. Physically different versions of the same token were used in order to promote phonetic or phonological as opposed to acoustic processing. Recall from §3.1.1 that the acoustic mode involves processing of non-linguistic distinctions; the listeners assess physical identity, for example, fundamental frequency, or amplitude. This is the type of processing we use in discriminating two bell tones, for example. If listeners were making decisions on an acoustic basis, this would not inform us about speech perception; thus, it was essential that we tap a linguistic mode of processing. The same talker's voice was used for all three words within each triad; however, different talkers were used across triads. The data were transferred into SoundEdit 16.01 at a sampling rate of 22kHz. The experimental paradigm was constructed using PsyScope 1.2 (Cohen et al. 1993).

3.2. Phonemic vs. Phonetic Feature Perception

In this section I give my predictions as to the perceptibility of each contrast. The purpose of the AXB task is to determine which featural contrasts listeners can perceive in each processing mode, phonological and phonetic.

Before proceeding, it is necessary to first describe what listeners are reckoned to do when completing an AXB task. I assume that during an AXB task, the listener assesses each sound in terms of its proximity to a native category, be it either phonetic or phonological. Thus, each sound is funneled into its closest L1 category. If both sounds are funneled into the same L1 category, then discrimination should be rather difficult. However, the degree of difficulty will depend on how well each sound fits into the category. If, on the other hand, each sound is funneled into a different L1 category, discrimination should be relatively easy. This interpretation is in essence equivalent to Best's Perceptual Assimilation Model (e.g. Best, McRoberts, & Sithole 1988, Best 1995).

The Perceptual Assimilation Model (PAM) suggests a number of ways in which pairs of (non-native) phones may be perceived in relation to each other. The three types of assimilation of relevance to the present work are:

1. *Two-Category Assimilation* (TC), where two phones are perceived as exemplars of two separate native categories. In this case, discrimination will be good.
2. *Category-Goodness Difference* (CG), where two phones are perceived as exemplars of one native category but are not equally good examples of this category. Here, discrimination should be moderate to good, depending on the degree of perceived differences in goodness of fit.
3. *Single-Category Assimilation* (SC), where two phones are again perceived as exemplars of one native category but are perceived as equally good (or bad) examples of that category. Discrimination is predicted to be poor.[41]

Figure 3.1 below gives a graphic explanation of these three types of perceptual assimilation:

[41] There are three other types of assimilation described in Best's model. The studies in this thesis do not provide a way of distinguishing these types from the types listed above; thus they are not included. These other three are: *Uncategorisable* (UU), where two phones both fall in between two or more phonetic categories of the native language. *Uncategorisable versus Categorisable* (UC), where one non-native phone is perceived as an exemplar of a native sound category and the other falls in between two or more native phonetic categories. *Non-assimilable* (NA), where both non-native speech sounds fall outside of the native speech domain altogether and are perceived as non-speech sounds.

Two-Category Assimilation			*Category-Goodness Difference*			*Single-Category Assimilation*		
Intake	{x}	{y}	Intake	{x}	{z}	Intake	{x}	{w}
	↓	↓		↓	↓		↓	↓
Distance	2	3	Distance	2	0	Distance	2	2
	↓	↓		↘	↙		↘	↙
L1 categories	[a]	[b]	L1 category	[a]		L1 category	[a]	

Figure 3.1. Auditory Distance and the Perceptual Assimilation Model

To exemplify the application of the PAM in concert with the ADM outlined in Chapter 2, let us take the case of EF listeners on the [f] vs. [ɸ] contrast, assuming that when listeners perform the AXB task, what they are doing is funneling each sound into an L1 category (either phonetic or phonological). During the funneling process, auditory distance is measured between the intake and each L1 category. So, in the case of phonetic processing for EF, intake {f} is funneled into the closest EF phonetic category, which is [f]; since the intake and L1 category are a perfect match, the distance is 0. Likewise, intake {ɸ} is funneled into the [f] phonetic category. This time the closest category is not a perfect match: the two sounds differ on the basis of LIP vs. DENTAL. Here the distance is 2. Since both [f] and [ɸ] are funneled into the same phonetic category, yet the goodness of fit into this category differs for each, this constitutes a *Category-Goodness Difference* type, and perception is predicted to be moderate: not as good as a *Two-Category Assimilation*, not as poor as a *Single-Category Assimilation*. For exact calculations of auditory distance within the context of the PAM for all contrasts and language groups, see Appendix B.

Let us now turn to the contrasts investigated in this study and the predictions PAM makes for each language group.

1. [f,v] vs. [ɸ,β]

The features involved in this contrast — LIP vs. DENTAL — do not serve to make a phonological distinction in any of the language groups tested in this study. It is predicted that no group will distinguish these sounds in the Long ISI (phonological) condition, and that all groups will have equal difficulty.

In phonetic processing, European French (EF), Québec French (QF), and Native English (NE) listeners should have less difficulty than the Japanese (JA) group on the [f,v] vs. [ɸ,β] contrast. This is because EF, QF, and NE all have the feature DENTAL in their phonetic inventories, so [ɸ,β] represents a *Category-Goodness Difference* by being a relatively poor exemplar of [f,v]. In JA, there is no [f,v] and all coronal segments are alveolar (Vance 2008). The phonetic feature DENTAL is thus absent in JA, and there is no conflict between LIP and DENTAL; therefore, JA listeners should merge [f] with [ɸ], even at the phonetic level (see §2.9.2).[42,43]

2. [f,v] vs. [θ,ð]

The features involved in this contrast — LABIAL vs. CORONAL — are distinctive for all languages. Thus it is predicted that participants will be able to discriminate this contrast in both conditions, since each member is funneled into a separate L1 category, a *Two Category Assimilation* (TC).

[42] Because the feature DENTAL is absent from both the phonetic and phonological inventories in Japanese, it is absent from the intake, and the feature LIP will not constitute a mismatch.

[43] Since JA does not have the voiced version of [ɸ], i.e. [β], it is predicted that Japanese speakers will merge [v] with [b].

3. [t̪,d̪] vs. [θ,ð]

The features contrasting these segments — STOP vs. CONTINUANT — are contrastive in all languages tested.[44] In the phonological condition, all language groups should easily perceive this contrast as it will fall under the TC type. In the phonetic condition, this contrast will also be a TC assimilation for EF, JA, and NE. However, QF should show a poorer discrimination rate since this will be a *Category-Goodness Assimilation*, with [θ,ð] being perceived as a poor exemplar of [t̪,d̪], since [θ,ð] will be associated with the closest QF phonetic representation, which is [t̪-d̪] according to the auditory distance algorithm.

4. [s̪,z̪] vs. [θ,ð]

The features of interest here — STRIDENT vs. MELLOW — are not contrastive for any of the non-native speakers. Thus, they should not be able to hear the difference in the phonological condition: each sound should be funneled into the same L1 category (*Single-Category Assimilation*), unlike for the NE listeners, who should easily perceive this difference, it being a TC assimilation. At the phonetic level, however, EF and JA should have more difficulty than QF and NE. For the former language groups, this would represent a *Category-Goodness Assimilation*; whereas, for the latter groups, it would be a TC assimilation.

---

[44] [t̪,d̪] was produced by the talkers as LAMINAL; thus both the [t̪,d̪] and [θ,ð] stimuli share this feature.

5. [t,d] vs. [θ,ð]

This contrast is related to the contrast in #3 above, [t̪,d̪] vs. [θ,ð]. However, in addition to testing STOP vs. CONTINUANT, it also examines ALVEOLAR vs. DENTAL and APICAL vs. LAMINAL. In the Long ISI, listeners should easily perceive this contrast based on STOP vs. CONTINUANT, which puts this contrast in the *Two-Category Assimilation* type for all language groups. In the Short ISI, discriminability should be further enhanced by the availability of the features ALVEOLAR vs. DENTAL and/or APICAL vs. LAMINAL. As with the contrast [t̪,d̪] vs. [θ,ð] in #3 above, it is expected that EF, JA, and NE, with TC assimilation, will out-perform QF, with a CG assimilation type. However, it is predicted that QF will show a somewhat better discrimination on this contrast involving apico-alveolar [t,d], than on the contrast in #3 involving lamino-dental [t̪,d̪]. This is because the latter is a perfect match to the QF coronal stop; whereas, the former is not. For JA and NE, it is the opposite.

6. [t,d] vs. [t̪,d̪]

The features involved in this contrast —ALVEOLAR vs. DENTAL and APICAL vs. LAMINAL — are not distinctive in any of the languages investigated. It is predicted that this contrast will not be perceived at the phonological level: all language groups should group these into a *Single-Category Assimilation* (SC). At the phonetic level, we should expect poorer performance for EF, for again this would be a SC type, given that both members of this pair are equidistant from the native EF apico-dental [t̺,d̺]. Conversely, for the other language groups at the phonetic level, this is a *Category-Goodness Assimilation*, since one of the members is a better exemplar of the native phonetic category than the other; hence we should expect somewhat better performance than for EF.

7. [t̪,d̪] vs. [t̪s̪,d̪z̪]

As in #4 above, the features distinguishing this pair are STRIDENT vs. MELLOW, but in this case these features are differentiating stop consonants rather than fricatives. In the phonological condition, all language groups are predicted to merge these two sounds into one category, SC assimilation; thus discrimination should be relatively poor. On the other hand, in the phonetic condition, all language groups are predicted to perceive this distinction, since it is predicted to be a *Two-Category Assimilation.*

In (5) below, I give the relative order of difficulty predicted by the auditory distance algorithm in conjunction with PAM, both in the phonological and phonetic conditions, from most errors to least errors.

(5) Predicted Order of Difficulty according to Auditory Distance Calculations and Perceptual Assimilation Model[45]

Rank	Phonological Processing				Phonetic Processing			
	EF	**QF**	**JA**	**NE**	**EF**	**QF**	**JA**	**NE**
Most Difficult (SC)	[f-ɸ] [s̪-θ] [t-t̪] [t̪-t̪s̪]	[f-ɸ] [s̪-θ] [t-t̪] [t̪-t̪s̪]	[f-ɸ] [s̪-θ] [t-t̪] [t̪-t̪s̪]	[f-ɸ] [t-t̪]	[t-t̪]		[f-ɸ]	
Moderately Difficult (CG)				[t̪-t̪s̪]	[f-ɸ] [s̪-θ]	[f-ɸ] [t-t̪] [t̪-θ] [t-θ]	[s̪-θ] [t-t̪]	[f-ɸ] [t-t̪] [t̪-t̪s̪]
Least Difficult (TC)	[f-θ] [t̪-θ] [t-θ]	[f-θ] [t̪-θ] [t-θ]	[f-θ] [t̪-θ] [t-θ]	[f-θ] [s̪-θ] [t̪-θ] [t-θ]	[f-θ] [t̪-θ] [t-θ] [t̪-t̪s̪]	[f-θ] [s̪-θ] [t̪-t̪s̪]	[f-θ] [t-θ] [t̪-θ] [t̪-t̪s̪]	[f-θ] [s̪-θ] [t̪-θ] [t-θ]

This section has given the predictions for each contrast without considering vowel context or voicing. We now turn to examine how the latter factors may influence the perception of the contrasts under investigation.

3.2.1. Context Effects

Research has shown that vowel context has an influence on consonant perception (e.g. Mann & Repp 1980, Ohala & Solé 2010). Therefore, we may expect L2 learners to have more difficulty in discriminating contrasts in certain vowel contexts. High vowels might be one such case. There are common diachronic and synchronic processes in which stops are assibilated

[45] For ease of exposition, only voiceless contrasts are presented.

to either fricatives or affricates before high vowels, and many of these processes may have perceptual underpinnings (Hall & Hamann 2003). Thus, we might expect L2 learners to be better at distinguishing between [t̪,d̪] and [t̪s̪,d̪z̪] before low vowels than before high vowels. Furthermore, this may be especially so for those groups for whom affrication exists as an allophonic process in the L1, viz. Québec French and Japanese (Muñoz-Sánchez 2003).

Shadle et al. (1996) (cited in Tabain 1998:109) suggest that non-sibilant fricatives are more subject to influence from the following vowel than are sibilant fricatives. In particular, the lower frequency ranges of the non-sibilant fricatives vary with vowel context. A spectral peak corresponding to F2 was found to be higher before the vowel [i] and lower before [u]. Furthermore, this effect was greater for [f] than for [θ].

On the other hand, Mann and Repp (1980) found an influence of following vowel on sibilant fricatives. Specifically, they found that in a [s-ʃ] series, native English listeners tended to interpret [ʃu] as [su]. The explanation is that the following [u] has the effect of lowering F2 on the preceding fricative; thus, listeners interpret the [ʃ] as being a backed [s].

If we integrate the information gleaned from both Shadle et al. and Mann & Repp, it is predicted that listeners in the present study would tend to interpret [ɸ,β] in the context of a NON-LOW FRONT vowel as a fronted [f,v]. However, in the context of a back rounded vowel, the inverse is not predicted to occur, since rounding extends the cavity in front of the constriction for [f,v], but not for [ɸ,β]. Therefore, this contrast should actually be enhanced in the context of a rounded vowel.

Regarding the fricative pairs involving interdentals, [f,v] vs. [θ,ð] and [s̪,z̪] vs. [θ,ð], it is predicted that these pairs will be more difficult to discriminate in the context of a low vowel. This is based on information from two studies: Soli (1981) found that fricatives before [a] have spectral characteristics of fricatives in isolation, and Jongman (1989) found that interdentals were poorly identified in isolation.

Not only might we expect vowel context to have an influence on perception of the contrasts in this study, but voicing condition should also affect the results. Vocal cord vibration in voiced fricatives limits the degree of oral pressure. High oral pressure is required to maintain the turbulence amplitude in strident fricatives (Cho & Giavazzi 2009). Thus, it is predicted that the contrast [z-ð] will be more difficult to distinguish than its voiceless counterpart due to reduced amplitude on [z], making it more similar to [ð] in terms of stridency.

Having reviewed the predictions made by the Auditory Distance Model and the Perceptual Assimilation Model, and of the effects of vowel and voicing conditions, let us now turn to the results from the AXB-1 study.

3.3. Results

This section reports the results from the AXB-1 task for both ISI conditions. Before we proceed, a comment is in order concerning the statistical analyses presented throughout the thesis. The reader will note that both (parametric) Anovas and non-parametric statistics are provided. The non-parametric statistics are provided because the data violate "homogeneity of variance", one of the assumptions of the Anova. Although it has been shown that the F statistic of the Anova is quite robust against violations of homogeneity of variance (Lindman 1974), it was decided to incorporate non-parametric tests which do not make such an assumption. Results will be reported as significant only when they are so in both tests. For the Anova, significance levels were set at $p = .05$. Since the non-parametric tests involve computing several small tests, similar to running multiple t-tests, the risk of Type I error increases; therefore, it was decided to set a stricter significance level for the non-parametrics, $p = .005$.[46,47]

Let us now continue with the results. The analyses presented in the following sections are based on each participant's error rate, which was calculated for each contrast.

3.3.1. Main Effects

This section reports the results for each factor individually, collapsing the effects of other factors. The factors that reached statistical significance are listed individually below. Throughout this dissertation, only significant results are reported.

Voicing: Contrasts involving voiceless pairs were compared with voiced pairs. Overall, there were significantly more errors in the Voiceless condition (12%) than in the Voiced condition (10%) [$X^2(1), p = .003$] (Anova, $F(1,15), p < .01$). This indicates that, all else being equal, voiceless contrasts are more difficult to perceive than voiced contrasts. However, as we shall see shortly in the section on interactions, the source of this difference appears to be restricted to two contrasts in particular.

Vowel: The nature of the following vowel also had significant influence on the perceptibility of contrasts [$X^2(2)$, $p < .001$] (Anova, $F(2,15)$, $p < .01$). The main effect means that contrasts are more difficult to perceive in the context of certain vowels if we collapse language groups, contrasts, and ISI.

---

[46] Type I error occurs when one falsely rejects the null hypothesis, i.e. when one finds a significance difference where none in fact exists.

[47] The unusual level of .005 is used because in one portion of the study, an Anova showed a significant interaction, and in the non-parametric "post-hoc" test, at less than .005, the result was not significant.

Contrasts in the context of a front (unrounded) vowel were most difficult to perceive, and those in the context of a back (rounded) vowel easiest to perceive.

Contrast: There was a significant main effect for Contrast, [Friedman, $X^2(6)$, $p<.01$; Anova $F(6,90)$, $p < .001$]. This means that all language groups found certain contrasts easier to perceive than others, across ISI, Voicing, and Vowel conditions. All contrasts were significantly different from each other at the .001 level except for the three pairs that had the fewest errors, viz. there was no significant difference between [t̪,d̪] vs. [t̪s̪,d̪z̪], [t,d] vs. [θ,ð], and [t̪,d̪] vs. [θ,ð]. Overall, performance was poorest on the [t,d] vs. [t̪,d̪] contrast and best on the [t,d] vs. [θ,ð].

Let us now turn to an analysis of how these factors interact with each other.

3.3.2. Interactions

Whereas main effects show that a particular factor has significant effects across-the-board; interactions demonstrate that the behavior of one (or more) factors may be dependent upon another factor.

Voicing x Contrast: A Wilcoxon test showed a significant interaction between Voicing and Contrast. The significance comes from two pairs: [f,v] vs. [ɸ,β] ($X^2(1)$, $p<.001$) and [s̪-θ] vs. [z̪-ð] (($X^2(1)$, $p=.001$) (also Anova ($F(6,90)$, $p<.001$). This is shown in Figure 3.2 below.

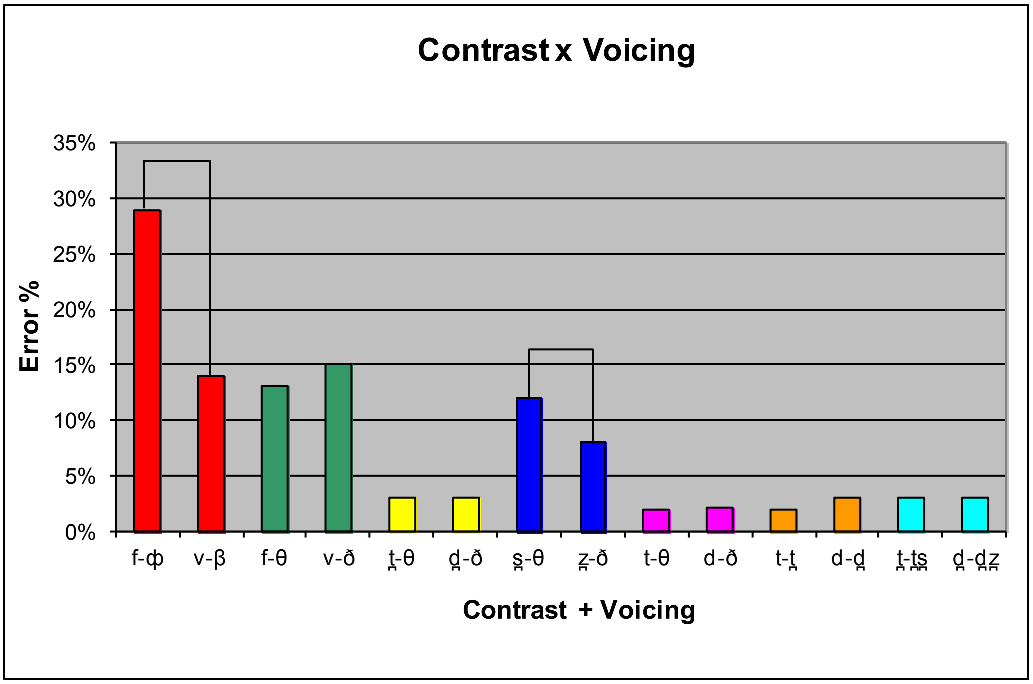

Figure 3.2. Contrast x Voicing Interaction. Connectors show significant differences for voicing within the same contrast.

In both cases, the voiceless contrast was more difficult to perceive than the voiced contrast. This trend holds across language, vowel, and ISI.

Vowel x Contrast: As can be seen in Figure 3.3 below, there was a significant interaction between Contrast and Vowel. The significance is due to four pairs, and in general, perception is best before the non-low back, rounded vowels (U).

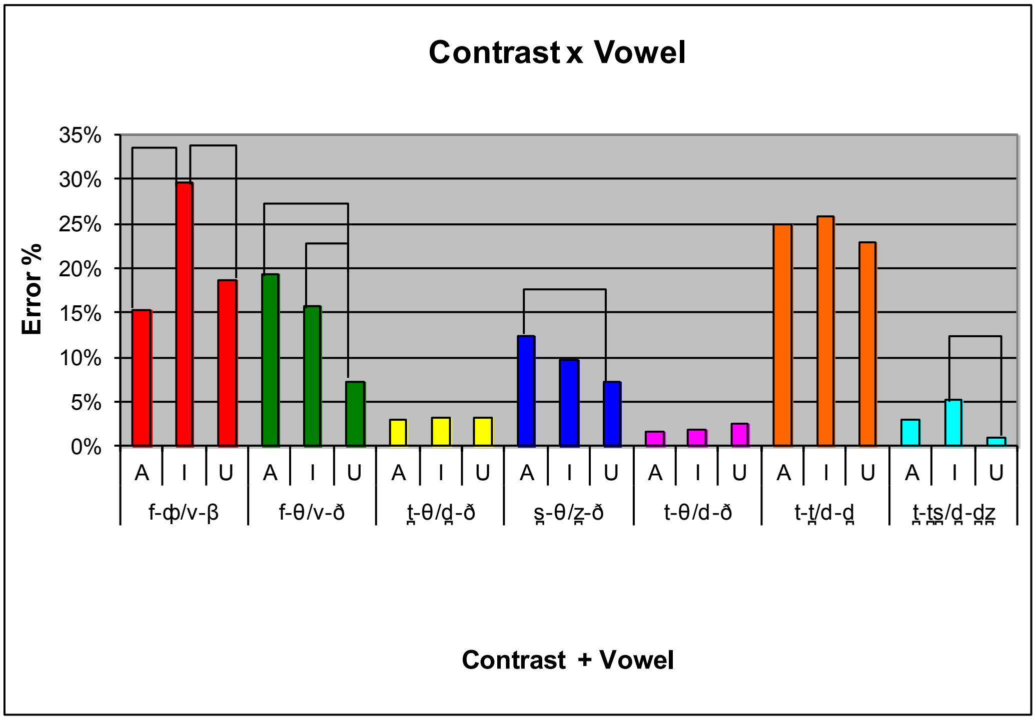

Figure 3.3. Contrast x Vowel Interaction. Connectors show significant differences between vowels within the same contrast.

For [t̪,d̪] vs. [t̪s̪,d̪z̪], performance on U was significantly better than on I (Wilcoxon $X^2(1), p < .001$). For [f,v] vs. [ɸ,β], the fewest errors were on A and next fewest on U, with the highest percentage of errors made on I (Wilcoxon $X^2(1), p < .001$). When the contrast [f,v] vs. [θ,ð] was followed by U, there were significantly fewer errors than when the contrast was followed by either A or I. And for [s̪,z̪] vs. [θ,ð], again U was the most favourable context, significantly different from the low vowel condition, A (Wilcoxon $X^2(1), p < .001$).

Language x Contrast: A Kruskal Wallis test showed a significant interaction between Language and Contrast (also Anova [$F(18,90) = 7.9$, $p < .001$]). Figure 3.4 below gives the error percentages on the test contrasts for each language group.

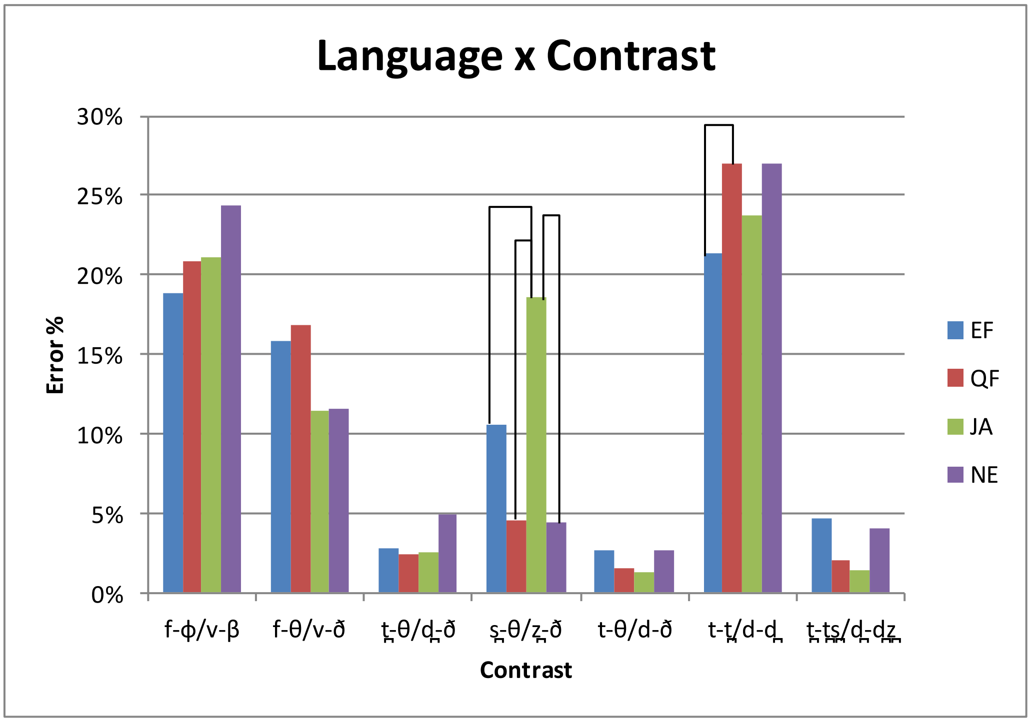

Figure 3.4. Contrast x Language Interaction. Connectors show significant differences between languages within a contrast.

As can be seen in Figure 3.4, there are differences in the ability to perceive certain contrasts depending on the native language. Most strikingly, Mann-Whitney tests showed that the STRIDENT vs. MELLOW fricative contrast ([s̪,z̪] vs. [θ,ð]) is more difficult for the JA listeners than it is for the other language groups. European French was second in terms of level of difficulty on this contrast. Let us now proceed to an interpretation of these results.

3.4. Discussion of Results for AXB-1

3.4.1. ISI

Recall that the ISI condition was intended to isolate phonemic from phonetic processing. This study showed no main effect and no interactions for ISI. This leaves us with the question: What level of processing was accessed during this experiment?

My hypothesis states that if listeners are adopting a phonemic processing mode, then they should fail to discriminate non-distinctive contrasts. On the other hand, if they are adopting a phonetic processing mode, then they should discriminate both distinctive and non-distinctive features. (Note that the fact that there was a significant Language x Contrast interaction suggests that listeners were not using an acoustic mode (see Section 3.1.1)).

If listeners are using a phonetic mode, they should be able to discriminate the non-contrastive pairs at a rate which is significantly better than chance. To verify this, I conducted t-tests comparing the means for each contrast against a hypothesized chance score of 50%. All contrasts were perceived significantly better than the 50% criterion ($p < .001$), except [f-ɸ] before the Vowel I in the Long ISI condition.[48]

Thus, on both phonemic and non-phonemic contrasts, participants had accuracy rates better than would be predicted by chance. This is not what we would expect if they were processing these contrasts in a phonological mode. Therefore, I will assume that they were using a phonetic mode throughout the experiment; that is, in both ISI conditions.

We must now ask why there was no difference between the ISI conditions. It seems that this study has run into the same problems as were encountered by Werker & Logan (1985) in a comparable within-subjects design. Since Werker & Logan did manage to tap phonological processing at 1500msec in the first two blocks of a between-subjects design, I reasoned that if this condition was presented first, I might be able to induce a

[48] [f-ɸ]-I in the Short ISI condition was significant at $p = .01$ ($t = -2.82$).

phonemic strategy.[49] This turned out to be unsuccessful. It seems that the AXB paradigm is not conducive to phonemic processing (see also Curtin, Goad, & Pater 1998).

In chapters 4 and 5 of the thesis, I will present other experiments that attempt to tap phonological processing. For the moment though, let us examine the next factor investigated in the present experiment.

3.4.2. Voicing

The main effect for Voicing shows that voiceless contrasts appear to be more difficult to perceive than their voiced counterparts; however, the interaction of Contrast and Voicing indicates that this difficulty is dependent on the contrast involved. Results show that the significance stems from two contrast pairs: [f-ɸ] vs. [v-β] and [s̪-θ] vs. [z̪-ð]. As shown in Figure 3.2 above, [f-ɸ] has twice the error rate as [v-β] (29% vs. 14%) and [s̪-θ] induced significantly more errors than [z̪-ð] (12% vs. 8%). Interestingly, these pairs contain fricatives; all of the contrasts involving stops showed no significance across voicing conditions. Voiced fricatives are less noisy than their voiceless counterparts due to reduced airflow caused by vocal fold vibration. Thus, the expectation was decreased rather than increased perceptibility in the voiced condition (Hancin-Bhatt 1994a,b).

A possible explanation for the finding that voicing enhances perception is that listeners may be attending to place information in the weak formants and formant transitions which are present on voiced, but

[49] I also tested for ISI differences for the first day of each ISI condition. Because participants heard a total of five hours of stimuli, they no doubt were trained to some degree to perceive non-native distinctions. In examining the first day of each ISI condition, the effect of training should be less evident. However, again there was no main effect for ISI.

not on voiceless fricatives (e.g. Ladefoged 2001). However, it is unclear why the [f-θ] vs. [v-ð] contrast does not participate in this trend.

Since there was no interaction between voicing and language, this means that all language groups behaved in a similar manner with respect to voicing; thus, there are no language particular effects for voiced versus voiceless pairs.

3.4.3. Vowel

In §3.2.1, it was predicted that [t̪,d̪] vs. [t̪s̪,d̪z̪] would be easier to discriminate before low vowels, than before high vowels. Results show a modestly higher error rate in the NON-LOW FRONT (I) condition, but this becomes significant not before LOW (A) vowels, but instead before NON-LOW BACK (U) vowels. Thus it is possible that high vowels do impede perception of this contrast; however, this does not appear to be tied to the existence of [t̪,d̪]-[t̪s̪,d̪z̪] allophony in the L1, since there is no interaction with Language. If such were the case, we would expect QF and JA to show more errors than EF and NE in the I- and U-vowel contexts respectively.[50]

It was also anticipated that the [f,v] vs. [ɸ,β] contrast would be more difficult to discriminate before a non-low front vowel due to [ɸ,β] being misinterpreted as a fronted [f,v]. This is borne out in the results, with a 30% error rate for [f,v] vs. [ɸ,β] in the NON-LOW FRONT VOWEL (I) condition as compared to 15% before A and 19% before U.

It was further predicted that [θ,ð] would have a greater tendency to be confused with [f,v] and [s̪,z̪] before a low vowel. This too seems to be confirmed by the results with [f,v] vs. [θ,ð] at 19% errors, significantly

[50] Note that the results here are confounded somewhat by the fact that the I- and U-vowel conditions included mid vowels in addition to high vowels. In Chapter 4, another AXB test was designed where this problem was resolved.

more than the 7% rate before U. Likewise, [θ,ð] vs. [s̪,z̪] is more difficult to perceive before A (12%) than before U (7%). The remaining pairs show no significant effects for vowel condition; thus, it appears that vowel context has a greater influence on the perceptibility of pairs of fricatives than it does when one member of the pair is a stop or affricate.

3.4.4. Contrast

For ease of reference, below I reproduce (5) as (6), which shows the predicted order of difficulty for each language group. Results are presented in (7). Checkmarks in both (6) and (7) indicate where the predictions coincided with the results.

(6) Predicted Order of Difficulty according to Auditory Distance Model and Perceptual Assimilation Model[51]

Rank	Phonological Processing				Phonetic Processing			
	EF	QF	JA	NE	EF	QF	JA	NE
Most Difficult (SC)	[f-ɸ]✓ [s̪-θ] [t-t̪]✓ [t̪-t̪s̪]	[f-ɸ]✓ [s̪-θ] [t-t̪]✓ [t̪-t̪s̪]	[f-ɸ]✓ [s̪-θ]✓ [t-t̪]✓ [t̪-t̪s̪]	[f-ɸ]✓ [t-t̪]✓	[t-t̪]✓		[f-ɸ]✓	
Moderate Difficulty (CG)				[t̪-t̪s̪]	[f-ɸ] [s̪-θ]✓	[f-ɸ] [t-t̪] [t̪-θ] [t-θ]	[s̪-θ] [t-t̪]	[f-ɸ] [t-t̪] [t̪-t̪s̪]
Least Difficult (TC)	[f-θ] [t̪-θ]✓ [t-θ]✓	[f-θ] [t̪-θ]✓ [t-θ]✓	[f-θ] [t̪-θ]✓ [t-θ]✓	[f-θ] [s̪-θ]✓ [t̪-θ]✓ [t-θ]✓	[f-θ] [t̪-θ]✓ [t-θ]✓ [t̪-t̪s̪]✓	[f-θ] [s̪-θ]✓ [t̪-t̪s̪]✓	[f-θ] [t-θ]✓ [t̪-θ]✓ [t̪-t̪s̪]✓	[f-θ] [s̪-θ]✓ [t̪-θ]✓ [t-θ]✓

[51] For ease of exposition, only voiceless contrasts are presented.

For the results in (7), in order to determine for each language group which contrasts belonged to which categories (SC, CG, or TC), statistical comparisons were run on pairs of contrasts (Wilcoxon tests, $p < .005$). Those contrasts that were significantly different from other contrasts within the same language group have been arranged into their separate categories. Thus, the classification of the results as SC, CG, or TC is relative; for example, [t-t̪] is classified as SC despite being perceived at a rate significantly better than chance.

(7) Results from AXB-1: Order of Difficulty[52]

Rank	EF	QF	JA	NE
Most Difficult (SC)	[f-ɸ] ≈ [t-t̪] ✓ ≈	[f-ɸ] ≈ [t-t̪] ≈	[f-ɸ] ✓ ≈ [s̪-θ] ≈ [t-t̪] ≈	[f-ɸ] ≈ [t-t̪] ≈
Moderate Difficulty (CG)	[f-θ] [s̪-θ] ✓	[f-θ]	[f-θ]	[f-θ]
Least Difficult (TC)	[t̪-θ] ✓ ≈ [t-θ] ✓ ≈ [t̪-t̪s̪] ✓	[s̪-θ] ✓ [t̪-θ] ≈ [t-θ] ≈ [t̪-t̪s̪] ✓	[t̪-θ] ✓ ≈ [t-θ] ✓ ≈ [t̪-t̪s̪] ✓	[s̪-θ] ✓ ≈ [t̪-θ] ✓ ≈ [t-θ] ✓ ≈ [t̪-t̪s̪] ✓

Key: ✓ phonetic prediction confirmed;
≈ phonemic prediction confirmed

[52] Since there were no interactions between language and voicing, I have grouped both voiced and voiceless contrasts in this analysis; however, for ease of exposition, only the voiceless symbols are shown.

When we compare the results in (7) with what was predicted in (6), we see that the phonetic predictions were borne out for [s̪-θ] (EF, QF, and NE) and for [t̪-t̪s̪] (all groups). On the other hand, the phonemic predictions were confirmed for [f-ɸ] (EF, QF, and NE), [t-t̪] (E=QF, JA, and NE), [s̪-θ] (JA), and [t̪-θ]/[t-θ] (QF).

Let us now analyze each contrast in turn.

1. [f,v] vs. [ɸ,β]

This contrast tests the ability of listeners to discern between the features DENTAL and LIP. As can be seen in (6) and (7) above, as predicted, this contrast was perceived as a *Single Category Assimilation* (SC) for JA since they are not purported to process the feature DENTAL, since this feature does not function to define phonological contrasts or to make phonetic specifications in JA. However, for the other language groups, DENTAL does play a phonetic role in the L1, so it was thought that this contrast would fall under a *Category Goodness Assimilation* (CG), and thus EF, QF, and NE should have shown better performance than JA on this contrast. Therefore, it seems that for all languages tested in this study, the features DENTAL vs. LIP are not perceived very well on LABIAL fricatives.

2. [f,v] vs. [θ,ð]

This pair tested the participants' abilities to perceive the difference between the Major Articulator features LABIAL versus CORONAL. All languages were predicted to do well on this contrast, given that all have LABIAL and CORONAL in their inventories. It was hypothesized that each member of this pair would be funneled into a separate category based on these features.

This prediction was not upheld by the data. All languages had moderate difficulty with this pair, classifying it as CG. This contrast is a difficult one cross-linguistically, even for Native English speakers, where these sounds are contrastive (e.g. Miller & Nicely 1955; Hancin-Bhatt 1994a,b; Cutler et al. 2004). This is likely due to the phonetic fact that place of articulation information is poorly encoded on the diffuse spectrum of non-strident fricatives.[53]

Thus, [θ,ð] may often get funneled into the LABIAL category, as seems to be evidenced in the production of certain dialectal variants of English, notably in African American Vernacular English and Cockney English (Babel & McGuire 2010). This would also explain why European French and Japanese speakers sometimes substitute [f,v] for [θ,ð] in production (Wenk 1979, Brannen 1998, Guion et al. 2000). However, it does not explain why [f,v] is not reported as a production substitute in Québec French.[54] However, I speculate that this may be partly explained by the fact that Québec French speakers generally have more contact with English speakers and English visual media, and that visual cues to the labiality of [f,v] and coronality of [θ,ð] are available for encoding in the representations for these segments (Babel & McGuire 2010). This predicts that French learners who acquire English using auditory materials alone would have a propensity to substitute [f,v] more frequently in perception and production. In Chapter 6, I present a study on the production of interdental fricatives, investigating whether we observe more instances of [f,v] for EF speakers.

[53] The lack of saliency for [f-v] vs. [θ-ð] means that this is a weak contrast in English. There are few minimal pairs involving these sounds in English, which is consistent with this.

[54] Although I have heard some QF students use this variant, there were no instances in the data from a production task that I conducted (see Chapter 6).

3. [t̪,d̪] vs. [θ,ð]

This contrast investigates listeners' abilities to discriminate the features STOP and CONTINUANT. It was predicted that EF, JA, and NE participants would do well on this pair, all classifying them as separate categories (TC). On the other hand, QF listeners were predicted to have more difficulty on this contrast, equating [θ,ð] as a poor exemplar of [t̪,d̪], i.e. a CG type.

These predictions were borne out for EF, JA, and NE, but not for QF. The QF group behaved similarly to the other groups with good discrimination of this contrast. However, it may be that these participants, who are, as the reader will recall, advanced learners of English, have already acquired the [θ,ð] category.

4. [s̪,z̪] vs. [θ,ð]

The features which distinguish this pair of phones are STRIDENT and MELLOW. The predictions for this contrast were that EF and JA would group them as CG, and therefore have more difficulty than QF and NE, who would group them as TC. It was anticipated that European French listeners would classify [θ,ð] as a poor exemplar of their native [s̪,z̪], while Japanese listeners were predicted to classify [θ,ð] as a perfect match for their native mellow [s̠,z̠], while considering [s̪,z̪] a poor exemplar of their native alveolar fricatives. On the other hand, the QF group was predicted to group [θ,ð] with their native [t̪,d̪], and equate [s̪,z̪] with their native [s,z], thereby classifying them as *Two Categories*.

For EF, QF, and NE, these predictions were largely borne out. However, as shown in (7), Japanese listeners had the most difficulty with this contrast. In fact, JA has significantly more errors than EF, who in turn has significantly more errors than QF on the [s̪,z̪] pair. This indicates that both [s̪,z̪] and [θ,ð] are perceived as being relatively close to the native [s̠,z̠] of JA.

As concerns EF, results are consistent with the analysis that the feature STRIDENT is present on their CORONAL fricative, but because it is a dental sibilant, it is less strident than an alveolar sibilant (DENTAL MUTES STRIDENT); thus STRIDENT only has a weight of 1 in EF. As such, target [θ,ð] is perceived as a poor exemplar of EF [s̪,z̪], a *Category Goodness Assimilation*.

In QF, the coronal sibilants are fully strident (ALVEOLAR ENHANCES STRIDENT); thus [θ,ð] is perceptually more distant from QF alveolar [s,z], and closer to [t̪,d̪].

5. [t,d] vs. [θ,ð]

This pair tested the ability to discriminate between STOP and CONTINUANT as well as APICAL vs. LAMINAL and ALVEOLAR vs. DENTAL. Similarly to #3 above, it was predicted that EF, JA, and NE would do better than QF, since the former would classify this contrast as a TC assimilation, whereas in QF, it would be a CG assimilation.

It was further predicted that, although QF would have significantly more errors on this contrast than the other language groups, they would have somewhat better discrimination on this contrast which involves an apico-alveolar [t,d], than on the contrast in #3 involving lamino-dental [t̪,d̪]. This is because for QF, there is some distance between their native lamino-dental stops and the target apico-alveolar stops. In other words, the QF listeners would perceive [t,d] as [t̪,d̪] (and [θ,ð] as [t̪,d̪]). For JA and NE, the opposite was predicted: [t,d] is a perfect match to their L1 coronal stops, with no distance incurred.

The results from this contrast, like those from #3 above, do not coincide with what was predicted, likely because these advanced QF learners have acquired the [t,d] vs. [θ,ð] contrast. Neither were significant differences found between #3, [t̪,d̪] vs. [θ,ð], and #5, [t,d] vs. [θ,ð]. Thus, it appears that the QF learners under examination have acquired the feature which allows them to form a separate phonological representation for [θ,ð].

Clearly, we need to investigate learners who are at a less advanced stage of acquisition in order to determine whether QF listeners ever do merge [θ,ð] with [t̪,d̪] or [t,d]. This is addressed in Chapters 4 and 5 of this thesis.

6. [t,d] vs. [t̪,d̪]

This pair tested the ability to perceive the features ALVEOLAR vs. DENTAL and APICAL vs. LAMINAL. Poor performance was predicted for the EF group, for whom this would be a SC type, given that both members of the pair are equidistant from the native EF apico-dental [t̺,d̺]. Conversely, for the other language groups, this was predicted to be a *Category-Goodness Assimilation*: although both sounds are funneled into one native phonetic category, one of the members is a better exemplar than the other; hence we should expect somewhat better performance than for EF.

Results for this contrast indicate that all language groups classified [t,d] vs. [t̪,d̪] as SC. Nonetheless, there was a significant difference between QF and EF, but in the opposite direction from what was predicted. Québec French listeners did significantly worse than EF listeners, with 27% errors vs. 21% respectively. This suggests that the [t̺,d̺] of EF is not equidistant from [t,d] and [t̪,d̪], but rather is closer to one or the other. On the other hand, it appears that the NE listeners do not perceive ALVEOLAR vs. DENTAL and APICAL vs. LAMINAL. The [t,d] stimuli were produced as English-like apico-alveolar stops; therefore, if this group did perceive ALVEOLAR and APICAL, we would expect the [t,d] stimulus to be a perfect match to the NE representation. On the other hand, [t̪,d̪] would be considered a poor exemplar of the NE representation and hence, this pair would be a CG assimilation. The results indicate that this is not the case, and that the NE listeners fail to perceive ALVEOLAR vs. DENTAL and APICAL vs. LAMINAL. Thus, I will assume that NE, QF, and JA are unable to distinguish the features ALVEOLAR vs. DENTAL and APICAL vs. LAMINAL on these stop consonants. As for EF, the results seem indeterminate between SC and CG.

7. [t̪,d̪] vs. [t̪s̪,d̪z̪]

This pair tested the ability to distinguish between MELLOW and STRIDENT on stop consonants. It was predicted that in the phonetic condition, this contrast would be easily perceived for all groups, since it is predicted to be a *Two-Category Assimilation.* This prediction was upheld, as seen in the low error rate.

The features MELLOW vs. STRIDENT were also examined in #4 above, [s̪,z̪] vs. [θ,ð]. There we saw that the Japanese listeners had considerable difficulty distinguishing these features on fricatives; however, here we see that this group does well discriminating these features on stops. If indeed affricates are specified as STRIDENT stops, this would indicate that the availability of a feature somewhere in the phonetic inventory of a language does not guarantee that this feature will be available throughout the inventory, contrary to Brown's (1997,1998) claim. However, it is possible that affricates are distinguished from stops based on duration as well.[55]

3.5. Summary and Discussion of AXB-1

This chapter has investigated the ability of listeners from EF, QF, JA, and NE backgrounds to perceive various contrasts involving anterior obstruents. The main hypothesis was that learners use more than just contrastive, phonemic features when assessing L2 input; they also use phonetic, non-contrastive information. I argued that L1 transfer, and specifically differential interdental substitution, is based upon an evaluation of both contrastive and non-contrastive features in the L2 input. In order to investigate this, two AXB tasks were carried out. One of them was intended to target phonemic processing, with a long interstimulus interval; the other was intended to target phonetic processing, with a short interstimulus interval. It was thought that we would observe differential behaviour in each of the two tasks, with listeners only perceiving features that are contrastive in their L1 in the long ISI task, while on the other hand additionally perceiving non-contrastive features in the short ISI task. This difference did not emerge in the results: both ISIs showed comparable performance. It was concluded that the AXB methodology is more conducive to phonetic processing.

Results from this study indicate that some phonetic features may be involved in differential substitution. Differences between language groups did emerge. In particular, Japanese listeners had significantly more difficulty with the [s̪,z̪] vs. [θ,ð] (STRIDENT vs. MELLOW) contrast than the

[55] To test whether STRIDENT or length is more important to the stop-affricate distinction, one could investigate the ability to perceive non-strident vs. strident affricates (e.g. [t̪θ] vs. [t̪s̪]).

other groups. Notably, the Québec French group perceived this distinction as well as the Native English group, indicating that STRIDENT and MELLOW are perceptually available to QF learners, despite their non-contrastive status in this language. This accounts for why QF speakers do not substitute sibilant fricatives in place of target English interdental fricatives. European French listeners had more difficulty with this contrast than did QF and NE, although this did not reach significance.

For EF, QF, and NE, the cross-linguistic direction of difficulty on the STRIDENT vs. MELLOW fricative contrast was as predicted by the Auditory Distance Model and Best's Perceptual Assimilation Model. European French participants consider [s̪,z̪] and [θ,ð] as belonging to one category, although the fit is not perfect, a *Category Goodness Assimilation.* Québec French and Native English participants, on the other hand, perceive these two sounds as belonging to two different categories, a *Two Category Assimilation.* For the Japanese participants, however, the models predicted better performance than was actually observed. It was argued that this can be attributed to the inability of JA listeners to process the feature STRIDENT on fricatives. The intermediate performance of the EF group was explained by the perceptual proximity of the EF dental fricative to the English interdental fricative. Thus, the feature STRIDENT is specified on the EF dental sibilant; however, it is less salient (DENTAL MUTES STRIDENT).

Another hypothesis investigated in this chapter was that misproduction is due to misperception. This hypothesis is supported for the Japanese group. Results reveal that JA listeners perceptually merge [s̪,z̪] and [θ,ð], corroborating reports in the literature on production and perception (Hancin-Bhatt 1994a,b; Brown 1997,1998). As predicted, European French listeners also had higher rates of errors on [s̪,z̪] vs. [θ,ð] than did QF listeners. This indicates a perceptual basis for differential substitution between these two dialects of French. However, as discussed, performance on [t̪,d̪] and [t,d] vs. [θ,ð] was equivalent to [s̪,z̪] vs. [θ,ð] for QF listeners. This does not mirror their stop substitution errors in production. It was noted, however, that these learners are advanced; thus it may be that they have acquired the relevant categories. In the next chapter, I investigate beginner/intermediate learners to see whether the predicted pattern holds for less advanced learners.

Another key result from this study is the propensity for confusion between [f,v] and [θ,ð] in all language groups tested, including NE. The labiodental fricative is occasionally reported in the production of European French speakers (Wenk 1979, Brannen 1998); although not to the degree of [s̪,z̪]. Apart from the existence of some anecdotal reports, [f,v] has never to my knowledge been reported in the literature for Québec French. It was suggested that visual cues to the labiality of [f,v] may lead speakers away from using this variant in their production.

Also investigated in this study were the effects of voicing and following vowel in the perception of these obstruent pairs. There were no interactions with language for these factors, indicating that the effects found applied to all language groups. It was found that for the contrasts [f,v] vs. [ɸ,β] and [s̪,z̪] vs. [θ,ð], the voiceless members were more difficult to perceive than the voiced members. This is somewhat unexpected, given the perceptual markedness of voiced fricatives. It was suggested that perhaps weak formants on the voiced fricatives may aid perception. As for vowel context, it was found that a back, rounded vowel facilitated perception of pairs that included the interdental fricative. In addition, the stop-affricate distinction was most difficult in the context of front non-low vowels, as expected.

The results of this study support the claim that when target [θ,ð] is substituted with either a labiodental or coronal fricative, this is due to perceptual confusion involving the features STRIDENT and MELLOW. For JA listeners, it is the failure to perceive these features which results in the merger. For EF and QF listeners, enhancement and muting effects involving these features play a role in choice of substitute. Importantly, EF and QF perceive STRIDENT and MELLOW, despite the fact that these features do not function contrastively in either language. This finding goes against Brown's 1997/1998 hypothesis that features which are non-contrastive in a language will not be perceived phonologically.

In the next Chapter, another AXB study will be presented (AXB-2). This study differs from AXB-1 in that it examines less advanced learners of English along with advanced learners. In addition, Russian learners of English are investigated in addition to EF, QF, and JA learners. Russian learners have been reported to substitute [t,d] in place of the English interdental fricatives, although some anecdotal reports suggest that some Russian speakers may substitute [s,z]. The results from AXB-2 will supplement those from AXB-1 and help clarify some of its less conclusive findings.

CHAPTER 4: AXB-2

4.1. Introduction

This chapter extends the investigation of the cross-linguistic perception of interdental fricatives. In Chapter 3 (AXB-1), the results for the Québec French group were not as expected on the stop-interdental contrasts: this was attributed to their advanced level of proficiency. The study reported in this chapter rectifies this by including learners with a lower proficiency in English. As well, a group of native Russian speakers is added in order to help establish which interdental substitute this group prefers. The reader will recall that there are conflicting reports in the literature as to whether Russian ESL learners substitute [t] or [s], which introduces the possibility that there is differential substitution within the Russian speech community. Finally, the AXB-1 study was constructed with different interstimulus intervals (ISI) in order to tap two levels of perceptual processing: a phonetic level and a phonemic level; however, no significant difference emerged for this factor. Therefore, in the current study, methodological changes were made in a further attempt to force this split in processing strategies.

The principal hypotheses from Chapter 3 still hold:

1. Differential substitution is due to transfer from the native language (L1);
2. Transfer in production is caused by transfer in perception;
3. Choice of substitute depends on a comparison of the phonetic properties of the target segment with phonetic properties of segments in the L1 sound system;
4. Transfer of non-contrastive, phonetic features is involved in substitution.

It is hypothesized that when learners initially hear a sound that is not part of their native L1 inventory, they will substitute the sound from their

L1 which is perceptually closest to that non-native sound. This substitute will be used in production; thus substitution in production is perceptually determined.

Results from Chapter 3 supported the hypothesis that some non-contrastive phonetic features play a role in differential substitution across two dialects of the same language, i.e. European French and Québec French. This is consistent with the idea that different dialects can store different phonetic featural representations for the "same" phonemes.

It is argued that these featural representations must somehow manifest themselves on the surface; otherwise, the L1 learner would have no evidence for their existence. Since different dialects (and languages) can have different phonetic expressions for the same phonemes, it is considered that these phonetic cues are what learners use to postulate different featural representations for their native sound inventory. Consequently, in these cases, different phonetic specifications in various dialects lead to distinct interpretations of the L2 input, and hence differential substitution.

In Chapter 3, we examined English learners with the following native language backgrounds: European French (EF), Québec French (QF), and Japanese (JA). All were classified as being at an advanced level of English proficiency.

For each of these languages, the following phonetic feature specifications are considered to define their anterior obstruents. Only voiceless segments are given, since AXB-2 does not test the voiced segments. The bolded and underlined features function contrastively in the language.

(1) European French

[f]	[t̪]	[s̻]
CONTINUANT	**STOP**	**CONTINUANT**
LABIAL	**CORONAL**	**CORONAL**
MELLOW	MELLOW	STRIDENT
DENTAL	DENTAL	DENTAL
	APICAL	LAMINAL

(2) Québec French

[f]	[t̻]	[s]	[t̻s̻]
CONTINUANT	**STOP**	**CONTINUANT**	**STOP**
LABIAL	**CORONAL**	**CORONAL**	**CORONAL**
MELLOW	MELLOW	STRIDENT	STRIDENT
DENTAL	DENTAL	ALVEOLAR	DENTAL
	LAMINAL	LAMINAL	LAMINAL

(3) Japanese

[ɸ]	[t]	[s̱]	[ts]
CONTINUANT	**STOP**	**CONTINUANT**	**STOP**
LABIAL	**CORONAL**	**CORONAL**	**CORONAL**
MELLOW	MELLOW	MELLOW	STRIDENT
LIP	ALVEOLAR	ALVEOLAR	ALVEOLAR
	APICAL	LAMINAL	LAMINAL

Russian (RU) was not included in the AXB-1 study reported in Chapter 3. The reader will recall from the conclusion of Chapter 3 that there is some uncertainty as to how Russian learners adapt the interdental fricatives. As a potential case of inter-dialectal differential substitution, this language group was added to the AXB-2 study reported in the present chapter. Therefore a description of the RU consonant inventory is called for. Russian has a rich inventory of consonants as shown in (4) below.

(4) Russian Phonemic Inventory (Jones & Ward 1969, Hamilton 1980, Maddieson 1984)

p pʲ		t̪ t̪ʲ				k	kʲ
b bʲ		d̪ d̪ʲ				g	
	f fʲ	s̪ s̪ʲ		ʃ		x	
	v vʲ	z̪ z̪ʲ		ʒ			
			ts	tʃ			
m mʲ		n̪ n̪ʲ					
		ɫ̪ l̪ʲ					
				r rʲ			
					j		

The RU phonetic inventory of anterior consonants includes a labiodental fricative, a dental stop (Jones & Ward 1969, Hamilton 1980, Maddieson 1984), a dental fricative (Hamilton 1980, Maddieson 1984) and an alveolar affricate (Jones & Ward 1969).[56] Except for the affricate, these segments have palatalized counterparts, which also function contrastively. As well, Russian has phonetic variants of the obstruents, which are labiovelarized before the vowels [u o]. The palatalized and labiovelarized segments are not included below since they were not included in the test

[56] Jones & Ward mention that there is inter-speaker variation in the production of the RU coronal stop and fricative; some speakers produce them as an apical, with the tongue tip up, while others produce them as a laminal, with the tongue tip down. Given this information, I represent the RU coronal stop and fricative as permanently underspecified for apicality and laminality.

items (see Appendix C for complete distance calculations including these segments).

(5) Russian Phonetic Representations

[f]	[t̪]	[s̪]	[ts]
CONTINUANT	**STOP**	**CONTINUANT**	**STOP**
LABIAL	**CORONAL**	**CORONAL**	**CORONAL**
MELLOW	**MELLOW**	**STRIDENT**	**STRIDENT**[57]
DENTAL	DENTAL	DENTAL	ALVEOLAR
PLAIN	PLAIN	PLAIN	PLAIN

With these representations in mind, let us now proceed to an examination of the predictions they yield.

4.2. Predictions

Predictions for both a phonemic assessment and a phonetic assessment are given, beginning with the phonemic assessment. Recall the hypothesis that substitution is based on a phonetic assessment.

4.2.1. Phonemic Predictions

In (6) and (7) below, I give the predictions based on a phonemic assessment, where only contrastive features are accessible for comparison. The reader is reminded that I consider phonological feature representations

---

[57] Note that to distinguish [ts] from [t̪], either STRIDENT/MELLOW or ALVEOLAR/DENTAL may be the relevant features. I have chosen STRIDENT/MELLOW since, given the choice, this is most commonly identified distinguishing feature (e.g. Rubach 1994).

to be categorical. In addition, enhancement does not enter into phonological calculations; thus only the number of featural conflicts is assessed.

(6) Predictions for EF, QF, and JA Based on a Phonemic Assessment

Intake {θ}	CONTINUANT CORONAL	
Potential Substitute	*Mismatches*	*# of Featural Conflicts*
☞ /s/		0
/t/	STOP	1
/f/	LABIAL	1

(7) Predictions for RU Based on a Phonemic Assessment

Intake {θ}	CONTINUANT CORONAL MELLOW	
Potential Substitute	*Mismatches*	*# of Featural Conflicts*
☞ /f/	LABIAL	1
☞ /t/	STOP	1
☞ /s/	STRIDENT	1

For RU, the intake is specified for more features than it is for EF, QF, and JA. This is because RU needs more phonemic features to differentiate the segments in its rich consonantal inventory. In RU, STRIDENT and MELLOW have distinctive status to contrast the coronal stop /t/ from the coronal affricate /ts/. PLAIN and SHARP serve to distinguish the plain obstruents from the palatalized obstruents.[58] As can be seen in (7), three segments differ minimally from the intake theta on a single feature; thus, in a phonemic assessment for RU, these three segments /f t s/ would be equally likely as potential substitutes, and one would expect free variation between them.

Let us now turn to the phonetic predictions for the test languages.

4.2.2. Phonetic Predictions

Recall from Chapter 2 the conventions used in the tables of assessment of potential substitute. A negative sign indicates a featural mismatch, i.e. where the feature of the substitute differs from the feature of the intake. A positive sign indicates that the features of the substitute and intake agree; however, their auditory salience differs.

The phonetic predictions for EF are given in (8):

---

[58] The features PLAIN and SHARP were proposed by Jakobson, Fant, and Halle (1963/1969). They are left out of the prediction as they have no effect on the outcome regarding target theta.

(8) European French Phonetic Predictions

Intake		CONT	+1
{θ}	*salient*	COR	+2
		MELL	+1
		DENT	+1
		LAM	+1

Potential Substitute	*Mismatches*			*Distance*	*Total Distance*
☞ [s̪]	*enhanced*	CONT	+2	1	
		STRID	-1	2	3
[f]	*salient*	LAB	-2	4	4
[t̪]	*enhanced*	STOP	-2	3	
		AP	-1	2	5

For EF, the intake is fully specified. It is predicted that [s̪] will be perceptually substituted for the target interdental fricative. The labiodental fricative is next with a distance of 4 due to a mismatch on the salient Major Articulator features. The coronal stop is perceptually farthest away from the target, with a mismatch on continuancy and the Minor Articulator features.

In (9), we see the phonetic predictions for QF:

(9) Québec French Phonetic Predictions

Intake			
{θ}		CONT	+1
	salient	COR	+2
		MELL	+1
		DENT	+1
		LAM	+1

Potential Substitute	*Mismatches*	*Distance*	*Total Distance*
☞ [t̪]	*enhanced* STOP -2	3	3
[f]	*salient* LAB -2	4	4
[t̪s̪]	STOP -1	2	4
	STRID -1	2	
[s]	*enhanced* CONT +2	1	
	enhanced STRID -2	3	
	ALV -1	2	6

For QF, the intake is fully specified. The lamino-dental stop is predicted to be substituted for target theta given that these two segments differ only on one feature and this feature is not salient on the intake. The labiodental fricative and affricate are perceptually farther away from the target, the first differing on an inherently salient feature, the second differing on two features. Finally, the coronal sibilant is auditorily farthest, with featural mismatches on STRIDENT and ALVEOLAR, as well as having the more salient CONTINUANT feature.

In (10) below, we see that for JA, the intake is specified for all features except DENTAL. This is because JA is considered to have no dental consonants in either its phonemic or phonetic inventories; thus this feature is considered to be permanently underspecified in Japanese and will not be parsed on the intake.

(10) Japanese Phonetic Predictions

Intake {θ}			
		CONT	+1
	salient	COR	+2
		MELL	+1
		~~DENT~~	~~+1~~
		LAM	+1

Potential Substitute	Mismatches			Distance	Total Distance
☞ [s̪]				0	0
[ɸ]	*salient*	LAB	-2	4	4
[t]	*enhanced*	STOP	-2	3	5
		AP	-1	2	
[ts]		STOP	-1	3	5
	enhanced	STRID	-2	2	

The coronal fricative [s̪] is predicted to be selected as the substitute for English theta; since the JA [s̪] is considered to be a mellow fricative, it agrees on all features with the target interdental. The JA phone [ɸ] is perceptually farther from the intake, differing on a salient feature. The coronal affricate and stop are yet farther, differing on two phonetic features.

The phonetic predictions for Russian are in (11):

(11) Russian Phonetic Predictions

Intake	
{θ}	CONT +1 COR +2 MELL +1 DENT +1 LAM +1

Potential Substitute	*Mismatches*	*Distance*	*Total* **Distance**
☞ [s̪]	*enhanced* CONT +2 STRID -1	1 2	3
☞ [t̪]	*enhanced* STOP -2	3	3
[f]	*salient* LAB -2	4	4
[ts]	STOP -1 *enhanced* STRID -2 ALV -1 AP -1	2 3 2 2	9

For RU, both [s̪] and [t̪] are equally plausible substitutes for English theta. The labiodental fricative is next, disagreeing on a salient feature. Finally, the affricate is farthest from the intake, differing on several features.

To summarize this section, we have seen predictions for both a perceptual assessment at the phonemic level and at the phonetic level. Predictions for a phonemic assessment, where only the number of featural mismatches of contrastive features is evaluated, indicate that European French, Québec French, and Japanese would all substitute /s/ for the English interdental fricative. Russian is predicted to show equal preference for /f/, /t/, and /s/ if only phonemic features are assessed.

If, on the other hand, featural assessment takes place at a phonetic level of representation, where non-contrastive features and featural enhancement and salience play a role, as hypothesized in this thesis, European French and Japanese learners are predicted to substitute either [s̻]/[s̠]. Québec French listeners are predicted to substitute [t], and for Russian, either [s̻] or [t̻].

4.3. Design and Method

4.3.1. Interstimulus Interval (ISI)

Recall that the AXB-1 study failed to find a significant difference between the short and long interstimulus interval (ISI) conditions. The results suggested that the AXB-1 task failed to tap a phonemic level of processing. As discussed here and below, several modifications were made to the AXB-2 design in an attempt to elicit the phonemic mode.

One of these changes is that different talkers produced the stimuli within each item. For example, the participant would hear A and B produced by one talker, and X produced by another talker. This methodology encourages phonemic processing, as the listener is forced to abstract across voices.

Another refinement is that the test words or non-words were embedded and presented to the participants within a carrier phrase, either "You hear ___" or "I learn ___." In the Long ISI condition, the ISI was 1500 msec; therefore, including the carrier phrase, the time between target words was 2 seconds, 500msec longer than that of the Long ISI condition of AXB-1. In the Short ISI condition, the ISI was only 50 msec; thus, including the carrier phrase, the time between target words was 250 msec, equivalent to that of the Short ISI condition in AXB-1. This relatively short interval between target words is considered adequate for phonetic processing (Werker & Logan 1985).

4.3.2. Wordhood

Both words (W) and Non-Words (NW) were used in the Short ISI condition. However, in the Long ISI condition, only real words (W) were used. This was done in order to favour lexical access, and thus phonological processing. By eliminating non-words, the contrasts involving [ɸ] and [t̪s̪] were not included in the Long ISI condition, because these sounds are not part of the native English inventory. However, the sounds [t̪] and [s̪] *were* included, despite being phonetically different from English [t] and [s] in being DENTAL rather than ALVEOLAR. These were retained because the results from AXB-1 gave an indication that DENTAL/ALVEOLAR might be perceptible to EF, and this is further investigated here. Also, while dental [s̪] was contrasted with [θ] in AXB-1, alveolar [s] was not, so it was decided to test it in both ISIs in the present study.

4.3.3. Proficiency Level

In AXB-1, only Advanced (A) learners of English participated. In the present experiment, Beginner-Intermediate (BI) learners were also tested. All non-native participants were classed as either BI or A based on the aural comprehension component of the standardized Michigan English Placement Test.[59] Comparing two proficiency levels will enable us to see if different patterns emerge for the BI learners as compared to the A learners, and to evaluate the effect of proficiency level with respect to the various contrasts tested.

[59] Beginner and Intermediate learners were grouped together. This was due to difficulty in finding sufficient Beginner participants, especially for the EF group.

In particular, the two proficiency levels will enable us to further investigate QF behaviour. Recall that in AXB-1, QF listeners did very well on the [t,d] vs. [θ,ð] and [t̪-d̪] vs. [θ,ð] contrasts. This went against the hypothesis that production errors are based on perceptual errors for this language group, because QF speakers produce [t̪,d̪] in place of the target interdental fricatives and yet AXB-1 showed that the advanced QF learners did not perceptually confuse the stops with [θ,ð]. By investigating Beginner-Intermediate learners, we can find out whether this pattern extends to less experienced L2 learners from QF backgrounds.

4.3.4. Vowel

Another difference in the AXB-2 task has to do with the vowels. Recall that in AXB-1, the vowels were classified into low (A), front non-low (I), which merged high front and mid front vowels, and back non-low (U), which merged high back and mid back vowels. In AXB-2, the divisions for the vowel conditions were redefined so that we now have the following three vowel conditions: HIGH FRONT (HF) [i ɪ], HIGH BACK (HB) [u], and NON-HIGH (NH) [eɪ ɛ ɜ ʌ oʊ æ aɪ ɑ]. By isolating the high vowels from the mid vowels, the effects of L1 allophonic processes could be examined. Specifically, in JA, there are phonological processes whereby /s/ → [ɕ]/__ [i]; /t/ → [ts]/__ [ɯ]; and /t/ → [tɕ]/__ [i]. In her dissertation, Muñoz-Sánchez (2003) found that JA listeners had more difficulty perceiving the [s-θ] contrast in the context of the high front vowel [i] as opposed to the low vowel [a]. Thus, we might expect the same pattern to emerge in the AXB-2 study, with higher errors for the JA group on [s-θ]/[s̪-θ] before HF vowels versus NH or HB vowels. In the same vein, it is anticipated that more errors will be observed for [t-θ]/[t̪-θ] before high vowels compared to non-high vowels because of JA speakers' lack of experience with [tu] and [ti].

In QF, there is affrication of /t,d/ before high front vowels, i.e. /t/ → [t̪s̪]/__[i,ɪ]; therefore, it is expected that [t-θ]/[t̪-θ] will be more difficult in the HF condition versus the other vowel contexts.

4.3.5. Language

For the Advanced proficiency condition, the same language groups were investigated as in AXB-1: European French (EF), Québec French (QF), Japanese (JA), with Native English controls (NE).

In the Beginner-Intermediate proficiency condition, these same groups were examined, but another group, Russian (RU) was added. Russian was added because observed discrepancies suggest that we might have a similar situation in Russian as we have with French, i.e. inter-dialectal differential substitution.

4.3.6. Contrast

In this AXB-2 study, most contrasts are the same as in AXB-1 (see §3.1.3.2. in Chapter 3). One difference in the present experiment is that all sounds are paired with [θ]: in the Short ISI condition -- [ɸ-θ f-θ s̪-θ s-θ t̪-θ t-θ t̪s̪-θ]; in the Long ISI condition -- [f-θ s̪-θ s-θ t̪-θ t-θ]. It was felt that this adjustment to the experimental design would be more truly representative of how people process featural contrasts. For example, it may be that listeners are able to hear the difference between DENTAL and ALVEOLAR on stop consonants; whereas, this difference is imperceptible on fricative consonants, or the difference may be perceptible on strident fricatives but not on mellow fricatives. In other words, the ability to discriminate certain features is likely dependent on the presence or absence of other features.

Also, in the current study, it was decided to use voiceless contrasts only. This was done for a few reasons. First, in AXB-1, the only significant results involving Voicing was an interaction between Voicing and Contrast, with [f-ɸ] and [s̺-θ] causing more difficulty than [v-β] and [z̺-ð] respectively. In the current design, [f-ɸ] was not tested. Second, no real word pairs can be constructed for [z̺-ð] in onset position. Finally, removing voiced contrasts aided in shortening the experiment, thus reducing the chances of non-representative results due to fatigue and habituation on the part of the participants.

There were a total of 168 test tokens presented.

(12) Test items used in AXB-2

TEST ITEMS AXB-2								
Distinguishing Feature(s)	**Segments**	**Number of tokens by vowel and wordhood**						
		NH		**HF**		**HB**		**Total tokens**
		w	**nw**	**w**	**nw**	**w**[60]	**nw**	
LABIAL vs. CORONAL LIP vs. DENTAL	[ɸ̪] vs. [θ]	0	6	0	6	0	6	**18**
LABIAL vs. CORONAL	[f] vs. [θ]	6	6	0[61]	6	0	6	**24**
STOP vs. CONT	[t̪] vs. [θ]	6	6	0[62]	6	0	6	**24**
STOP vs. CONT ALVEOLAR vs. DENT APICAL vs. LAMINAL	[t] vs. [θ]	6	6	6	6	0	6	**30**
STRIDENT vs. MELL	[s̪] vs. [θ]	6	6	6	6	0	6	**30**
STRIDENT vs. MELL ALVEOLAR vs. DENT	[s] vs. [θ]	6	6	6	6	0	6	**30**
STRIDENT vs. MELL STOP vs. CONT	[t̪s̪] vs. [θ]	0	6	0	0	0	6	**12**
Total per Vowel and Wordhood		**30**	**42**	**18**	**36**	**0**	**42**	Grand Total **168**

There were 144 filler items ([w-j], [ʃ-tʃ], [m-n], [t-k], [t-p]).

---

60 There are no frequent words in English beginning with [θu].

61 This cell is empty because all relevant pairs involve infrequent words.

62 An error in test construction resulted in this cell being empty, as well as that for [t̪s̪-θ] NW.

4.4. Participants

A total of 79 individuals participated in the AXB-2 experiment (EF-BI 8; EF-A 12; QF-BI 10; QF-A 13; JA-BI 10; JA-A 7; RU-BI 10; NE 9). Each participant did both the Short ISI and Long ISI tests. Participants were removed as outliers from either the Short and/or Long ISI if their error rate on the filler items (in the particular ISI condition) was two standard deviations from the mean of that language group. Thus, a participant could potentially be removed from the Short, the Long, or both ISI conditions. Participants removed on this basis were: one QF-BI from Short ISI; one EF-A from Short ISI; two QF-A, one from Short ISI, another from Long ISI; and one JA-A from Short ISI.

4.5. Procedure

The procedure followed in AXB-2 is much the same as that in AXB-1. There were a couple of changes however. First, unlike AXB-1, in the present study, both the Long and Short ISI were completed on the same day. Second, as mentioned above in §4.3.1, the test words or non-words were embedded and presented to the participants within a carrier phrase, either "You hear ___" or "I learn ___." Carrier phrases were not controlled, so sometimes all three phrases were the same while in other cases the carrier phrases differed within an item. This was done as it was felt that allowing the carrier phrases to vary would promote linguistic as opposed to non-linguistic, acoustic processing by forcing the listeners to abstract away from the context. In other words, the listener could not rely on physical identity between two of the phrases.[63]

Other changes to the design were discussed in §4.3.

4.6. Predictions for Phonetic vs. Phonological Processing

1. [ɸ] vs. [θ]

This contrast was tested in the Short ISI condition only, so only the phonetic predictions are given in (13) below. The two features which distinguish this pair of sounds are LABIAL/CORONAL and LIP/DENTAL. For EF, QF, and NE, {ɸ} should be equated with [f], and for JA, with [ɸ].[64] Concerning {θ}, it is predicted that EF will associate it with [s̪]. Québec French is predicted to associate {θ} with [t̪], Japanese with [s̠], Russian with either [s̪] or [t̪], and of course NE should identify it correctly.[65]

The table below gives the predicted outcomes for each language. It shows the type of assimilation expected according to the Perceptual Assimilation Model (as described in Chapter 3) and whether discrimination of the contrast will be good, moderate, or poor. For the contrast [ɸ] vs. [θ], since each should be assimilated to separate L1 categories for all languages in the phonetic condition, performance should be good.

---

[63] Note that listeners could not make distinctions based on carrier phrase alone. Different talkers used different carrier phrases, and the same talker used different carrier phrases, even for the same item.

[64] Recall that the curly brackets represent the intake segment.

[65] As mentioned at the outset of this chapter, Russian was included to explore the possibility that differential substitution exists between dialects of this language. However, in these predictions, I have not taken this potential differential substitution into account since I have been unable to find adequate phonetic descriptions of different Russian dialects.

(13) Predictions for [ɸ] vs. [θ]

Language	Phonemic (Long ISI)	Phonetic (Short ISI)
EF	not tested	TC (good)
QF		TC (good)
JA		TC (good)
RU		TC (good)
NE		TC (good)

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

2. [f] vs. [θ]

This contrast was included in the Long and Short ISI conditions, so predictions for both phonemic and phonetic processing are given below. Since [f] and [ɸ] are predicted to be merged for all language groups in the phonetic condition, the Short ISI predictions outlined in #1 will be the same for this contrast as well.

In the phonemic condition, JA should merge {f} with either /p/, /s/, or /h/, or with /ɸ/ if this segment has phonemic status in JA (Vance 2008). In the other languages, {f} constitutes a match with their native /f/. European French and Québec French are predicted to merge {θ} with /s/, given that these sounds are the same phonologically, sharing the contrastive features CONTINUANT and CORONAL; thus, this contrast will be a TC assimilation for EF and QF. Japanese is also predicted to equate {θ} with /s/, so depending on which association is made for {f}, we can expect either CG or TC for this group. For RU, several substitutes are possible for {θ}, as outlined in (7) above: if /f/ is selected, performance should be moderate, but if any of the others are chosen, then performance should be good.

(14) Predictions for [f] vs. [θ]

Language	Phonemic (Long ISI)	Phonetic (Short ISI)
EF	TC (good)	TC (good)
QF	TC (good)	TC (good)
JA	CG (moderate) or TC (good)	TC (good)
RU	CG (moderate) or TC (good)	TC (good)
NE	TC (good)	TC (good)

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

3. [t̪] vs. [θ]

The features contrasting these segments — STOP vs. CONTINUANT — are contrastive in all languages tested. At the phonemic and phonetic levels, it is anticipated that all languages will equate {t̪} with their native /t/. In phonemic processing, European and Québec French, and Japanese should identify {θ} with /s/; and for RU, there are four potential substitutes, /f s t x/, all differing from {θ} on the basis of one contrastive feature. So it is predicted that EF, QF, JA, and NE will have a TC assimilation in the Long ISI condition, while for RU, it will be either CG or TC.

In phonetic processing however, {θ} is predicted to be associated with [s] for EF and JA; while QF is predicted to associate {θ} with [t̪], and Russian with either [s̪] or [t̪], and NE should identify it correctly. This means we should expect good performance for EF, JA, and NE; either moderate or good performance for RU, and only moderate performance for QF.

(15) Predictions for [t̪] vs. [θ] and [t] vs. [θ]

Language	Phonemic (Long ISI)	Phonetic (Short ISI)
EF	TC (good)	TC (good)
QF	TC (good)	CG (moderate)
JA	TC (good)	TC (good)
RU	CG (moderate) or TC (good)	CG (moderate) or TC (good)
NE	TC (good)	TC (good)

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

4. [t] vs. [θ]

The predictions for this contrast are the same as those in the table in (15) above; although, in the Short ISI condition, [t̪-θ] should be more problematic for QF than [t-θ], since the pair of sounds in the former is phonetically closer than those in the latter.

5. [s̪] vs. [θ]

The features of interest here — STRIDENT vs. MELLOW — are not contrastive for the EF, QF, and JA groups. Thus, they should be unable to hear the difference in the phonological condition. For RU, it is predicted that this group will associate {θ} with either /f s t x/; thus if associated with /s/, discrimination would be moderate, these two segments differing only in stridency, but if associated with any of the other native segments, discrimination should be good.

By hypothesis, these features are available at the phonetic level, but nonetheless, EF and JA are predicted to associate {θ} with [s], albeit a CG association; QF is predicted to associate {θ} with [t̪], and Russian with either [s̪] or [t̪], and NE should identify it correctly. This means we should expect moderate performance for EF; good performance for QF; either moderate or good performance for RU and JA. For JA, a TC assimilation is possible here because although target dental {s̪} should be associated with JA [s̠], target alveolar {s} could be associated with the JA allophone [ts] before a HB vowel and with JA [s̠] in other vowel contexts. Finally, good performance is expected for NE.

(16) Predictions for [s̪] vs. [θ] and [s] vs. [θ]

Language	Phonemic (Long ISI)	Phonetic (Short ISI)
EF	SC (poor)	CG (moderate)
QF	SC (poor)	TC (good)
JA	SC (poor)	CG (moderate) or TC (good)
RU	CG (moderate) or TC (good)	CG (moderate) or TC (good)
NE	TC (good)	TC (good)

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

6. [s] vs. [θ]

The predictions for this contrast are the same as those in the table in (16) above; although, in the phonetic condition, for EF, [s̪-θ] should present more difficulty than [s-θ], given that the sounds in the former pair are phonetically closer than those in the latter.

7. [t̪s̪] vs. [θ]

In the present experiment, this contrast was tested in the Short ISI Non-

Word condition only. The features involved in this pair are considered to be STOP vs. CONTINUANT and STRIDENT vs. MELLOW. Québec French and Japanese are expected to do well on this contrast. The European French group should merge [t̻s̻] with their native [s̻], given that these sounds differ only in the feature CONTINUANT, which is subject to DENTAL MUTES CONTINUANT, thus rendering this feature relatively non-salient. Even though Russian has a native apico-alveolar [ts], it is predicted that this group will instead associate {t̻s̻} with either [t̻] or [s̻], given that the latter are lamino-dental. Recall from (8) above that EF is predicted to substitute intake {θ} with [s̻]: both native categories are equidistant from the target, thus an SC assimilation. For RU, this merger should be a CG assimilation. For Native English, it is possible that {t̻s̻} would be associated with [θ], in which case we might expect moderate performance.

(17) Predictions for [t̻s̻] vs. [θ]

Language	Phonemic (Long ISI)	Phonetic (Short ISI)
EF	not tested	SC (poor)
QF		TC (good)
JA		TC (good)
RU		CG (moderate)
NE		CG (moderate)

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

Let us now turn to the results from this experiment.

4.7. Results

Each participant's error rate was calculated for each contrast. The design used in the experiment presented in this chapter involves several instances where the number of levels for a given factor is uneven. This can be seen in (18) and (19) below. Under Language, there is no Advanced (A) group for Russian due to difficulty in recruiting Russian participants who were advanced learners of English. Native English serves as the control group to be compared with both Beginner-Intermediate (BI) and A test groups. Under ISI, there is no Non-Word (NW) variable in the Long ISI condition, as discussed in §4.3.2. Under Wordhood, there are no contrasts involving [ɸ] or [t̪s̪] in the Word (W) condition. Under Contrast, there are levels of Vowels omitted in the Word condition. This is because there are few or no frequent real words with the combination of test contrast and class of vowel. Finally, the HF vowel was excluded from [t̪s̪] NW and [t̪] W due to an error in the test construction.

The tables in (18) and (19) below give the factors that were analyzed in this study: between-subjects and within-subjects. Shaded cells represent factors that were omitted.

(18) Between-subjects design

Between-Subjects	
EF	BI
	A
QF	BI
	A
JA	BI
	A
RU	BI
	A
NE	n/a

(19) Within-subjects design

Within-Subjects			
ISI	**Wordhood**	**Contrast**	**Vowel**
Short	NW	ɸ	HB
			HF
			NH
		f	HB
			HF
			NH
		s	HB
			HF
			NH
		s̪	HB
			HF
			NH
		t	HB
			HF
			NH
		t̪	HB
			HF
			NH
		t̪s̪	HB
			HF
			NH
	W	f	HF
			NH
		s	HF
			NH
		s̪	HF
			NH
		t	HF
			NH
		t̪	HF
			NH
Long	W	f	HF
			NH
		s	HF
			NH
		s̪	HF
			NH
		t	HF
			NH
		t̪	HF
			NH
	NW		

In order to present a clear view of the data, it is necessary to section the design and present the results from several different perspectives. This was done in order not to confound interactions and main effects. This required that the data be split in 18 different ways or “data sets”. Below, the results will be presented for each of these 18 data sets (Data Sets A through R), organized by factor. The factors are: Interstimulus Interval (ISI), Proficiency level, Wordhood, Vowel, Contrast, and Language. For example, all data sets that include both ISI levels are included under ISI; whereas, any data sets which have only one level of ISI are excluded. This means one data set will be investigated from the perspective of different factors; for example, a particular data set might be examined under ISI as well as under Contrast and Language.

Both (parametric) Anovas with or without Bonferroni correction and non-parametric statistics are often provided. Results will be reported as significant only when they are so in both tests. For the Anova, significance levels were set at $p = .05$; for the non-parametric tests, it was set at $p = .005$ (see Chapter 3, §3.3).

4.7.1. Statistical Results for Interstimulus Interval (ISI)

The first six data sets enable us to see the effect of Interstimulus Interval (ISI).

Data Set A. (EF/QF/JA/RU/NE; BI; Short/Long; W; [f s s̪ t t̪]; NH). Only real words were compared for the effect of ISI, since in the Long ISI condition, only real words were used. Since Contrast and Vowel were not evenly balanced, i.e. Contrasts [s-θ], [s̪-θ], and [t-θ] occurred before both HF and NH vowels, but Contrasts [f-θ] and [t̪-θ] only occurred before NH vowels, an Anova was run on NH only. Only BI is run for this analysis, since there are no Advanced learners from the RU group.

Neither the Anova nor the non-parametric tests show a significant difference between Short and Long ISI. There were no significant interactions either.

Data Set B. (EF/QF/JA/NE; A; Short/Long; W; [f s s̺ t t̺]; NH). This analysis is the same as above, but instead examines the Advanced learners.

The Anova test for the effect of ISI in this second analysis is not significant. There are no significant interactions involving ISI.

Data Set C. (EF/QF/JA; BI/A; Short/Long; W; [f s s̺ t t̺]; NH). This analysis is the same as above on the dimensions of ISI, Wordhood, Contrast, and Vowel; however, this time, proficiency levels are also compared; therefore RU and NE are omitted, since they only have one level of proficiency.

In this analysis, the Anova does show a significant main effect of ISI (Anova, $F(1,48), p<.05$). There are more errors in the Long ISI condition, as expected, due to the great load on memory imposed by the long interstimulus interval. However, there are no significant interactions involving ISI.

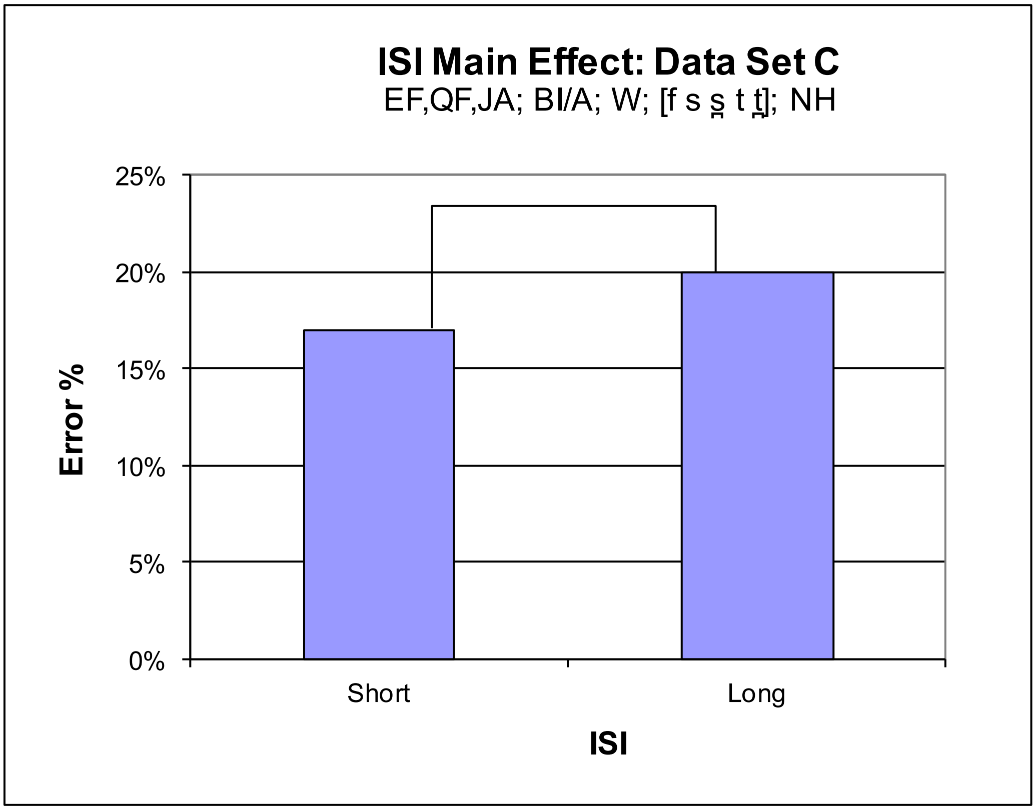

Figure 4.1. Data Set C. ISI Main Effect.
Significant differences are indicated with a connector line.

Results indicate that with an increased ISI, the error rate increases across the board. Since there are no interactions with ISI, this suggests that the longer ISI puts a burden on general memory, and does not indicate that different levels of representation are being tapped.

Data Set D. (EF/QF/JA/RU/NE; BI; Short/Long; W, [s s̺ t]; HF/NH). This analysis is similar to Data Set A above, except that it includes both HF and NH vowels. The Anova does not report significance for Main Effect or Interactions for ISI.

Data Set E. (EF/QF/JA/NE; A; Short/Long; W; [s s̺ t]; HF/NH). This analysis looks at the same factors as Data Set D above, but this time for the

Advanced learners. Here the Anova shows a significant main effect for ISI, (Anova F(1,33),p=.002), but there are no significant interactions involving ISI.

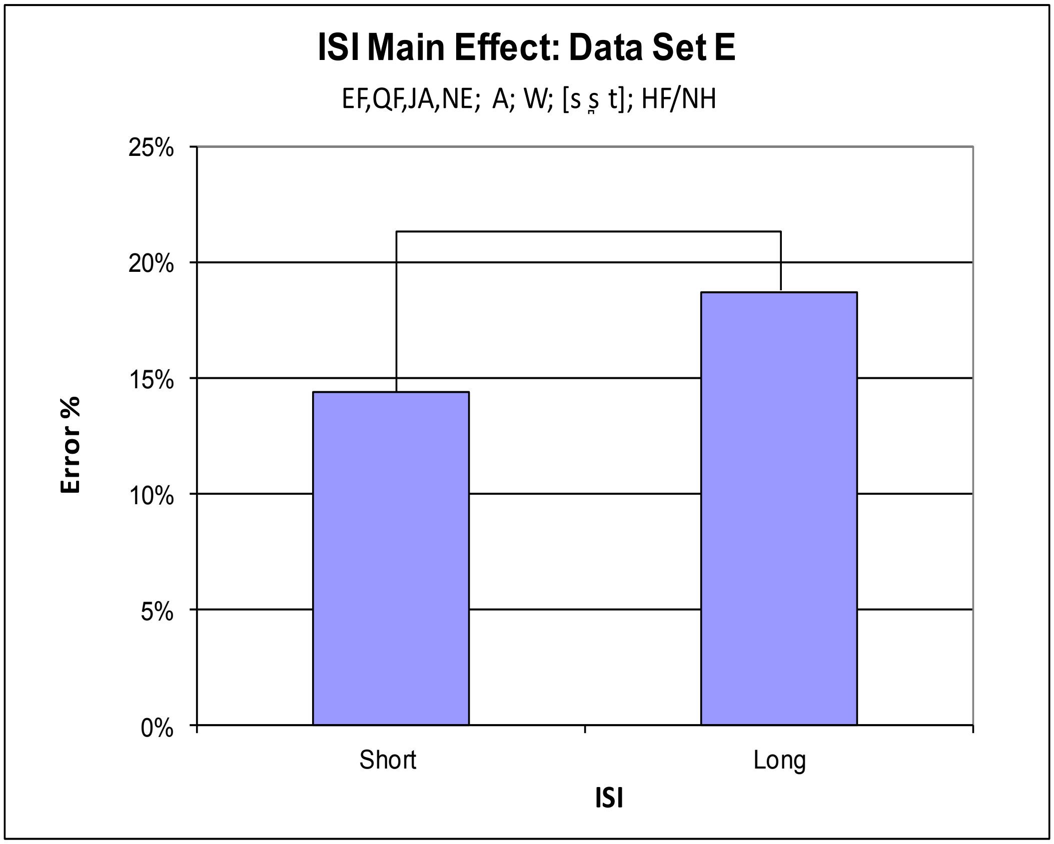

Figure 4.2. Data Set E. ISI Main Effect.
Significant differences are indicated with a connector line.

Data Set E again shows more errors in the Long condition as opposed to the short condition, as expected. Once more, however, since there were no interactions between ISI and Language or Contrast, this finding does not support the hypothesis that the Long ISI taps underspecified phonemic representations and the Short ISI taps phonetic representations. If that were the case, we would expect cross-linguistic differences due to different underlying inventories, which would show up in interactions.

Data Set F. (EF/QF/JA; BI/A; Short/Long; W; [s s̪ t]; HF/NH). This data set is similar to Data Sets D and E above; however, here both levels of proficiency are investigated.

The main effect of ISI is significant in the Anova (Anova $F(1,48), p<.05$), as is ISI x Proficiency. However, Mann-Whitney tests do not identify any significant relationships for the interaction at the .005 level.

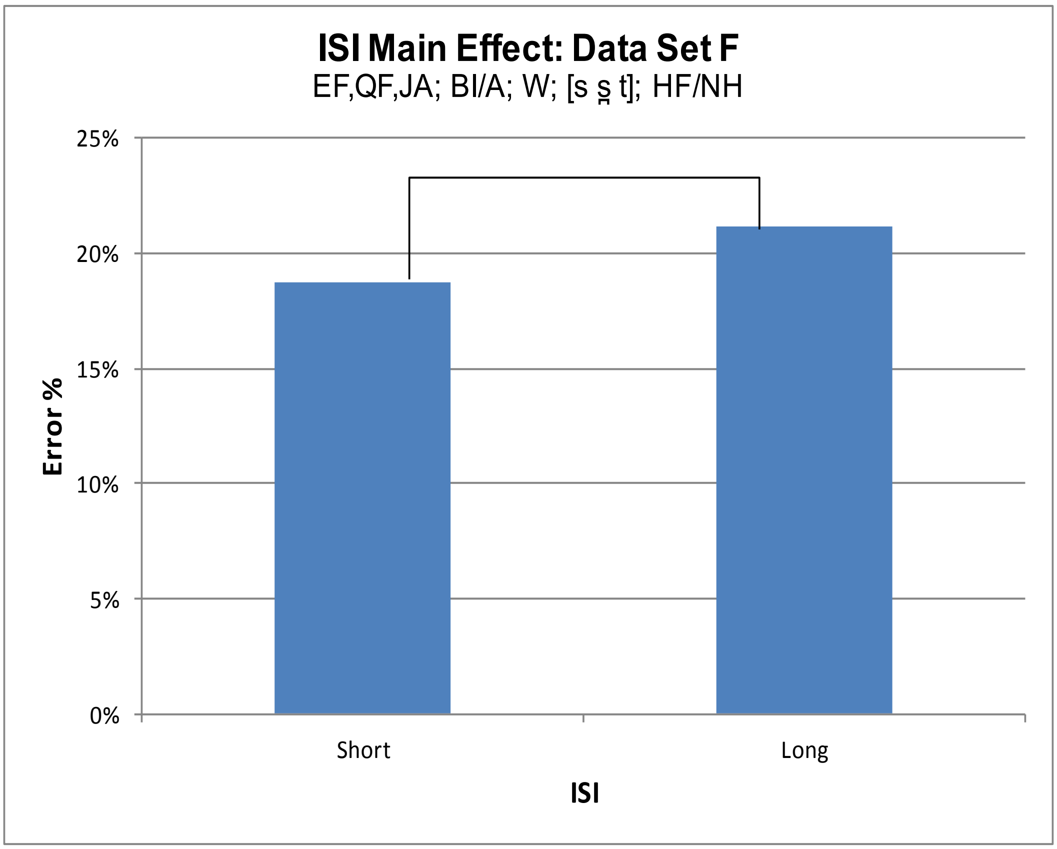

Figure 4.3. Data Set F. ISI Main Effect.
Significant differences are indicated with a connector line.

4.7.1.1. Summary of Data Involving ISI

Data Sets C, E, and F show a main effect for ISI with more errors in the Long ISI condition as compared to the Short ISI condition. Recall that the results in this section are all based on real words.

These results suggest that features fade in short term memory with an increased time delay. However, they fail to show that the Short ISI taps phonetic representations and that the Long ISI taps phonemic representations, or at least not in the way predicted by these theoretical constructs (where only contrastive features are available in the phonemic mode as compared to the phonetic mode). Had the predictions held, then we would have expected the EF, QF, and JA groups to do better on the STRIDENT vs. MELLOW contrasts ([s] vs. [θ] and [s̪] vs. [θ]) in the Short ISI compared to the Long ISI because these features are non-distinctive in these languages, and therefore predicted to be perceptible only in the phonetic condition. Instead of finding language-specific differences across ISI conditions, the results from the present experiment indicate an across-the-board decline in perceptibility with a longer ISI. We would have expected more language-particular effects if the two intervals in the AXB task were tapping distinct levels, i.e. phonetic vs. phonological.

Because there is no interaction with other factors such as Contrast and/or Language, it seems that phonemic features fade along with phonetic features with longer time-spans between stimuli.

4.7.2. Statistical Results for Proficiency

Only EF, QF, and JA could be analyzed for the effect of proficiency on error rates. This is because, for RU, only data on Beginner/Intermediate learners was collected, and, for the control group, NE, of course, there is only one level of proficiency. The following data sets provide information on main effects and interactions involving Proficiency level.

Data Set C. (EF/QF/JA; BI/A; Short/Long; W; [f s s̺ t t̺]; NH). For this data set, the Anova shows a significant main effect for Proficiency (Anova F(1,48),p = .001). This is also significant in a Mann-Whitney test (X^2(1),p = .001). There are no significant interactions involving Proficiency. Instead, we see an increased error rate across the board for the Beginner/Intermediate learners as compared to the Advanced learners, as expected.

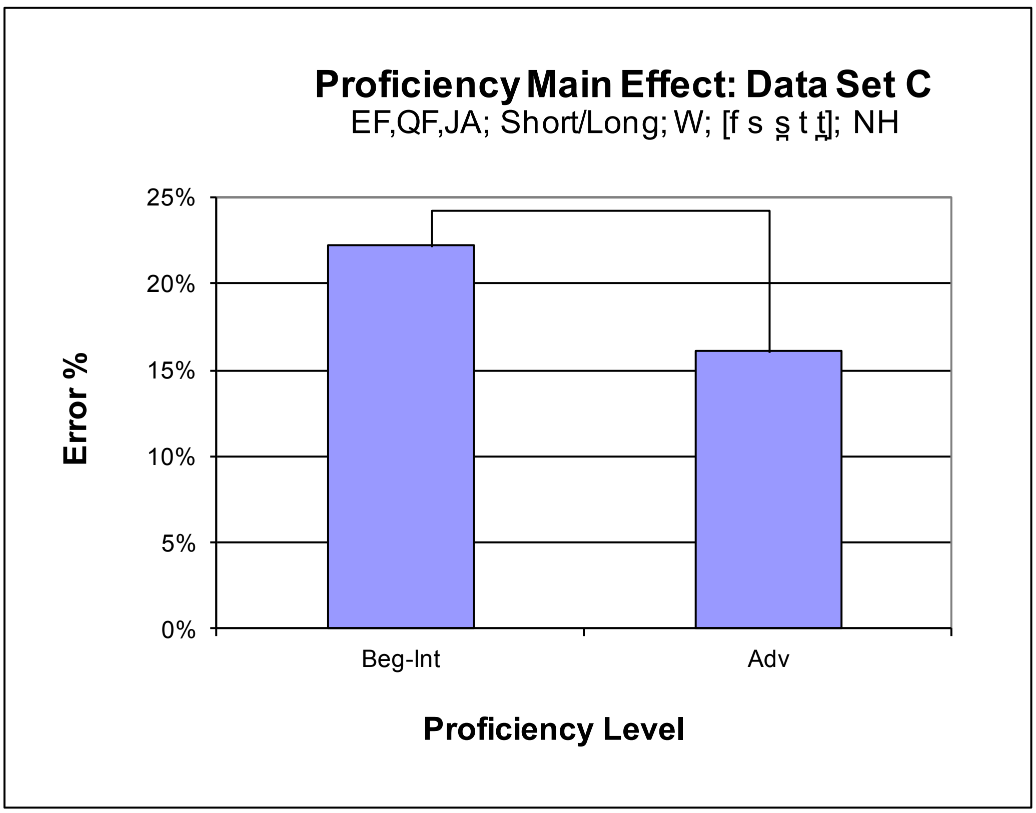

Figure 4.4. Data Set C. Proficiency Main Effect.
Significant differences are indicated with a connector line.

Data Set F. (EF/QF/JA; BI/A; Short/Long; W; [s s̩ t]; HF/NH). The Anova here indicates a significant Vowel x Language x Proficiency interaction, (Anova $F(2,48), p<.03$). Independent Sample T-Tests show that the significant differences occur between languages on the HF vowel ($p<.005$).[66] The Beginner/Intermediate JA group has higher error rates in the HF vowel condition than do QF-BI, EF-A, QF-A, and JA-A. In addition, the JA-A group has more errors in the HF condition than does QF-BI.

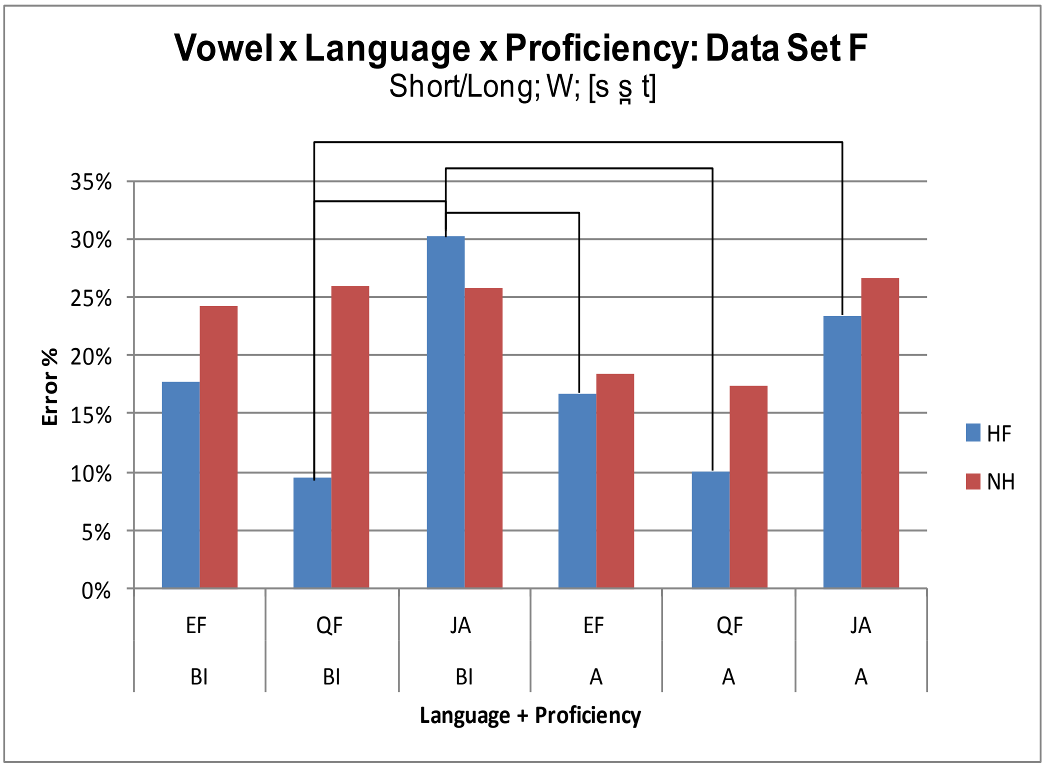

Figure 4.5. Data Set F. Vowel x Language x Proficiency.
Significant differences are indicated with a connector line.

[66] T-tests (equal variances not assumed) were used here instead of non-parametric tests because this set of data did not violate normality.

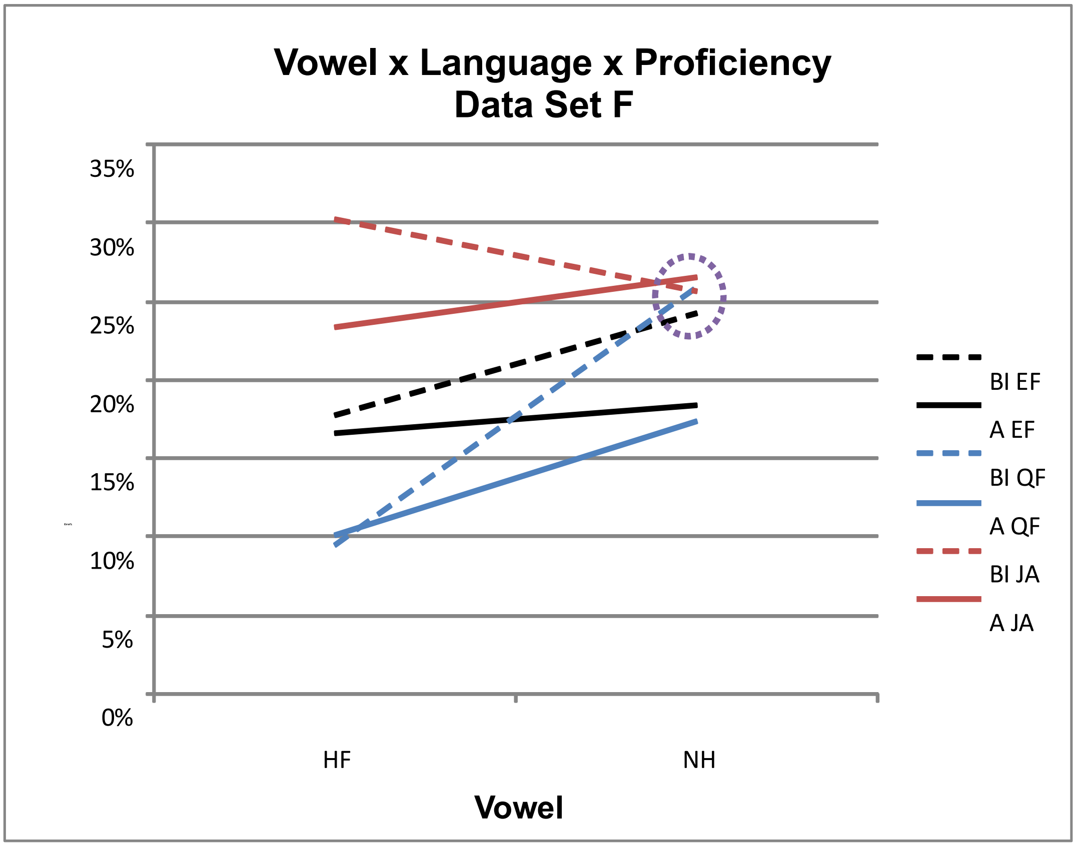

Figure 4.6. Data Set F. Vowel x Language x Proficiency.

Data trends are shown in Figure 4.6. We see here that while the other language groups at both proficiency levels tend towards a higher error rate in the NH condition as compared to the HF condition, only Japanese BI shows a lower rate.

Data Set G. (EF/QF/JA; BI/A; Short; NW/W; [f s s̪ t t̪]; NH.) This data set shows a Contrast x Wordhood x Proficiency interaction, (Anova, $F(4,200), p<.04$). Mann-Whitney tests shows the significance to lie with the Contrast [f-θ] in the real Word condition, with more errors for the BI group than the A group ($X^2(1), p<.001$), as shown below in Figures 4.7 and 4.8.

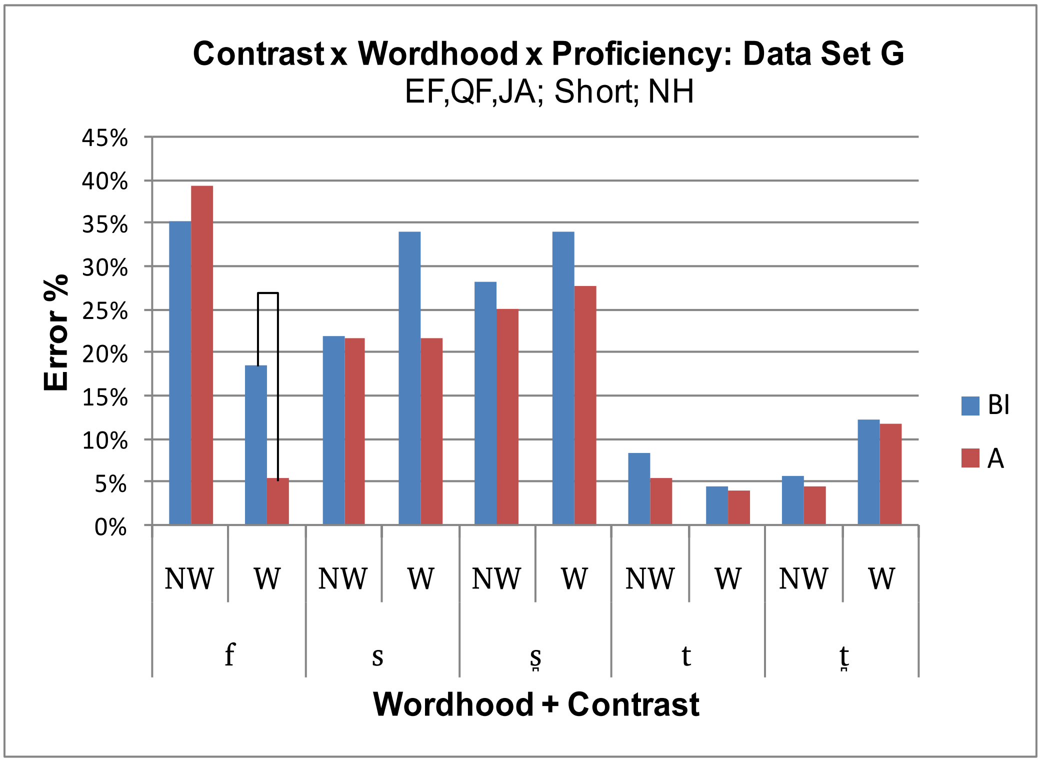

Figure 4.7. Data Set G. Contrast x Wordhood x Proficiency.
Significant differences are indicated with a connector line.

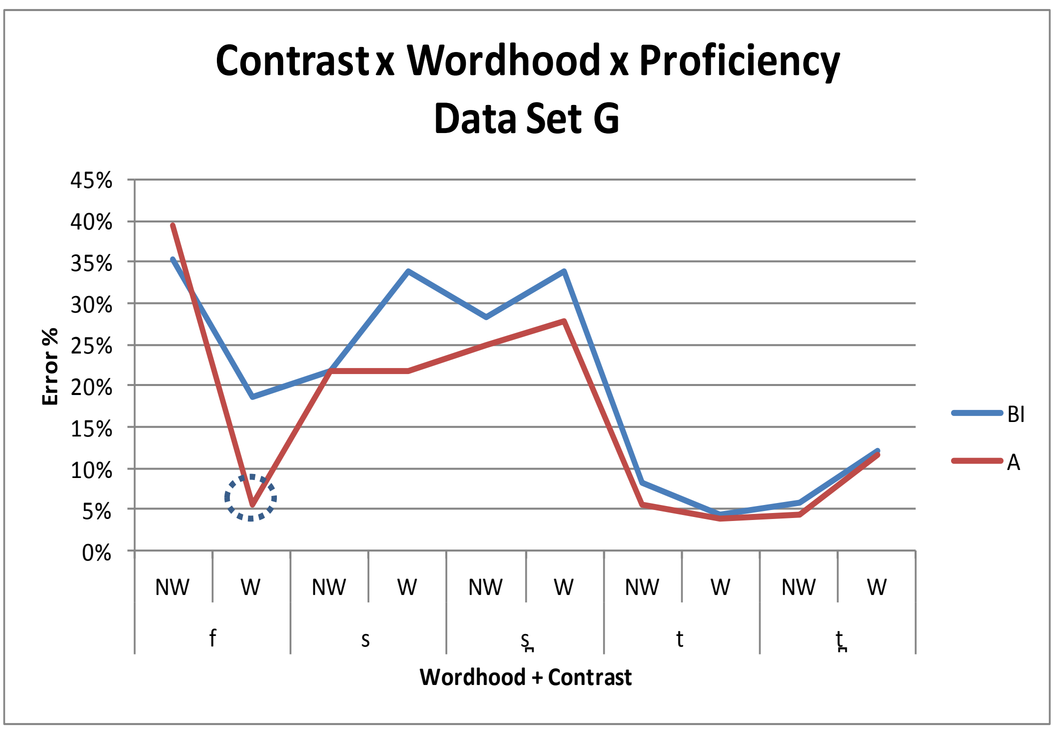

Figure 4.8. Data Set G. Contrast x Wordhood x Proficiency.

In the interactions here, we see that the Proficiency effect is dependent on Contrast and Wordhood, particularly with respect to the Contrast [f-θ] in the real Word condition, with Advanced learners outperforming the Beginner/Intermediate Group. This points to an influence of experience with lexical items as opposed to non-words, especially with this difficult contrast.

Data Set H. (EF/QF/JA; BI/A; Short, NW/W; [s s̪ t]; HF/NH.) This data set is similar to Data Set G, except that here the HF vowel condition is added, and thus the contrasts [f-θ] and [t̪-θ] are absent, since these stimuli were not included in the HF condition, as discussed in §4.7. In the current data set, there is a significant main effect for Proficiency, (Anova, $F(1,50), p<.02$); Mann-Whitney, ($X^2(1), p<.02$), but no significant interactions involving Proficiency.

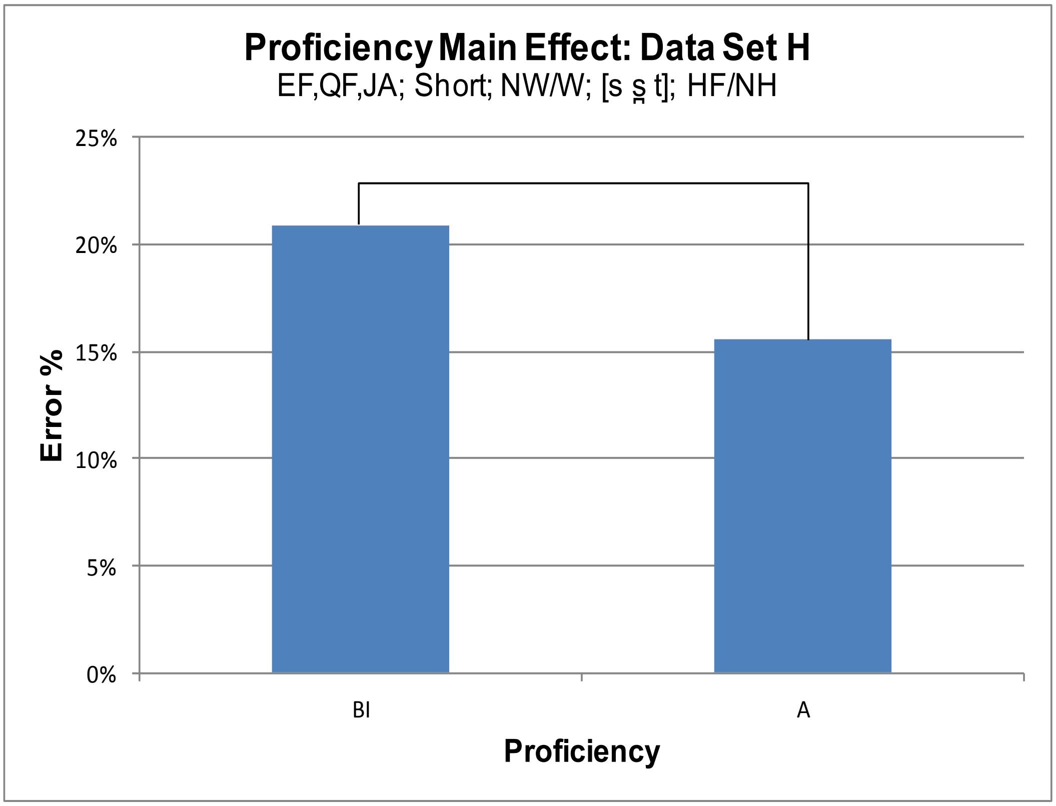

Figure 4.9. Data Set H. Proficiency Main Effect.
Significant differences are indicated with a connector line.

We see again that as learners become more proficient, their perceptual errors decrease. This is as expected.

Data Set I. (EF/QF/JA; BI/A; Short; NW; [f ɸ s s̪ t t̪ t̪s̪]; HB/NH). In this data set, an Anova was run on the Non-Word condition alone; this time all contrasts are included, but only before the vowels HB and NH, since there were no stimuli involving [t̪s̪] before HF.

Here all factors interact, giving a significant Contrast x Vowel x Language x Proficiency interaction, (Anova F(12,300),p=.004). Mann-Whitney tests show the differences to lie between EF-BI and QF-BI on Contrasts [t̪s̪-θ] before HB and [t-θ] before HB and NH; between EF-A and QF-BI on Contrasts [ɸ-θ] before HB and [t-θ] before HB and NH; between

EF-A and QF-A on Contrasts [s-θ] and [t-θ] before NH; between QF-BI and QF-A on Contrast [ɸ-θ] before HB, with higher error rates for the BI group; between QF-BI and JA-BI on Contrasts [ɸ-θ] before HB and [t-θ] before HB and NH; between QF-BI and JA-A on Contrast [t-θ] before HB and NH. For clarity, this interaction is presented in separate graphs.

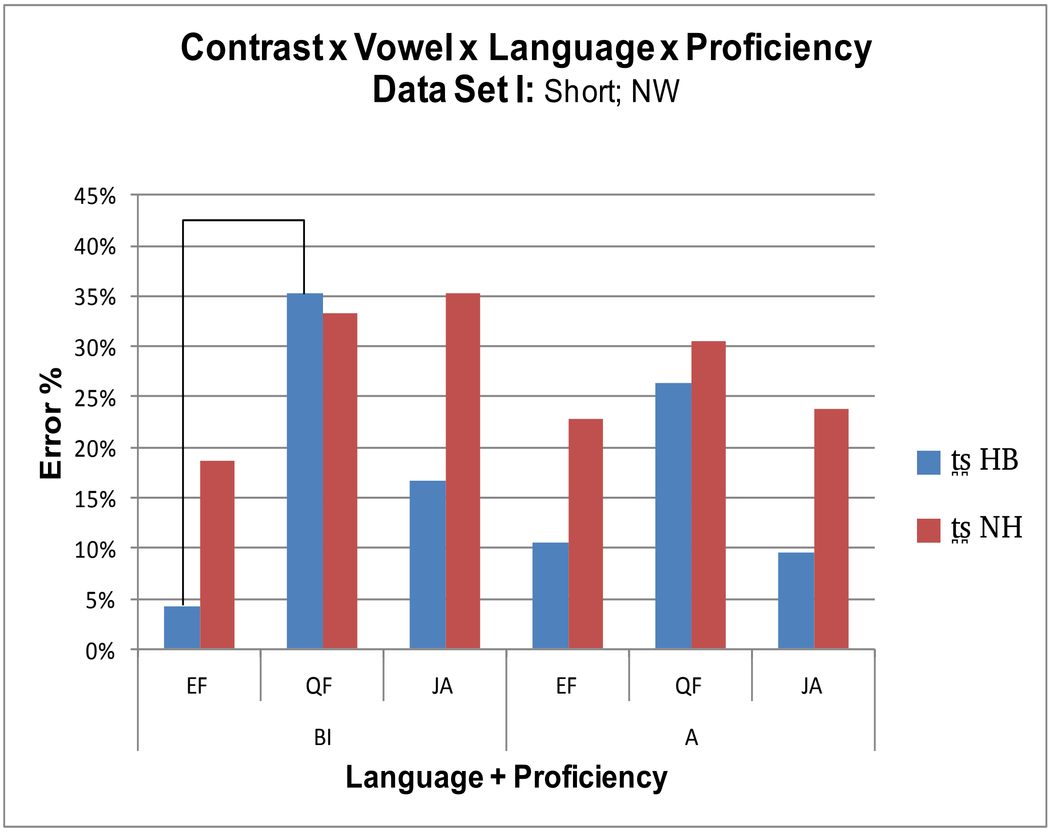

Figure 4.10. Data Set I. Contrast x Vowel x Language x Proficiency. Significant differences are indicated with a connector line.

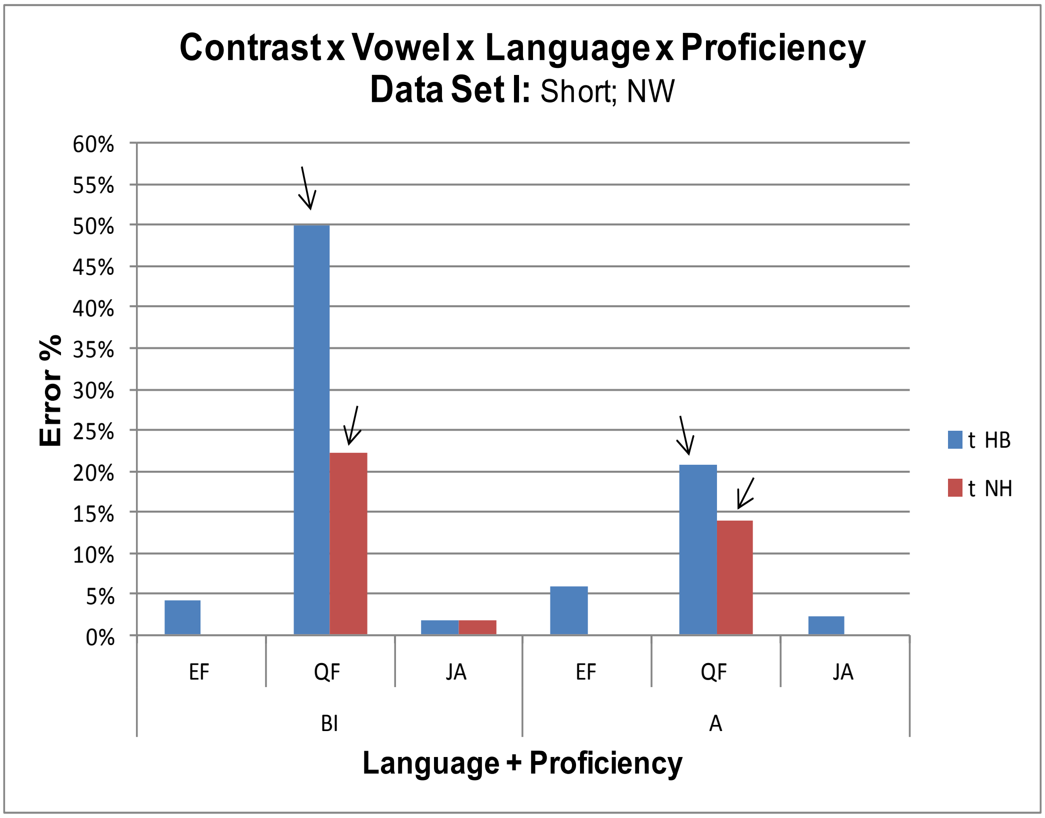

Figure 4.11. Data Set I. Contrast x Vowel x Language x Proficiency.
Arrows show source of significance.

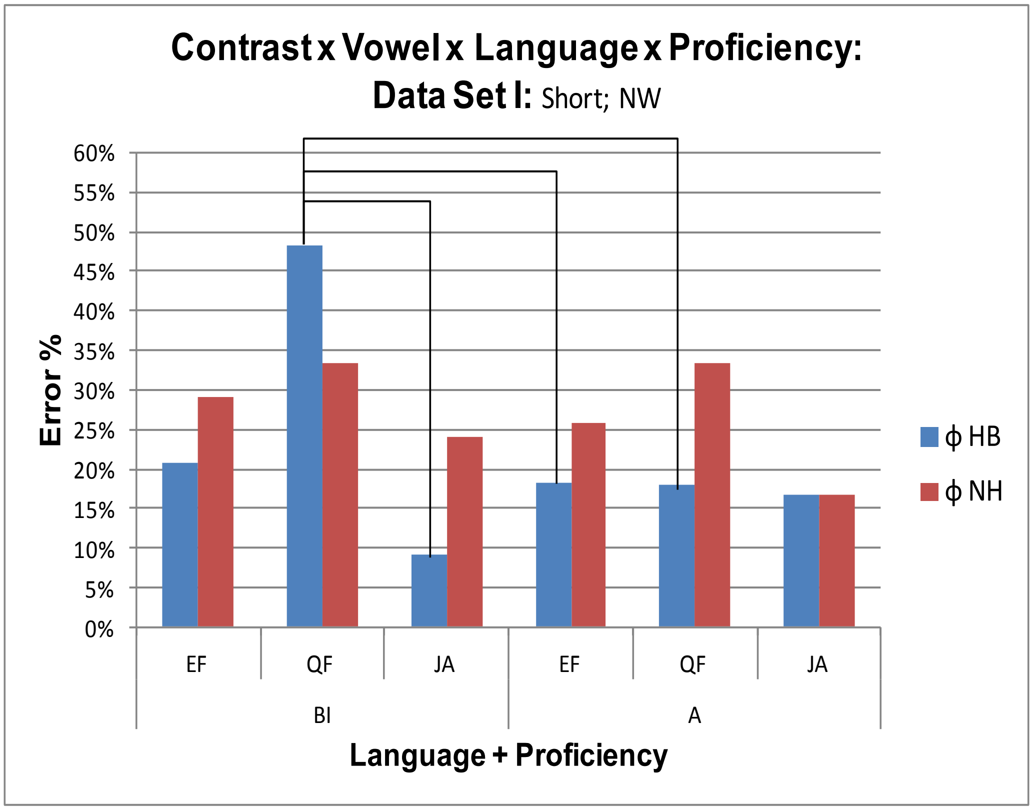

Figure 4.12. Data Set I. Contrast x Vowel x Language x Proficiency. Significant differences are indicated with a connector line.

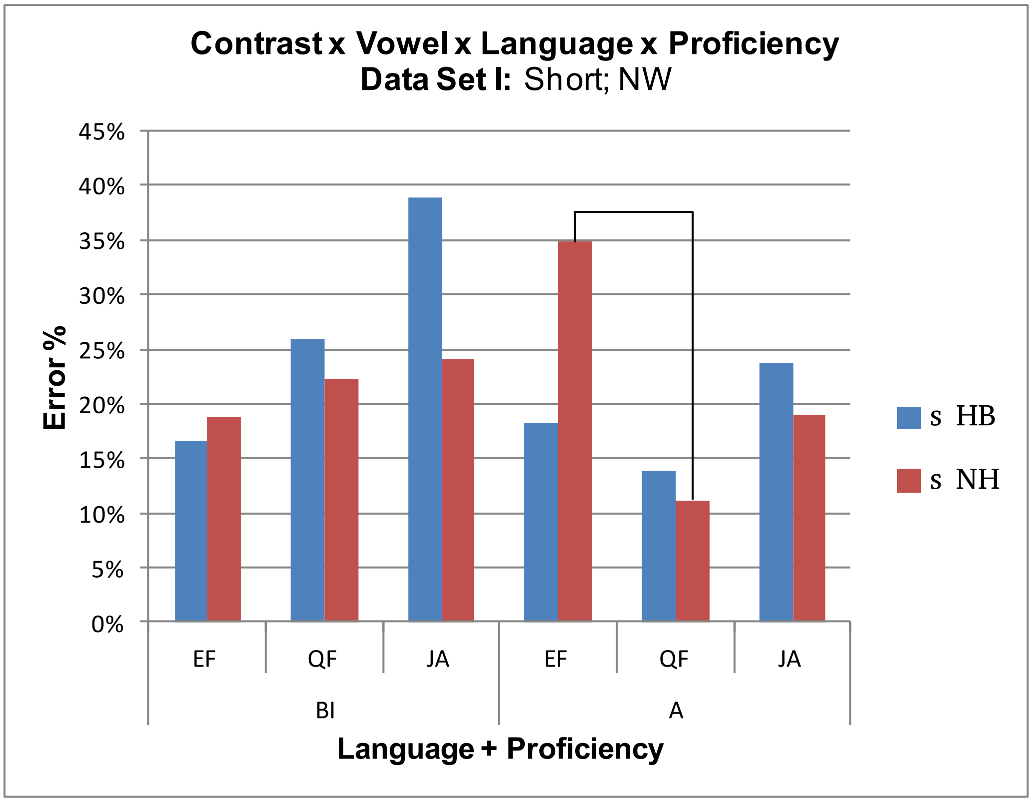

Figure 4.13. Data Set I. Contrast x Vowel x Language x Proficiency. Connector lines show significant differences.

As shown in Figures 4.10-4.13 above, there is a difference in error rate between languages, but this is a function of Proficiency level. Note that the QF-BI group have higher error rates overall as compared to the EF-BI group. This may suggest that in general, the QF-BI participants were less advanced than the EF-BI participants. If so, this is likely due to the fact that most of the EF-BI group were in Canada as part of an exchange program. In order to qualify for this exchange program, they had to achieve a certain level of proficiency in English, and as such, were a more homogeneous group than the QF-BI participants.

Data Set I also reveals inter-language differences on certain contrasts, but these are a function of vowel and proficiency level. On the [t̪s̪-θ] contrast, we see that QF-BI listeners have considerably more errors than

EF-BI listeners. This may be related to the fact that QF has [t̪s̪] as an allophone of /t/, but that neither the HB nor NH vowel is the allophonic context for the QF affricate. It may be this mismatch between allophonic variant and context which is causing the difficulty for the QF group. In a similar vein, it is interesting to note that the JA listeners, for whom [ts] is an allophone of /t/ before high back [ɯ], have more facility on this contrast before a HB vowel, a context which resembles that of the native JA conditioning context for [ts], than before a NH vowel, which is not a conditioning context for this allophone.

This data set reveals the relative difficulty that the QF group has with the [t-θ] contrast, especially for the Beginner/Intermediate learners, and notably in the context of a HB vowel. Recall the lack of such a finding in the AXB-1 experiment, where only Advanced QF learners were tested. Thus it seems that the perceptual basis for the stop substitute reported in QF production more strongly emerges when less advanced learners are investigated. Nonetheless, even for the Advanced groups in the present study, we see significant differences between QF and EF on [t-θ], something that was not seen in AXB-1.

In Chapter 3, we saw that learners tended to perceptually confuse [ɸ] with [f], and [f] with [θ]; and again this emerges in the current data set. However, it is unclear why the QF-BI group experiences more difficulty than do the other language groups on the [ɸ-θ] contrast in the context of a HB vowel, other than through an appeal to this group being somewhat less proficient than their counterparts.

It is notable that among language groups, the QF-A group has the best performance on the [s-θ] contrast. This suggests that they are most able to make use of the STRIDENT-MELLOW contrast. The significant difference between EF-A and QF-A on this contrast supports the hypothesis that for EF, [s-θ] is a *Category-Goodness Assimilation*, while for QF, it is a *Two*

Category Assimilation. Compare the relatively good performance for the QF listeners on [s-θ] with their relatively poor performance on the [t̪s̪-θ] contrast: this suggests that different cues are being attended to for [s] vs. [t̪s̪], and as discussed above, the allophonic status of [t̪s̪] may be an impediment.

Data Set J. (EF/QF/JA; BI/A; Short; NW; f ɸ s s̪ t t̪]; HB/HF/NH). This data set is similar to Data Set I except that the High Front vowel context is also investigated, and as such the contrast [t̪s̪] is absent, since there were no stimuli constructed for this contrast in the HF condition. In the current data set, again there is an interplay between all factors for a significant Contrast x Vowel x Language x Proficiency interaction, (Anova $F(20,500), p < .03$). The results from Data Set I hold, and the only additional significant result that emerges by adding the HF condition is that the QF-BI group has fewer errors than JA-BI on [s̪-θ] in this vowel context. In the interests of clarity, only the relevant portion of this complex interaction is shown.

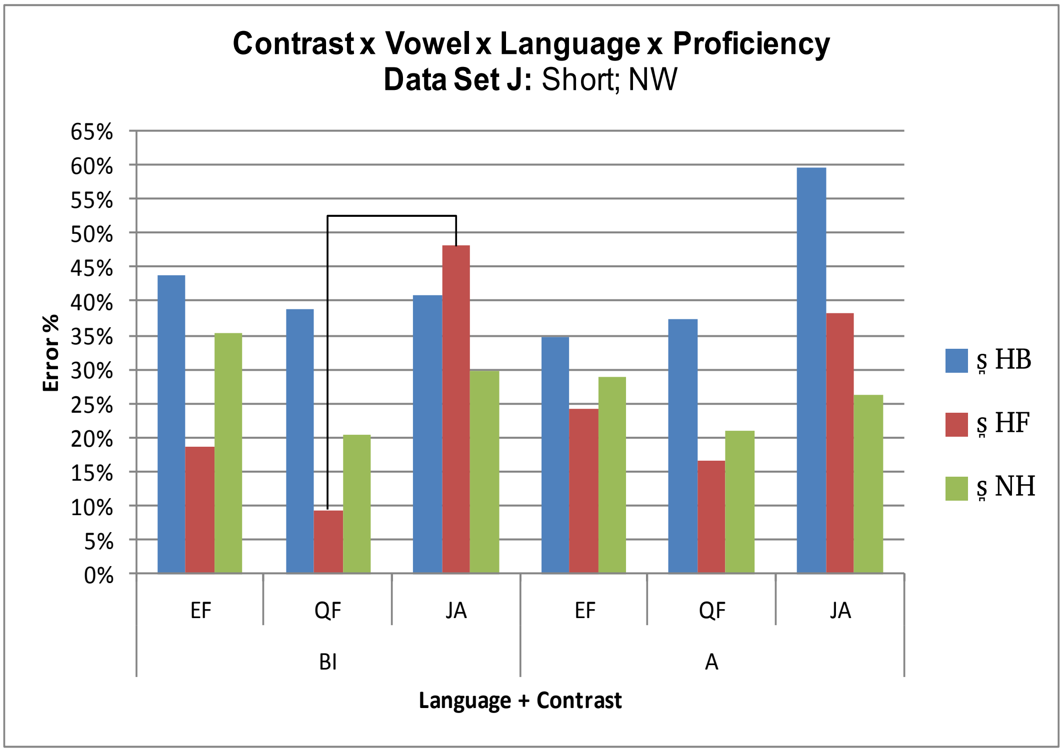

Figure 4.14. Data Set J. Contrast x Vowel x Language x Proficiency [s̪-θ].
Connector lines indicate significant differences.

The main finding for Data Set J is that the JA-BI group experiences considerably more difficulty on the [s̪-θ] contrast as compared to the QF-BI group. This is consistent with the hypothesis that production mirrors perception in differential substitution. It is noteworthy that the JA-BI group has more trouble with the contrast in the context of a HF vowel than with other vowels. This may be related to the presence of an inappropriate allophone in this vowel context. Recall that Japanese has an allophonic process whereby /s/ becomes [ɕ] before the vowel [i]. In an AXB task, Munoz-Sanchez (2003) found that JA listeners had more difficulty with [s-θ] in the context of [i] than in the context of [a]. She suggests that when a contrast involves a non-native sound, in this case [θ], and the allophonic context is wrong for the other sound, here [s], then the listeners are unable to extend their allophonic knowledge. Thus, it may be that the unfamiliar sequence [s̺i] poses a challenge to this language group.

4.7.2.1. Summary of Data Involving Proficiency

The analyses in this section investigated both Beginner/Intermediate and Advanced learners of English, and they indicate an overall improvement in the Advanced learners versus the Beginner-Intermediate learners. This is an expected result, and lends credence to the position that the participants were correctly categorized by the proficiency tests.

The Contrast x Wordhood x Proficiency interaction is due to the Contrast [f-θ], with significant improvement in the Advanced group versus the Beginner-Intermediate group, but only in the real Word condition. In the Non-Word condition, there is no statistical improvement between the BI and A groups. A similar, but non-significant, pattern between words and non-words is seen for [s-θ] and [s̺-θ], but not for the contrasts involving stop consonants. These results therefore suggest that increased lexical familiarity facilitates the ability to distinguish between word pairs, and this effect may be more prominent for fricative contrasts.

The interactions in this section reveal that perception of contrasts depends on the vowel context, the L1, and Proficiency level. We have seen that all language groups show poorer performance for all contrasts in the NH condition as compared to the HF condition, except for JA-BI. This may be related to findings by Jongman (1989) which suggested that a low vowel inhibits perception of non-strident fricatives. The increased difficulty that the JA group experiences before a HF vowel may be related to this group's lack of experience with [si], [s̺i], and [ti] due to the JA allophonic processes /s/ → [ɕ]/_[i] and /t/ → [tɕ]/_[i]. That this difficulty is diminished for the JA-A group may be an indication that learners can overcome problems due to L1 allophonic processes.

We have observed that the Québec French listeners have significantly more errors than the other language groups on [t-θ], particularly in the case of the Beginner/Intermediate groups, but also between the Advanced Québec and European French groups in the context of a Non-High vowel. This inter-language difference supports the hypothesis that differential substitution has its roots in perception. Thus, we see that by testing Beginner/Intermediate learners of English, perceptual differences between Québec and European French begin to emerge.

On [ɸ-θ], the QF-BI participants have the highest error rate. For the Advanced group, it may be that increased familiarity with [θ] has led to improved performance on [ɸ-θ]. It is interesting to note that JA-BI listeners have the lowest error rates on [ɸ-θ]. There is a tendency towards better performance before a high back vowel, a compatible context for the native allophone [ɸ]: in Japanese /h/ → [ɸ]/_[ɯ]. We see higher error rates before a non-high vowel, an incompatible context for this allophone. This is similar to the situation with [t̪s̪] discussed above. This difference disappears for the Advanced Japanese group, perhaps due to increased exposure to the [f] plus non-high vowel sequence in English.

Another significant finding in this analysis is that Advanced European French listeners have more difficulty than their Québec French counterparts on [s-θ] before a Non-High vowel. Thus here again, we see perceptual differences between these two dialects of French which follow the direction of their production differences.

In summary, the data presented in this section show improvement in the perception of contrasts involving the interdental fricative for all languages tested. There are indications that improvement is aided by wordhood status: real words appear to have a facilitating effect for certain contrasts. In addition, it is speculated that the existence of allophonic processes in the L1 may inhibit the perception of contrasts where there is a conflict between allophone and contextual environment. On the other hand, there may be a facilitating effect where the allophone and contextual environment agree. Finally, in examining less proficient Québec French listeners, we see perceptual differences emerge between the two French dialects.

4.7.3. Statistical Results for Wordhood

Some Wordhood effects have already been discussed in other interactions; this section looks at the remainder. In this section, only those contrasts which occur in both the Word and Non-Word conditions were examined. These contrasts are: [f-θ]; [s-θ], [t-θ], [s̪-θ], and [t̪-θ]. Also, only those vowel contexts that allowed for frequent real words are included in the Word condition. This means that for [f-θ] only NH was used. Also, for [t̪-θ], only NH was used. The other contrasts occur before both NH and HF vowels. Finally, only the Short ISI condition was subjected to statistical analysis of Wordhood, since in the Long ISI, only real words were used.

The results from analyzing Wordhood and interactions with this factor will indicate whether the lexical status of words affects perception.

Data Set G. (EF/QF/JA; BI/A; Short; NW/W; [f s s̪ t t̪]; NH.) This analysis revealed a Contrast x Wordhood x Proficiency interaction which was discussed in the previous section entitled Proficiency.

Data Set H. (EF/QF/JA; BI/A; Short, NW/W; [s s̪ t]; HF/NH.) There is no main effect or interactions for Wordhood in this analysis.

Data Set K. (EF/QF/JA/RU/NE; BI; Short; NW/W; [f s s̪ t t̪]; NH). This data set shows a significant Contrast x Wordhood x Language interaction, (Anova, $F(16,156), p < .001$). Mann-Whitney tests also find that the differences between languages are dependent on Wordhood and Contrast. The source of significance here is the [t-θ] contrast in the NW condition alone, with the QF group having more errors than all other language groups. The relevant result is shown in Figure 4.15.

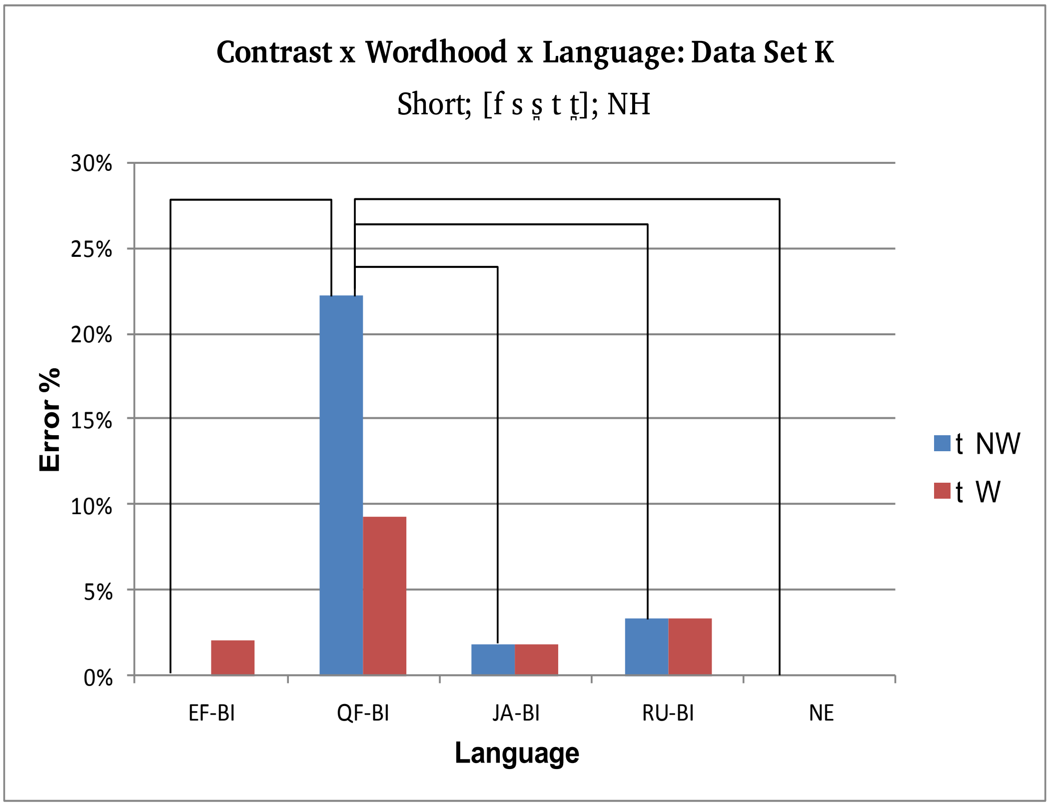

Figure 4.15. Data Set K. Contrast x Wordhood x Language [t-θ] . Connector lines indicate significant differences.

The poorer performance of the QF-BI group on [t-θ] in the Non-Word condition demonstrates that real words appear to be somewhat easier to discern for this group in this particular context, which again would suggest that lexical experience facilitates perception. For the other languages and contrasts, this effect is not seen, with perhaps the exception of [f-θ], although the latter does not reach significance. It is unclear why Wordhood would have an influence on only [t-θ] and [f-θ].

Data Set L. (EF/QF/JA/NE; A; Short; NW/W; f s s̪ t t̪]; NH). This data set is the same as Data Set K, except this time advanced listeners are examined. Results here indicate a significant Contrast x Wordhood interaction, (Anova, $F(4,136), p < .001$).

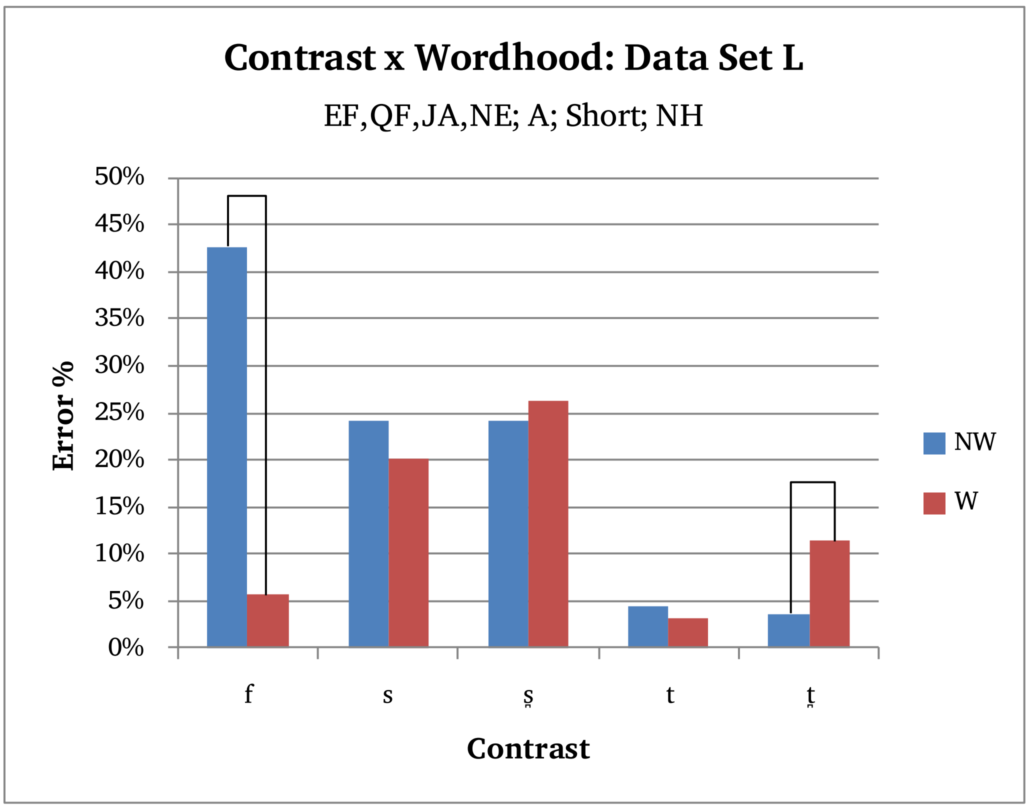

Figure 4.16. Data Set L. Contrast x Wordhood.
Connector lines indicate significant differences.

Wilcoxon tests show the significance to lie within the Contrast [f-θ], with higher error rates in the Non-Word condition, and within Contrast [t̪-θ], with higher error rates in the Word condition.

When all languages are considered together here, there is a huge effect for Wordhood for the Contrast [f-θ].

Data Set M. (EF/QF/JA/RU/NE; BI; Short; NW/W; [s s̪ t]; HF/NH). The main effect for Wordhood is not significant in this analysis in either the Anova or the Wilcoxon test. Neither are there any significant interactions involving Wordhood.

Data Set N. (EF/QF/JA/NE; A; Short; NW/W; [s s̪ t]; HF/NH). This data set is the same as Data Set M, except that here we investigate the advanced participants.

The Anova for main effect for Wordhood in this analysis is not significant. It does show a significant Contrast x Wordhood interaction, (Anova $F(2,68), p<.04$); however, this fails to reach significance in Wilcoxon posthoc tests.

4.7.3.1. Summary of Data on Wordhood

The data examined in this section have shown that the perception of contrasts is dependent on the L1 and whether the stimuli are real words or non-words.

We observed fewer errors in the real Word condition than the Non-Word condition for [f-θ], but more errors in the Word versus Non-Word condition for [t̪-θ]. This possibly indicates that lexical status plays a role in speech perception, specifically that real words are more easily perceived than non-words, particularly for the [f-θ] pair. However it must be mentioned that for the [f-θ] contrast, the real word stimuli consisted of only one type of word pair: *first* versus *thirst*, repeated six times, while the other stimuli consisted of two or more word pairs. Thus it may be that frequency effects have aided in the discrimination of [f-θ] in the Word condition.

The data in this section have revealed an interesting situation for the QF listeners involving the contrasts [t-θ] and [t̪-θ]. The QF group has difficulty on [t-θ] in the Non-Word condition, but good performance on [t-θ] in real Words; on the other hand, the QF listeners as well as other language groups have difficulty on [t̪-θ] in the real Word condition, but good performance on [t̪-θ] in the NW condition. In other words, it seems that problems occur when the English apico-alveolar [t] is associated with a non-English word and when the non-English lamino-dental [t̪] is associated with a real English word.

A possible explanation for this is that these participants are more likely to have native-like representations for both [t] and [θ] in real English words. Thus, when presented with [t̪-θ] in real words, the interdental [θ] is correctly represented, but the non-English lamino-dental [t̪] is more likely to be confused with [θ], resulting in a *Category Goodness Assimilation*. If this interpretation is correct, it means that non-contrastive information may be available in lexical representations.

4.7.4. Statistical Results for Vowel

Data Sets D-F, H-J, and M-R serve to evaluate the remainder of the significant results involving vowel context that have not been discussed so far. Recall the vowel categories that are examined: High Front (HF) [i, ɪ], High Back (HB) [u], and Non-High (NH) [eɪ ɛ ɜ ʌ oʊ æ aɪ ɑ].

Data Set D. (EF/QF/JA/RU/NE; BI; Short/Long; W, [s s̺ t]; HF/NH). The Anova for this data set shows a significant Contrast x Vowel x Language interaction, (Anova F(8,74), p=.001). Mann-Whitney tests show that JA has significantly more errors than QF and NE on [s-θ] HF and [s̺-θ] HF. Japanese listeners also had more errors than NE on [s-θ] NH.

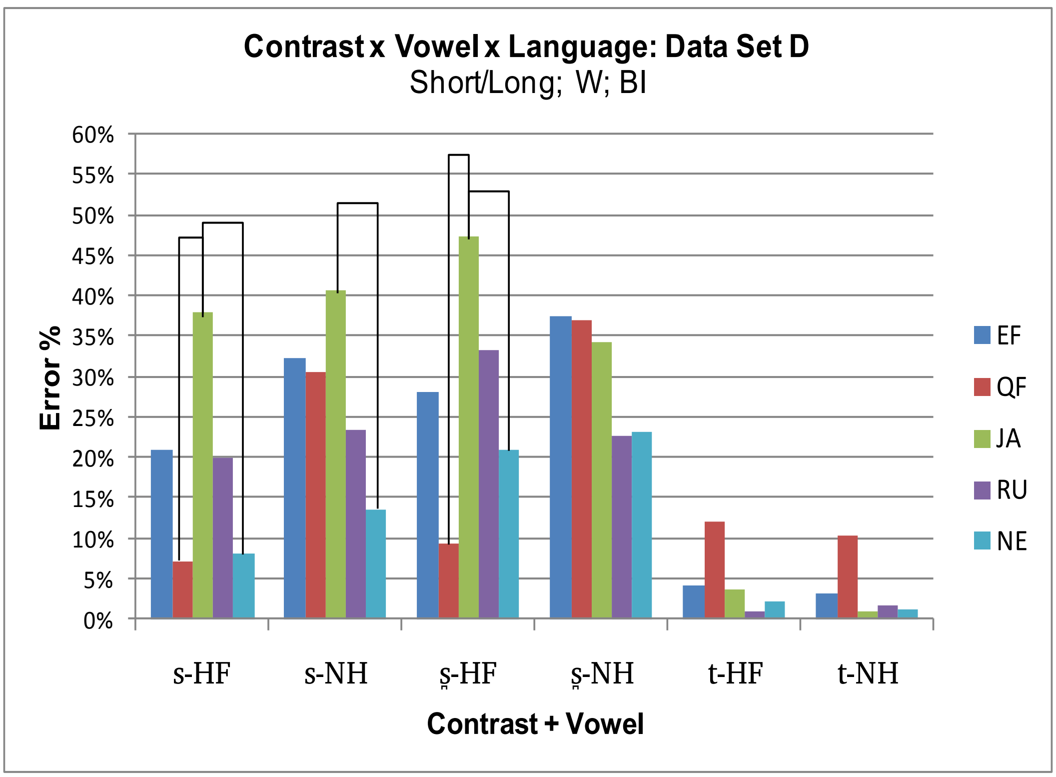

Figure 4.17. Data Set D. Contrast x Vowel x Language interaction. Significant differences are indicated with a connector line.

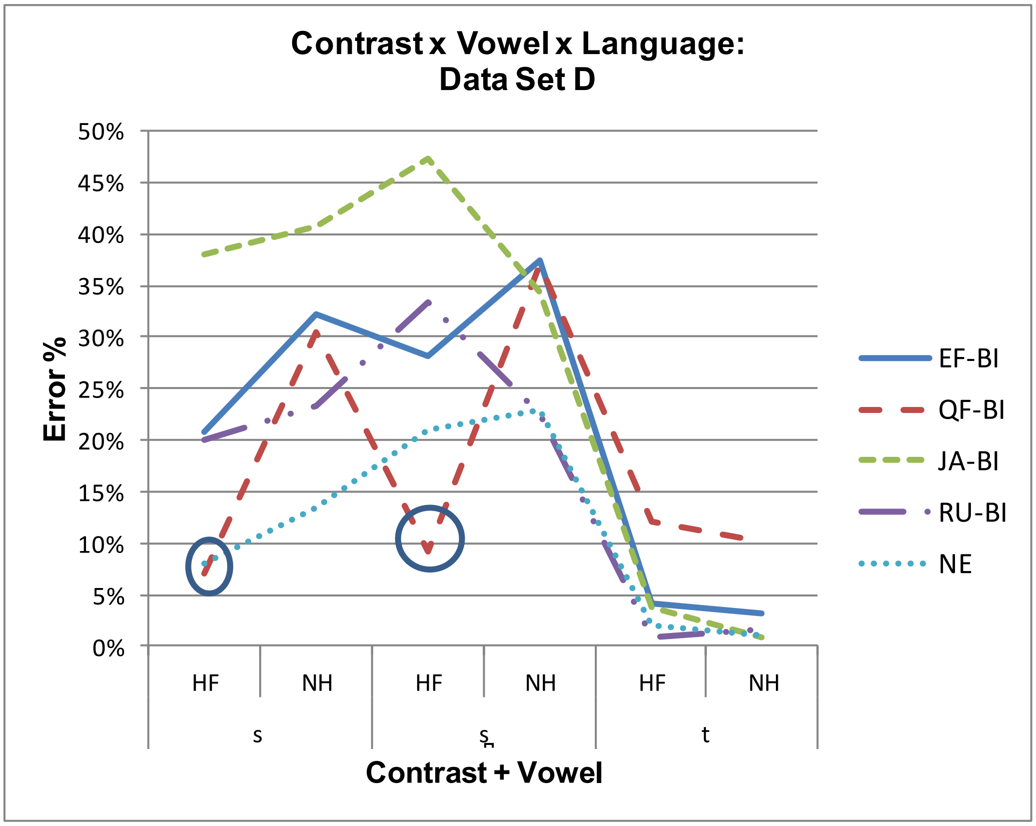

Figure 4.18. Data Set D. Contrast x Vowel x Language interaction.

Figure 4.18 illustrates the trend of the interaction. What stands out in particular is the relatively low error rates for QF on [s-θ] and [s̻-θ] in the context of a HF vowel.

These data examining BI learners' perception of the contrasts [s s̻ t] indicate that the JA listeners have particular difficulty with distinctions involving STRIDENT vs. MELLOW fricatives. This difficulty seems to be exacerbated in the context of a HF vowel for the [s̻-θ] contrast, perhaps due to the allophonic relationship in JA, as discussed earlier.

Data Set E. (EF/QF/JA/NE; A; Short/Long; W; [s s̻ t]; HF/NH). This analysis looks at the same factors as Data Set D above, but this time for the advanced learners.

For Data Set E, there is a significant main effect for Vowel, (Anova $F(1,34), p<.05$); with more errors overall in the NH condition.

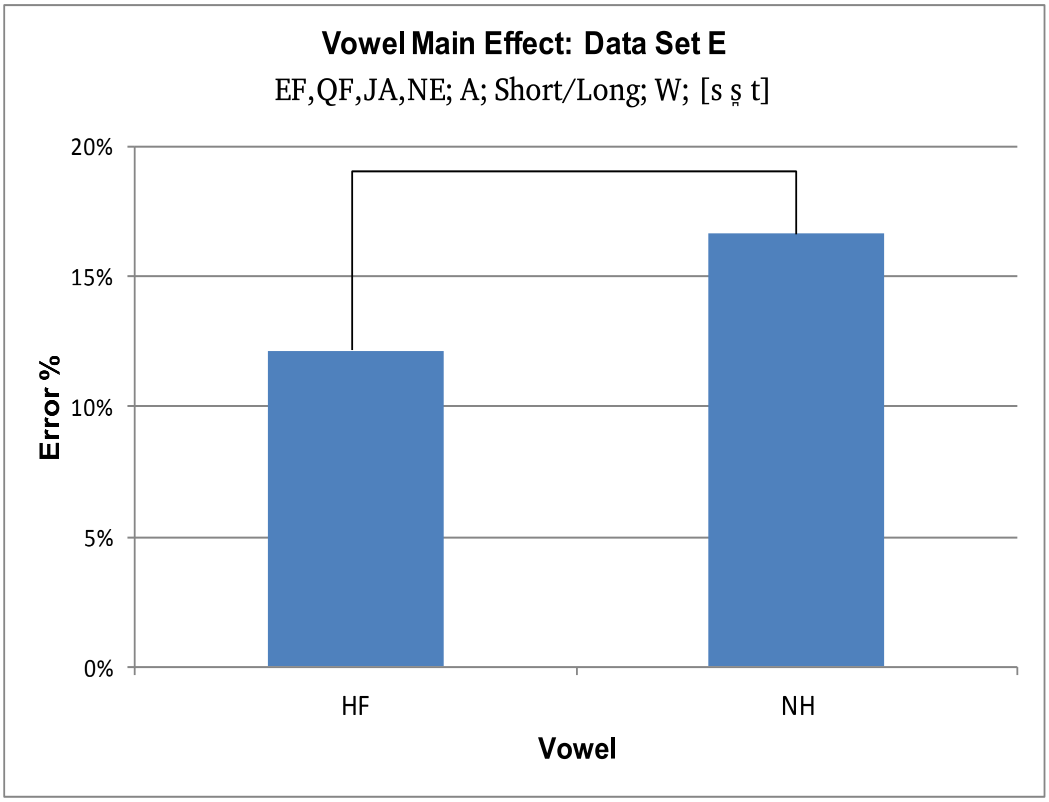

Figure 4.19. Data Set E. Vowel Main Effect.
Significant differences are indicated with a connector line.

There are no interactions with Vowel, so the difficulty before the Non-High vowel holds across languages and across different contrasts.

Data Set I. (EF/QF/JA; BI/A; Short; NW; [f ɸ s s̪ t t̪ t̪s̪]; HB/NH). The results for the data set were discussed in §4.7.2, Figures 4.10-4.13.

Data Set J. (EF/QF/JA; BI/A; Short; NW; [f ɸ s s̪ t t̪]; HB/HF/NH). This data set was discussed under in §4.7.2 on Proficiency, Figure 4.14.

Data Set M. (EF/QF/JA/RU/NE; BI; Short; NW/W; [s s̺ t]; HF/NH). Here, the Anova reveals a significant Contrast x Vowel x Language interaction, (Anova F(8,78),p=.001). Significant differences as determined by two-tailed Mann-Whitney tests are shown in Figure 4.20. The JA group has more errors than QF and NE on the Contrasts [s-θ] and [s̺-θ] before a HF vowel. The RU listeners have a higher error rate than NE on [s-θ] before a HF vowel. On the other hand, the QF group has more errors than EF on [t-θ] in both vowel contexts, as well as more than JA and NE in the context of a NH vowel.

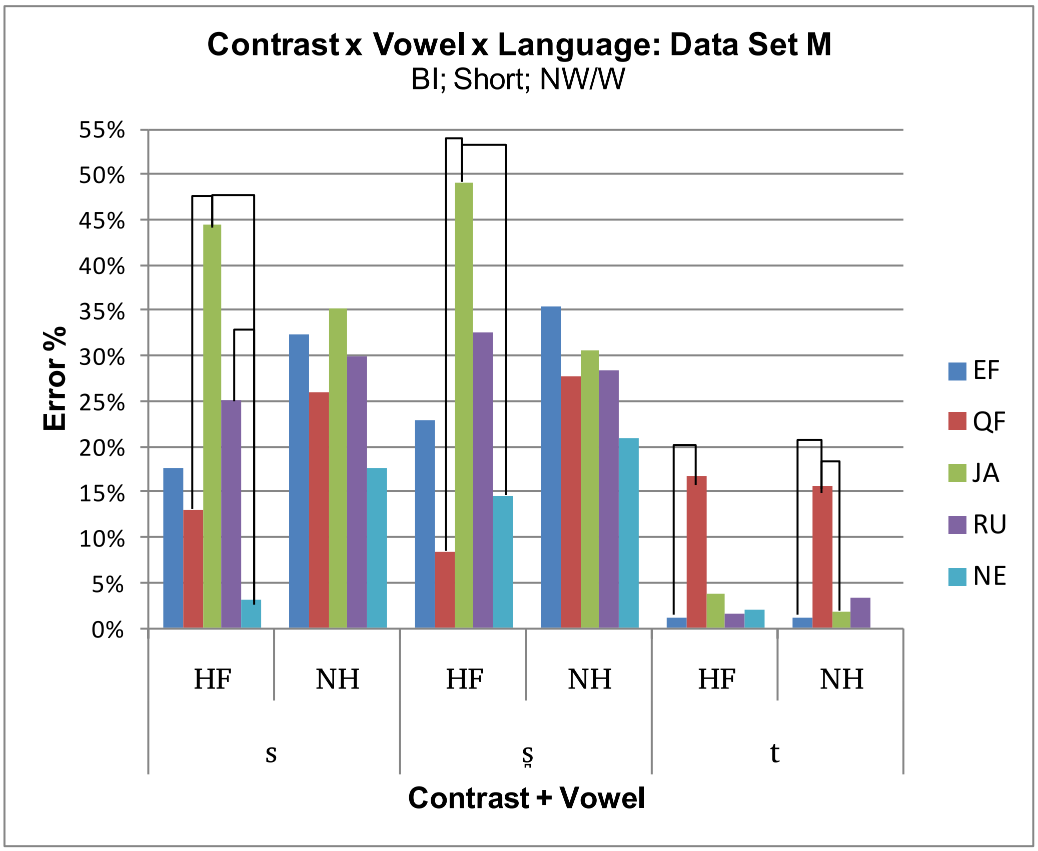

Figure 4.20. Data Set M. Contrast x Vowel x Language.
Connector lines indicate significant differences.

Thus, we again see a strong effect for vowel context, especially for the JA group, which shows that HF vowels inhibit this language group's ability to make discriminations. This effect is stemming from the fricative contrasts. This is consistent with the findings from AXB-1 and other research that finds that fricative perception is sensitive to vowel context. In addition, this may be related to the JA allophonic process in the context of the vowel [i] as discussed earlier; although it is not clear why vowel context does not seem to influence the QF listeners' ability to perceive the [t-θ] contrast before HF versus NH vowels, since QF has an allophone of [t̺] before high front vowels.

Additionally, we observe large cross-linguistic differences in the perception of these contrasts. Notably, the JA group has more difficulty on the [s-θ] and [s̺-θ] pairs. The latter in particular shows an interesting trend, with JA showing many more errors than QF and NE, while EF and RU fall in the middle. And with [t-θ], we see again that the QF group has considerably more difficulty than all the other language groups. These data support the hypothesis that perceptual difficulties underlie production errors.

Data Set N. (EF/QF/JA/NE; A; Short; NW/W; [s s̺ t]; HF/NH). This data set is the same as Data Set M, except that here we investigate the advanced participants.

These data reveal a significant Contrast x Vowel x Language interaction, (Anova $F(6,68), p < .03$).

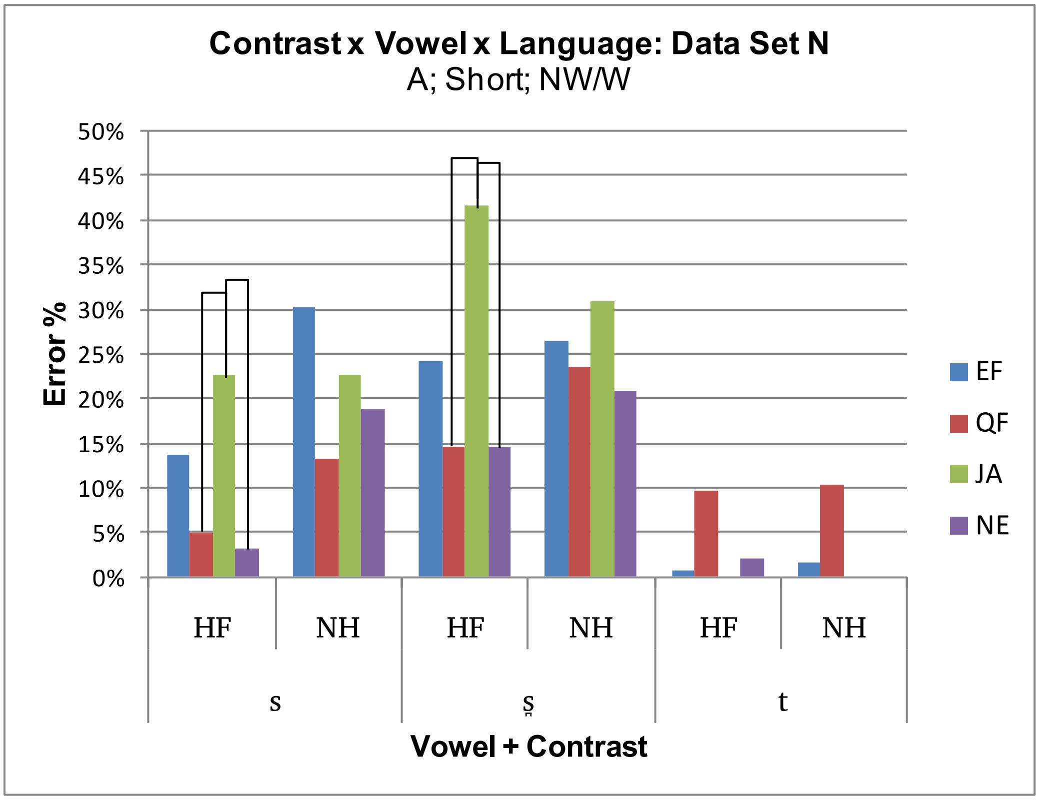

Figure 4.21. Data Set N. Contrast x Vowel x Language.
Connector lines indicate significant differences.

Two-tailed Mann-Whitney tests show the difference to lie within [s-θ] and [s̪-θ] HF, with the JA group having more errors than both QF and NE. Within [s-θ] NH, the EF group has more errors than QF.

This data set has again shown that the JA group has difficulty with contrasts before High Front vowels, and in particular, with the STRIDENT/MELLOW fricative contrasts. This indicates that QF and NE are using different auditory cues than JA when attempting to distinguish these contrasts.

Data Set O. (EF/QF/JA/RU/NE; BI; Short; NW; [f ɸ s s̪ t̪s̪ t t̪]; HB/NH.) For Data Set O, an Anova shows a significant Contrast x Vowel x Language interaction, (Anova $F(24,234), p<.02$). Two-tailed Mann-Whitney tests show the difference to be due to higher error rates for the QF group as compared to the other groups on certain contrasts in the context of certain vowels, as seen in Figures 4.22-4.24. The QF group has more errors on [ɸ-θ] HB as compared to the JA and NE groups; more errors than EF on [t̪s̪-θ] in the context of a HB vowel (Figure 4.23); and higher error rates on [t-θ] in both vowel conditions as compared to the other language groups (Figure 4.24). The relevant portions of this interaction are shown in the graphs below.

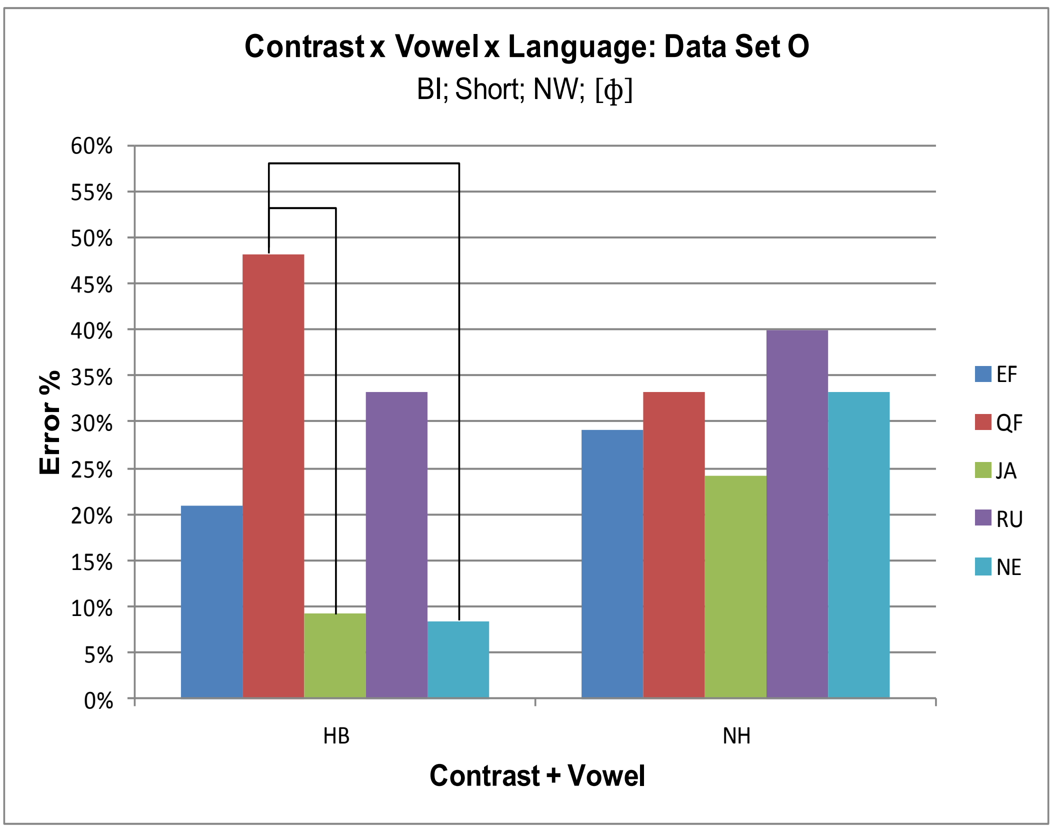

Figure 4.22. Data Set O. Contrast x Vowel x Language.
Connector lines indicate significant differences.

The QF listeners have higher error rates on [ɸ-θ] HB as compared to the JA and NE listeners. It has been discussed earlier that the relatively low errors for the JA group on this contrast may be due to their L1 allophonic experience in this vowel context.

For NE, note how the low error rate in the context of an HB vowel compares to the NH contrasts. The situation is similar for [f-θ]. It seems that a high back (rounded) vowel facilitates the discrimination of these mellow fricatives for Native English speakers. This may be because for [f] and [ɸ] in the context of a HB vowel, the articulation starts as relatively unrounded and becomes rounded as it progresses into the vowel. However, for [θ], given the lack of involvement of the lips, rounding can occur earlier during the articulation of [θ]. It is unknown why QF is unable to make use of these rounding cues for [ɸ-θ].

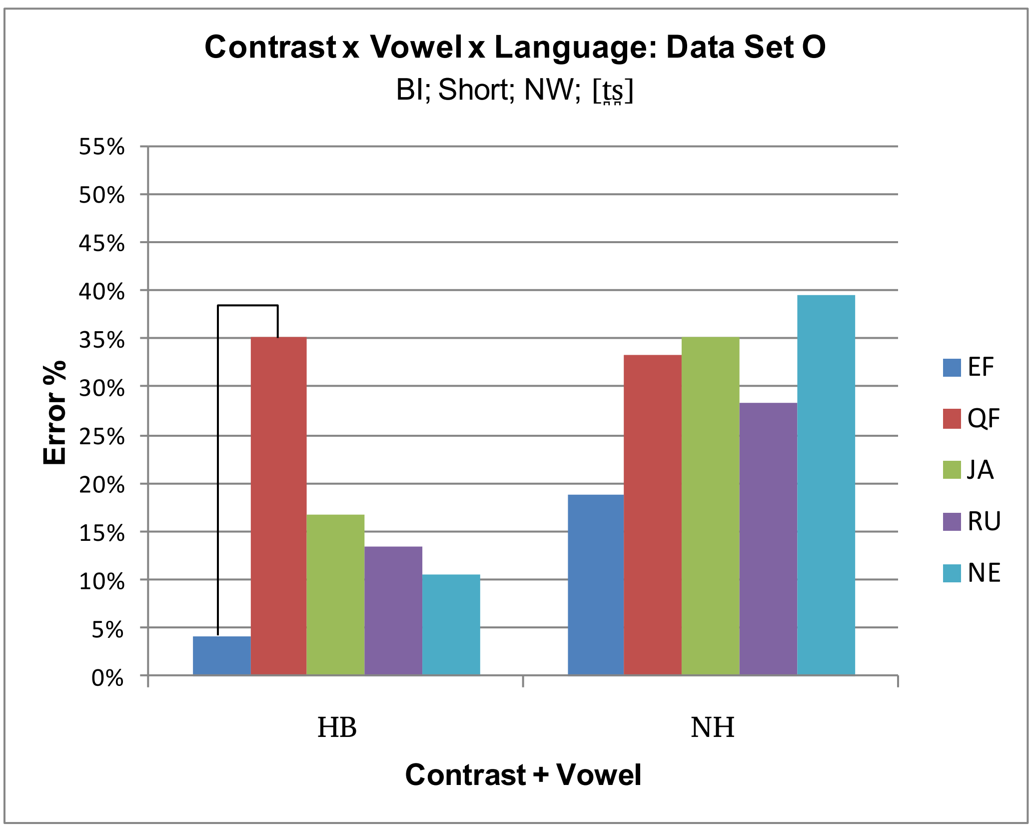

Figure 4.23. Data Set O. Contrast x Vowel x Language.
Connector lines indicate significant differences.

Figure 4.23 shows that QF has more errors than EF on [t̪s̪-θ] in the context of a HB vowel. As discussed earlier, this may be because QF has [t̪s̪] as an allophonic variant of [t̪] in their L1, but not in this environment. Since [θ] is equated with [t̪] in QF, in effect both [t̪s̪] and [θ] are equated with [t̪], thus making these two sounds difficult to discriminate. Since this allophonic variant does not play a role in the EF phonetic system, this may be why we observe relatively good discrimination of these sounds for the EF group.

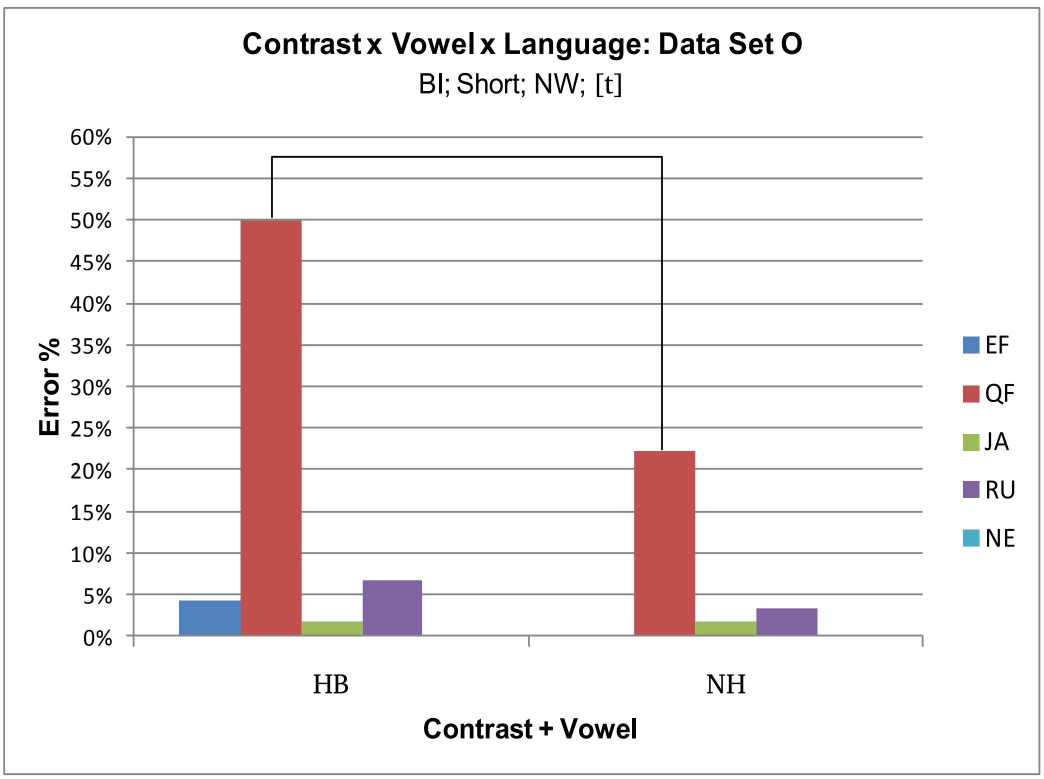

Figure 4.24. Data Set O. Contrast x Vowel x Language.
Connector lines indicate significant differences.

Figure 4.24 shows the interaction for [t-θ] arising within QF with significantly more errors before HB than NH; whereas, the other groups are close to ceiling before both vowels.

Data Set P. (EF/QF/JA/NE; A; Short; NW; f ɸ s s̪ ts̪ t t̪]; HB/NH.) The Anova indicates a significant Contrast x Vowel x Language interaction, (Anova $F(18,204), p < .03$). Two-tailed Mann-Whitney tests reveal the significance to lie within two contrasts ($p < .005$). The first is the [s-θ] contrast, with the EF group displaying significantly more errors than the QF group, but only before a NH vowel. The second is the [s̪-θ], with the JA group having significantly more errors that the NE group before the HB vowel. The relevant portions of this interaction are graphically represented below.

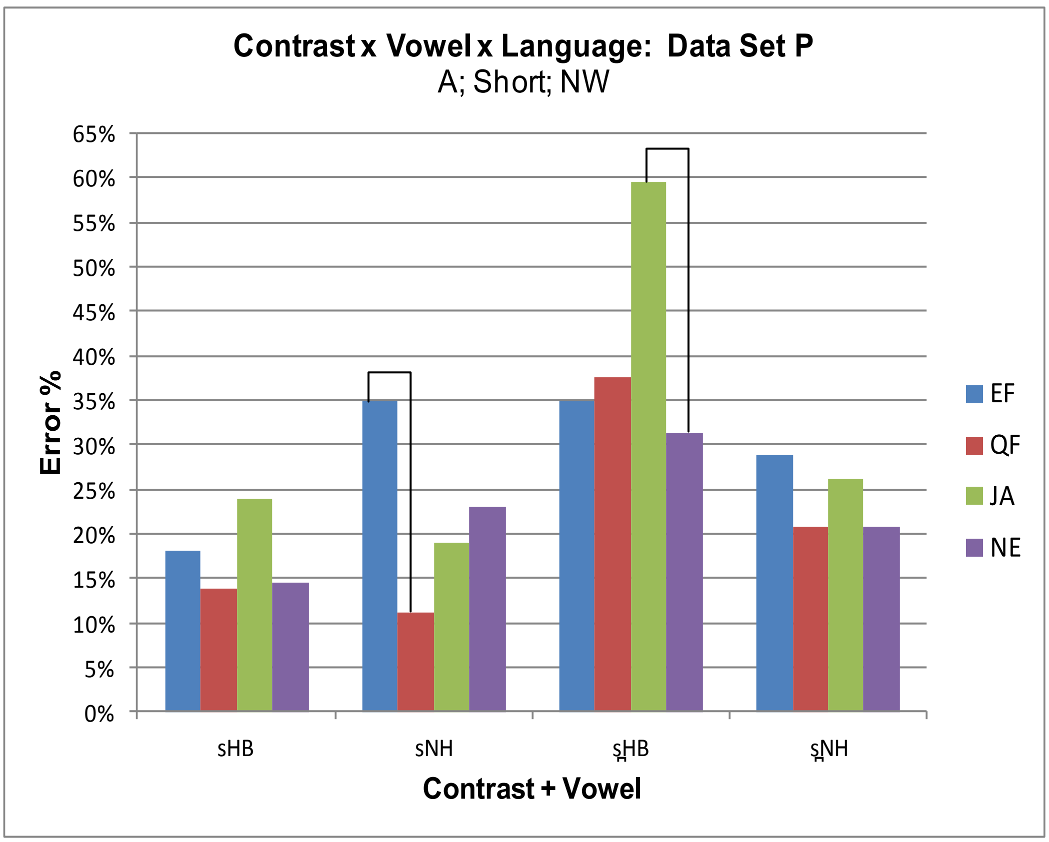

Figure 4.25. Data Set P. Contrast x Vowel x Language.
Connector lines indicate significant differences.

Data Set Q. (EF/QF/JA/RU/NE; BI; Short; NW; f ɸ s s̪ t t̪]; HB/HF/NH.) The Anova here shows a significant Contrast x Vowel x Language interaction, (Anova F(40,390),$p<.001$). Two-tailed Mann-Whitney tests reveal the source of significance in the following areas ($p<.005$): For the contrast [ɸ-θ] in the context of a high back vowel shown in Figure 4.26, the QF group has more errors than JA and NE; and the RU group has a higher error rate than JA. For [s-θ] in the context of a high front vowel only, the JA listeners have more difficulty than RU and NE as seen in Figure 4.27. With regards to the [s̺-θ] contrast in Figure 4.28, the JA group has more errors than QF and NE, but this is statistically significant only before a HF vowel. Finally, for the [t-θ] contrast in Figure 4.29, the QF listeners have more difficulty than all the other language groups; however, this is significant in the context of HB and NH vowels only. The relevant portions of this interaction are shown in the graphs below.

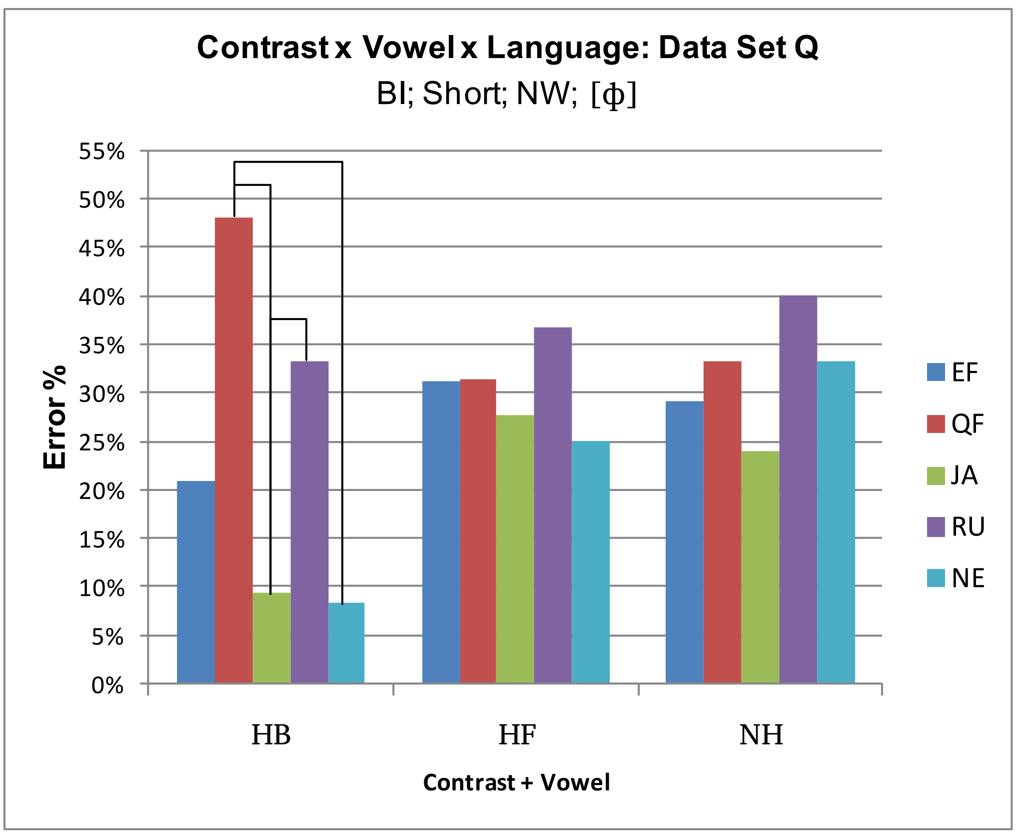

Figure 4.26. Data Set Q. Contrast x Vowel x Language.
Connector lines indicate significant differences.

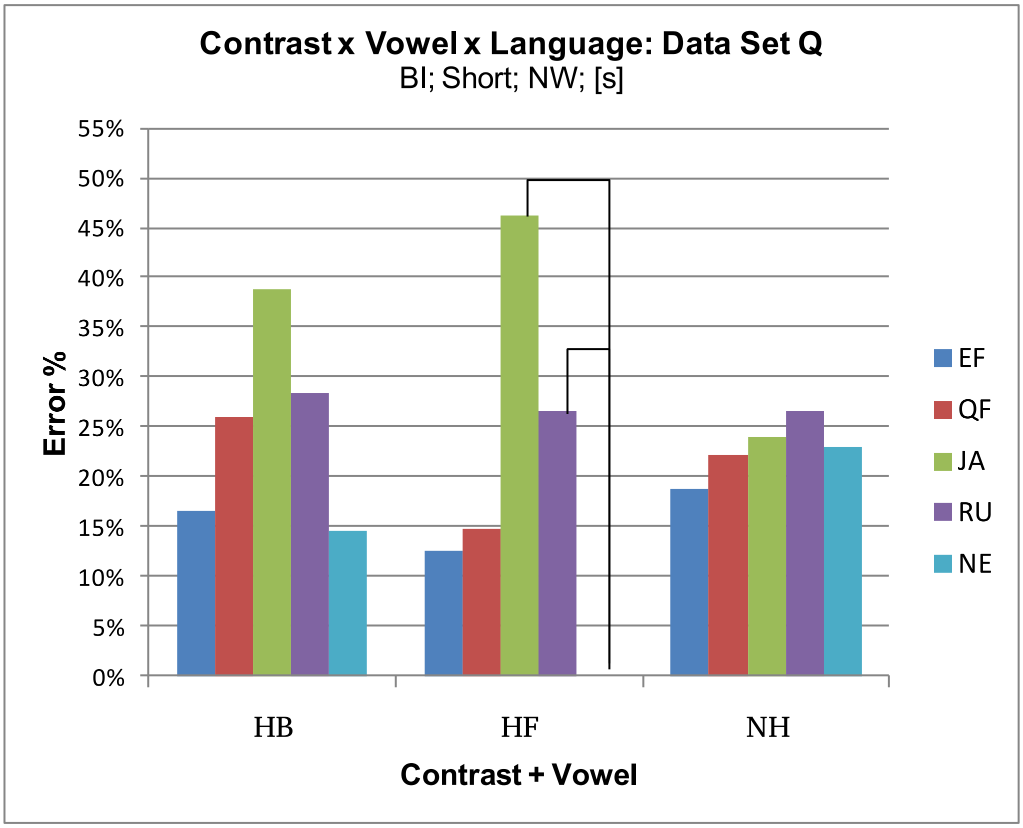

Figure 4.27. Data Set Q. Contrast x Vowel x Language.

Connector lines indicate significant differences.

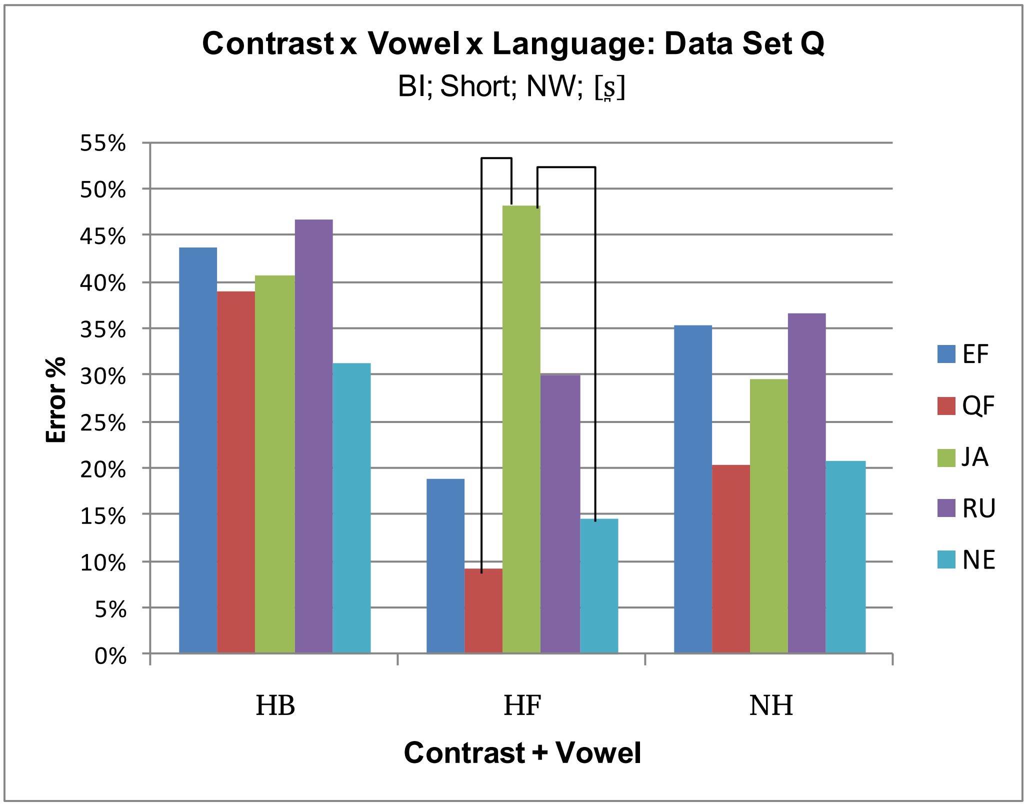

Figure 4.28. Data Set Q. Contrast x Vowel x Language.
Connector lines indicate significant differences.

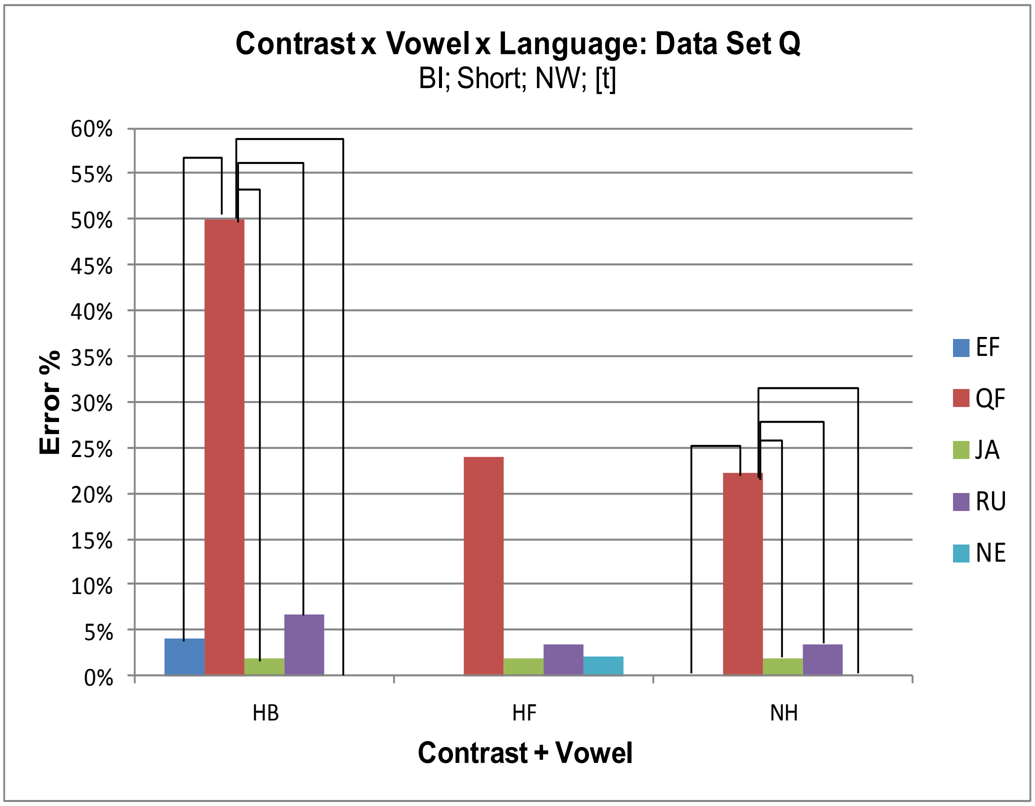

Figure 4.29. Data Set Q. Contrast x Vowel x Language.
Connector lines indicate significant differences.

Data Set R. (EF/QF/JA/NE; A; Short; NW; f ɸ s s̪ t t̪]; HB/HF/NH.) This data set yields a significant Contrast x Vowel interaction, (Anova $F(10,340), p<.001$).

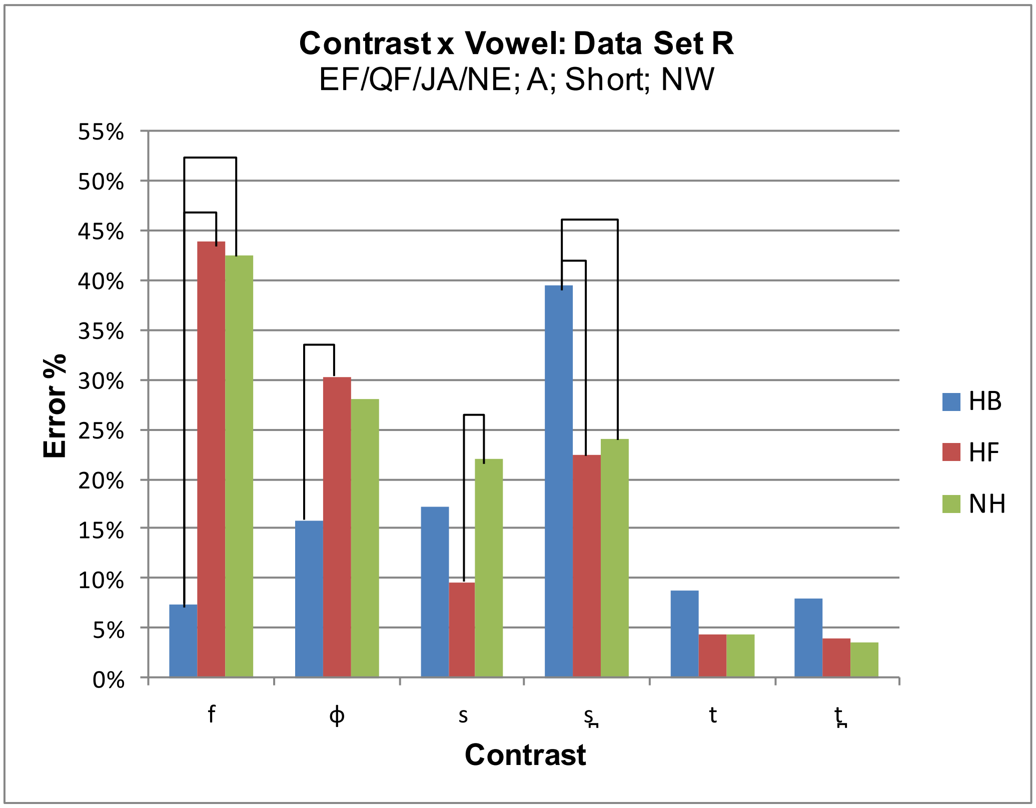

Figure 4.30. Data Set R. Contrast x Vowel.

Connector lines indicate significant differences.

Two-tailed Mann-Whitney tests ($p < .005$) reveal the significance to lie within Contrast [f-θ], with higher error rates for the HF and NH vowels as compared to the HB vowel; within Contrast [ɸ-θ], with higher error rates for the HF vowel as compared to the HB vowel; within Contrast [s-θ], with higher error rates for the NH vowel as compared to the HF vowel; and within Contrast [s̺-θ], with higher error rates for the HB vowel as compared to HF and NH.

This data set analyzed advanced learners in the Short ISI, Non-Word conditions. We see the same general trends in the Contrast x Vowel interactions as seen in previous data sets: higher error rate in the NH condition for the [f-θ] and [ɸ-θ] contrasts and higher error rates in the context of an HB vowel for the [s̺-θ] contrast.

4.7.4.1. Summary of Results on Vowel

Data in this section indicate that vowel context influences the perception of contrasts. We have seen that overall, performance is worse in the context of a non-high vowel as compared to a high-front vowel. This is perhaps related to the findings of Jongman (1989), discussed earlier, whereby interdental fricatives may be more poorly perceived in the context of a low vowel.

One instance of vowel influence on perceptibility is in the case where a non-high vowel impedes the discrimination of [f-θ] and [ɸ-θ] in comparison to a high back vowel context. It is proposed that a lack of cues to [θ] hinders discrimination in the context of low vowels; whereas, high back (rounded) vowels may help increase the distinction by the fact that lip rounding can begin earlier in the articulation of [θ], but not so for [f] and [ɸ].

Another way in which vowel context plays a role in perception again involves a negative impact of a non-high vowel, this time on the ability to perceive the [t̪s̪-θ] contrast, as compared to a high back vowel. However, the facilitating effect of a HB vowel is not explained by earlier lip rounding in either [t̪s̪] or [θ], as lip rounding can begin early in both of these sounds.

The [s̪-θ] contrast shows the opposite pattern, with more errors in the HB condition than in the NH condition. This does not have a ready explanation either.

In conjunction with vowel context, the [t̪s̪-θ] contrast has also revealed inter-language differences. We have seen that QF has more errors than EF on [t̪s̪-θ], and while the QF group's errors are comparable in both the HB and NH conditions for this contrast, the EF group does considerably better in the HB condition. In other words, the HB vowel aids EF in discriminating [t̪s̪-θ], but does not help QF. This may be related to the fact that QF has the affricate as an allophonic variant of [t̪], and that [t̪] is the substitute QF uses in place of [θ]. This may be further exacerbated in a context which is not compatible with the [t̪s̪] allophone.[67] It is unclear why NH vowels inhibit discrimination of this contrast for the EF group, although this is perhaps due to the negative influence of NH vowels on the perception of [θ] as discussed earlier.

On the other hand, for the Japanese group, for whom [ts] is an allophone of /t/ before a High Back vowel, we see a tendency towards better performance in the allophone appropriate high back condition as compared to the inappropriate non-high condition for both proficiency levels.

Vowel context also negatively influences the JA listeners, who have considerable difficulty with the [s-θ] and [s̪-θ] contrasts, especially in the context of HF vowels. This stands in contrast to QF, for whom these STRIDENT/MELLOW contrasts are quite well perceived in the context of a HF vowel. The EF and RU groups are in an intermediate position between JA and QF. The negative influence of HF vowels for JA has been attributed to the fact that there is an allophonic process which palatalizes the coronal fricative before [i] in JA, so Japanese listeners have relatively little experience with [s] or [s̪] before high front vowels.

In addition, I speculate that stridency is enhanced in the context of HF vowels, because of closer contact with the alveolar ridge, creating a narrower constriction. This further enhancement of the feature STRIDENT increases the distance between [s]/[s̪] and [θ]. The effect is greatest for QF, for whom the L1 [s] is already a strong STRIDENT (ALVEOLAR ENHANCES STRIDENT), compared to EF and RU, whose L1 [s̪] is a weak STRIDENT (DENTAL

---

[67] Recall that [t̪s̪] was not tested before HF vowels; therefore, this interpretation is speculative.

MUTES STRIDENT). And, for JA, STRIDENT is hypothesized to be non-existent on their native [s̠], so the HF would not have an effect.

Let us now turn to the remaining results.

4.7.5. Statistical Results for Contrast and Language

This section includes results that only involve the factors of Contrast and Language. Many results involving Contrast and Language have already been discussed if it was found that these factors interacted with other factors. For these cases, consult the section that discusses the specific interaction, either ISI, Proficiency, Wordhood, or Vowel.

Data Set A. (EF/QF/JA/RU/NE; BI; Short/Long; W; [f s s̪ t t̪]; NH). An Anova reveals a significant Contrast x Language interaction (Anova $F(4,38), p < .001$).

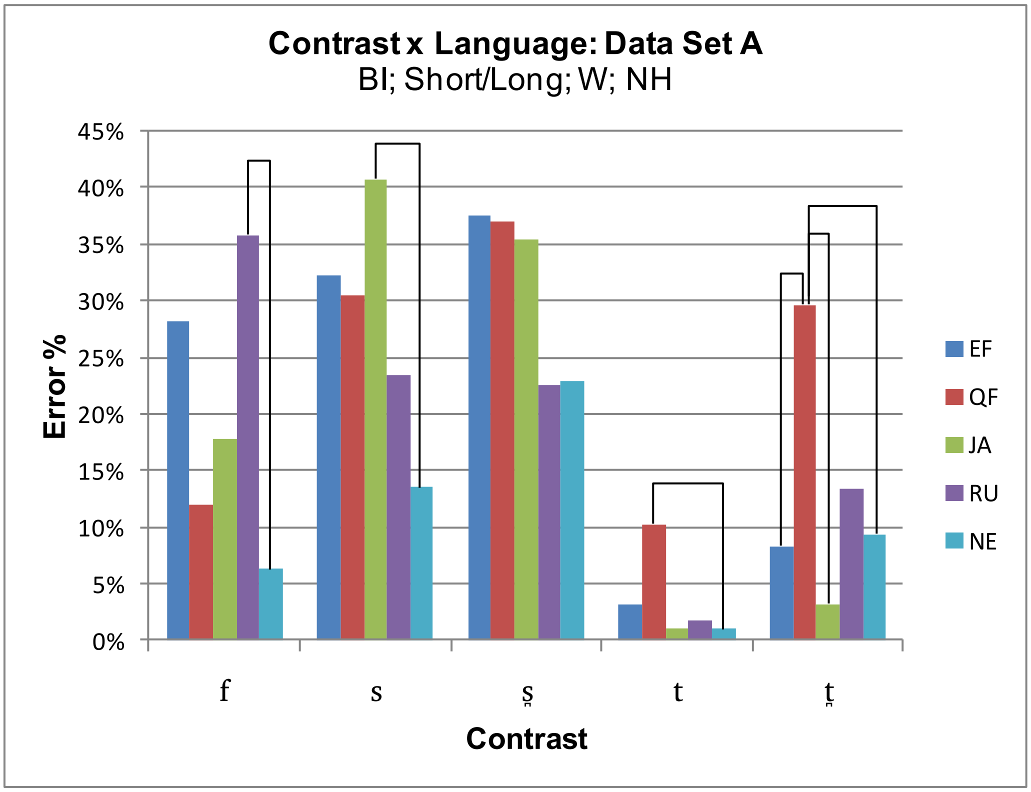

Figure 4.31. Data Set A. Contrast x Language interaction. Significant differences are indicated with a connector line.

Two-tailed Mann-Whitney tests show the difference to lie within the Contrast [f-θ], with RU having more errors than NE; within the Contrast [s-θ], with JA having more errors than NE; within the Contrast [t-θ], with QF having a higher error rate than NE, and within the Contrast [t̪-θ], with QF having more errors than EF, JA, and NE.

All groups had considerable difficulty with [s̪-θ], and, to a lesser extent, with [s-θ], including Native English listeners. For EF-BI, JA-BI, and RU-BI, these results pattern in a manner consistent with their production behaviour: all three of these groups are reported to substitute target theta with their native coronal fricative.

Data Set B. (EF/QF/JA/NE; A; Short/Long; W; [f s s̪ t t̪]; NH). The Contrast x Language interaction is again significant, (Anova $F(12,132), p < .001$); as depicted in Figure 4.32 below, Mann-Whitney tests show the significance to lie between EF and QF on [f-θ]; European French and Japanese on [s̺-θ].

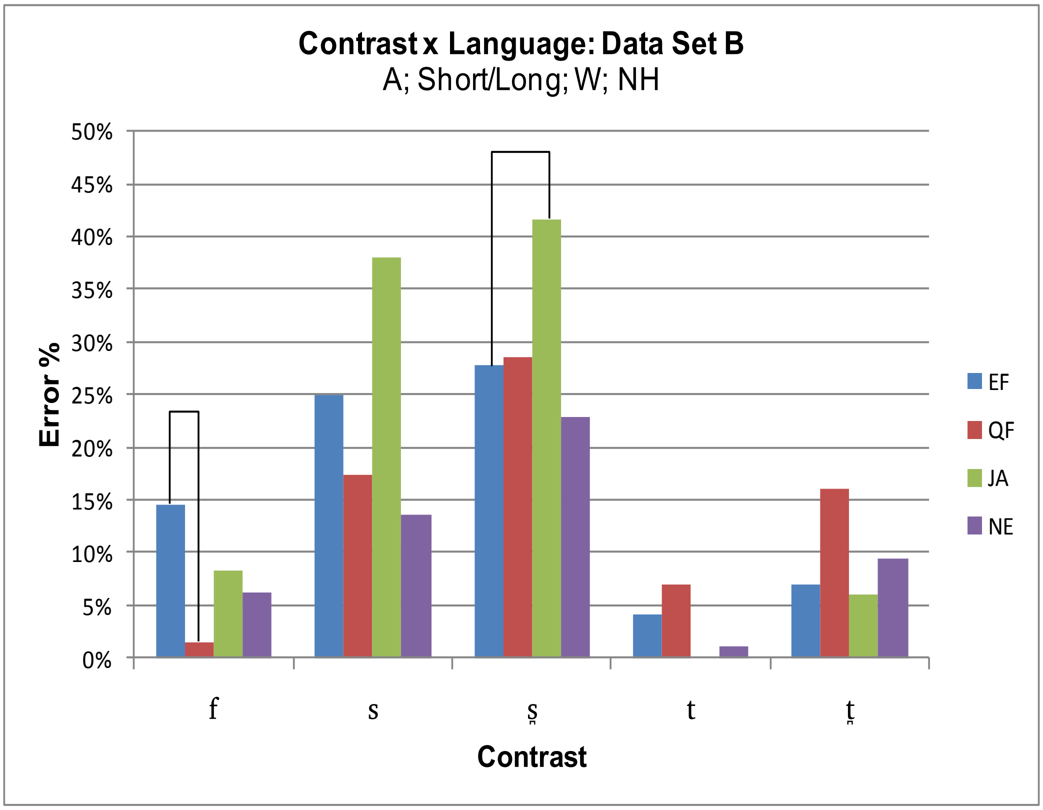

Figure 4.32. Data Set B. Contrast x Language interaction.
Significant differences are indicated with a connector line.

Data Set B shows that the [s̺-θ] contrast is the most difficult to perceive for all language groups. However, we see that JA has significantly more errors on this contrast than does EF. Performance is moderate for the EF group. The difference between EF and QF on [f-θ] indicates that EF has a greater propensity to associate [θ] with [f], while QF does not. This latter finding goes against the QF results from AXB-1, instead showing that for QF, [θ] and [f] are distinct.

Looking within each language group, results show that each of them has significantly more errors on [s̪-θ] as compared to [t-θ], and for EF and JA, as compared to [t̪-θ] as well. Both [θ] and [s̪] are fricatives and both have a dental place of articulation; these two sounds differ only in stridency. But [θ] and [t̪] also share all but one feature: they are both DENTAL sounds and both are MELLOW; they differ only in continuancy. These findings, especially regarding EF and JA, suggest that overall differences in stridency are less perceptible than differences in continuancy. However, the better performance of EF on [s̪-θ] compared to JA, is consistent with the hypothesis that EF [s̪] is specified as STRIDENT, while JA [s̠] is not.

Data Set E. (EF/QF/JA/NE; A; Short/Long; W; [s s̪ t]; HF/NH). Contrast x Language interaction, (Anova $F(8,74), p < .001$). Mann-Whitney tests show the difference to lie within the Contrasts [s-θ], with JA having more errors than QF and NE, and within [s̪-θ], with JA having significantly more errors than the other language groups.

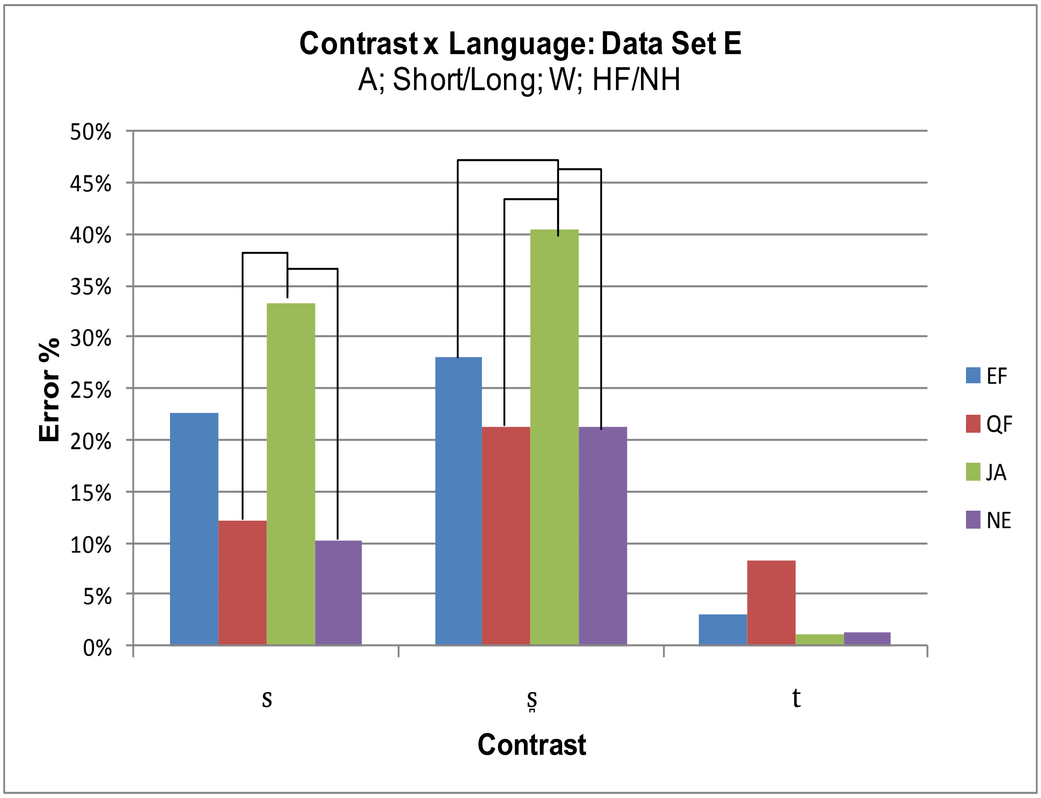

Figure 4.33. Data Set E. Contrast x Language.
Significant differences are indicated with a connector line.

In Data Set E, we again see that the Contrast x Language interaction is due to JA listeners having relatively poor performance on the STRIDENT/MELLOW contrasts compared to the other language groups. The poorer performance of the JA group on [s̪-θ] compared to EF suggests that EF [s̪] is specified for the feature Strident, whereas JA [s̪] is not.

Data Set L. (EF/QF/JA/NE; A; Short; NW/W; f s s̪ t t̪]; NH). Contrast x Language interaction, (Anova F(12,136),p<.001).

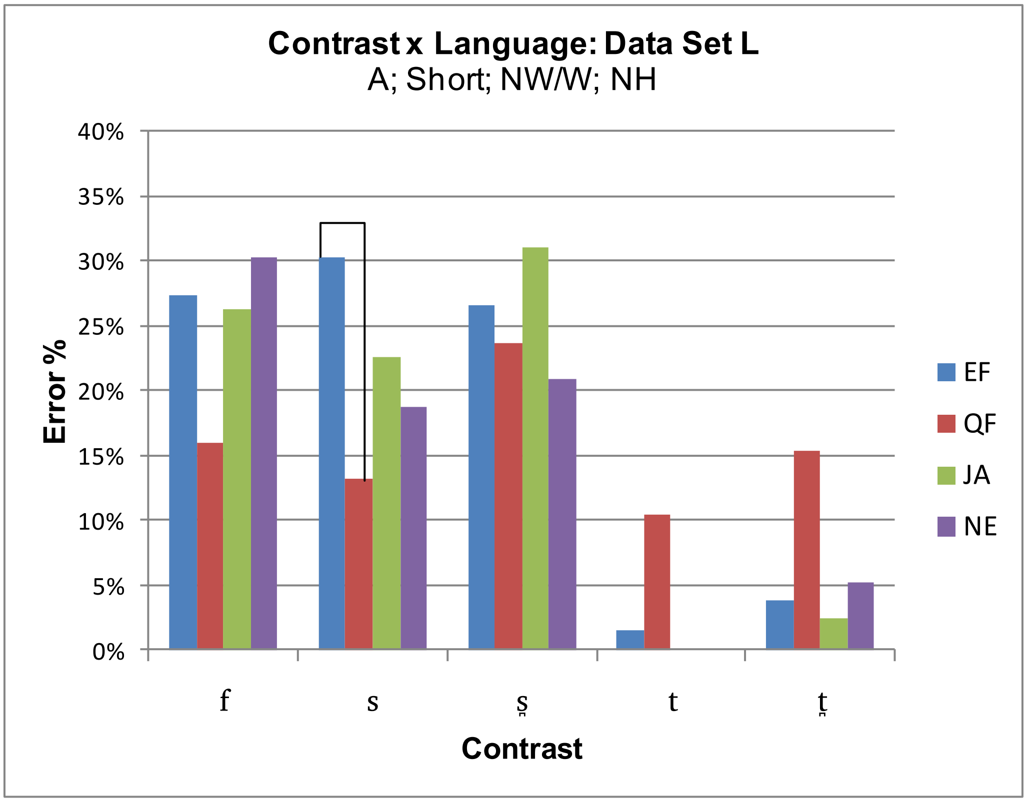

Figure 4.34. Data Set L. Contrast x Language.
Connector lines indicate significant differences.

Mann-Whitney tests show the significance to lie within Contrast [s-θ], with higher error rates for EF compared to QF.

For these groups of advanced learners, we see a striking difference between EF and QF listeners on the contrast [s-θ], with the EF group surpassing the QF group in terms of number of errors. This trend is less evident for the [s̪-θ] contrast, pointing to a universal lack of saliency in the STRIDENT/MELLOW contrast at a dental place of articulation.

Data Set R. (EF/QF/JA/NE; A; Short; NW; f ɸ s s̪ t t̪]; HB/HF/NH.) Contrast x Language interaction, (Anova F(15,170),p<.001).

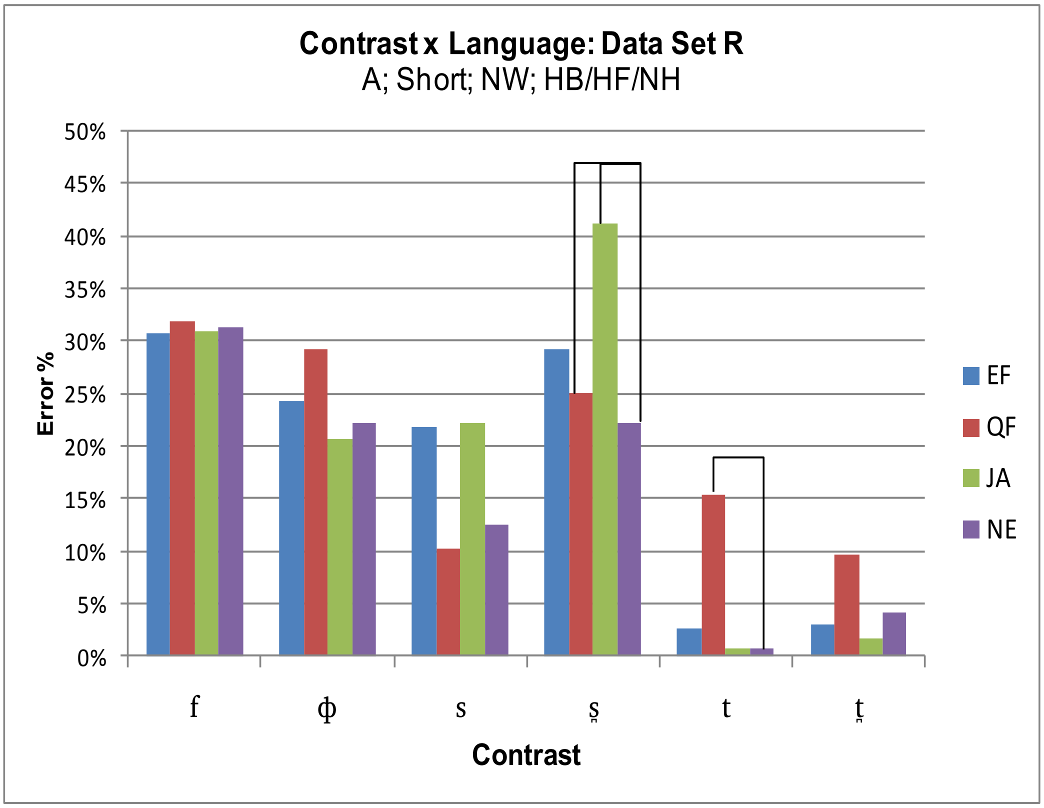

Figure 4.35. Data Set R. Contrast x Language.
Connector lines indicate significant differences.

Two-tailed Mann-Whitney tests show the significance to lie within the Contrast [s̪-θ], with JA having more errors than QF and NE; and within the Contrast [t-θ], with QF having more errors than NE, as seen above in Figure 4.35.

The Contrast x Language interaction shows that across these groups, including the NE group, [f-θ] and [ɸ-θ] are difficult distinctions to perceive. These results also show that [s-θ] and [s̪-θ] (the STRIDENT-MELLOW distinction) are generally more difficult than the [t-θ] and [t̪-θ] (the STOP-CONTINUANT distinction). However, the QF group has more difficulty with the STOP-CONTINUANT distinction than all other groups. In fact, it is striking that other groups are near ceiling on the [t-θ] and [t̪-θ] contrasts, making the QF group stand out with 15 % and 10% errors respectively. On the [s-θ] and [s̪-θ] contrasts, the QF group is comparable to the NE group in comparison to EF and RU, who have more errors for these contrasts. The results for [t-θ]/[t̪-θ] and [s-θ]/[s̪-θ] are compatible with the view that production errors are rooted in perception. The error patterns generally follow what is observed in production with respect to [t] and [s] versus [θ].

4.7.5.1. Summary of Results on Contrast and Language

This section has demonstrated that the ability to perceive certain contrasts depends on the native language of the participant. We have seen that Russian listeners have the most errors on [f-θ], and this is the only language that differs significantly from the Native English control group on this contrast. In addition, the EF group has more errors than the QF group on this contrast. This differs from the findings from AXB-1 (Chapter 3), which indicated that EF and QF had equal difficulty with [f] and [θ].

Japanese stands above the other language groups on the [s-θ] contrast. These findings are consistent with the hypothesis that JA listeners will misperceive target {θ} as a continuant.

This situation holds also for the [s̪-θ] contrast, and is consistent with the idea that the JA native [s̪] is not specified as STRIDENT. The fact that EF has significantly fewer errors than JA on [s̪-θ] is consistent with the hypothesis EF [s̪] is specified as STRIDENT, albeit a weak strident (DENTAL MUTES STRIDENT). We have also seen that EF has more errors than QF on the [s-θ] contrast, which is expected if QF [s] is specified as a strong strident (ALVEOLAR ENHANCES STRIDENT).

The finding for the contrast [t-θ] shows that the QF group has the highest error rate of all the language groups. And the same pattern is seen for [t̪-θ]. These results suggest that QF listeners do have difficulty with the contrast, and we see here a clear difference between the QF and EF groups. These results differ from those in AXB-1 (Chapter 3), which investigated advanced learners only. In general, the results here lend support to the hypothesis that differential substitution is perceptually based, although this must be tempered by the observation of a higher error rate on [s̪-θ] for QF as will be discussed in §4.8.

The fact that we see such cross-linguistic differences generally supports the notion that production errors are grounded in perception.

4.7.6. Statistical Results for Non-Contrastive Features

To test whether the participants were able to perceive non-contrastive features, two pairs of contrasts were investigated: [s-θ] vs. [s̪-θ] and [t-θ] vs. [t̪-θ]. These pairs allow us to examine whether listeners are able to distinguish between apico-alveolar and lamino-dental obstruents.

For each data set, a Repeated Measures Anova was run for each language group on the contrasts [s-θ] vs. [s̪-θ] and [t-θ] vs. [t̪-θ], with other variables collapsed. The Anovas include Bonferroni corrected pairwise comparisons ($p < .05$) to compensate for data that violate homogeneity of variance as well as correcting for multiple comparisons. Results are presented by language group.

For the European French group, it was found that [s̪-θ] has a significantly higher error rate than [s-θ] in four data sets (I,J,O,Q). This was as predicted. Furthermore, it appears this difference stems largely from the difficulty that the BI group has in the context of a HB vowel (see Figures 4.27 and 4.28). In other words, for the EF-BI group the apico-alveolar [s] seems to be a poorer fit to their native lamino-dental [s̪] especially in the context of a high back vowel, rendering moderate discrimination of [s-θ] (*Category Goodness*) versus poor discrimination of [s̪-θ] (*Single Category*).

This same significance also emerged for the Québec French group: [s̪-θ] has a significantly higher error rate than [s-θ] in six data sets (F,H,I,J,N,R), e.g. Figures 4.34-4.35. The influence of vowel here is less clear, with perhaps some inhibitory effect from NH and HB vowels. There doesn't seem to be any particular link to Proficiency. This pattern is not consistent with the predictions, as it was expected that [s-θ] would be more difficult, as alveolar {s} was hypothesized to be closer to the native QF fricative.

The QF group also shows significant differences between [t-θ] vs. [t̪-θ]. It appears that the BI group has more difficulty with [t̪-θ] than with [t-θ] (Data Sets A, C, and G for real Words only), e.g. Figure 4.31. This indicates that apico-alveolar [t] is a poorer fit to their native lamino-dental [t̪]. The predictions were upheld here. Interestingly however, this situation is reversed in other data sets. In Data Sets J and O, [t-θ] presents significantly more difficulty than does [t̪-θ]. These latter data sets involve Non-Words only; which may partially explain this reversed trend.

For the Japanese group, again [s̪-θ] has a significantly higher error rate than [s-θ] in three data sets (H for Advanced only, N, R), e.g. Figure 4.21. It is unclear as to why these differences appear limited to the Advanced JA group.

The results for the Russian group do not reveal any significant differences between these pairs of contrasts, which lends support to the idea that the features APICAL and LAMINAL are underspecified in RU phonetic representations (see footnote 55).

The Native English control group also shows the same pattern as EF, QF, and JA on [s-θ] vs. [s̪-θ], but only in two data sets (D, E), e.g. Figures 4.17 and 4.33. Again the source of this is unclear.

The fact that EF, QF, JA, and especially NE all have a tendency to confuse [s̪] with [θ] compared to [s] versus [θ] may indicate a universal component to this merger. If this is the case, it may that in the AXB task, these sounds are being directly compared to each other rather than being compared against an internal L1 representation, as the Perceptual Assimilation Model would suggest. If this idea is correct, it supports the hypothesis that lamino-dental [s̪] is perceptually closer to [θ] for listeners from all languages. The results on [t-θ] and [t̪-θ] for the QF-BI group also tend to follow this direction, yet the exceptionally good performance for the other language groups on these contrasts indicate that a language-specific factor is involved here.

In general, these findings, and particularly with respect to EF and QF, suggest that listeners are able to perceive non-contrastive features in the AXB task, and that this ability may be related to vowel context, proficiency level and wordhood.

4.7.7. Analysis of Russian Participants

Recall that the reason for including Russian participants in this study was to get an idea as to whether there exists inter-dialectal variation amongst this language group. The impetus behind their inclusion came from conflicting reports in the literature on the production of English by Russian speakers. The majority of these reports state that Russian speakers substitute the English interdental fricative with a coronal stop (e.g. Weinberger 1988, Lombardi 2003); however, one study reports the coronal fricative (Teasdale 1997), and anecdotal reports have also mentioned the coronal fricative (see §3.5).

The Russian participants in this study were from the following regions:

(20) Regional Affiliation of Russian Participants

RU-1: Ukraine	RU-6: Ukraine
RU-2: Kirov	RU-7: Moldova
RU-3: Belarus	RU-8: Kirov
RU-4: Kirov	RU-9: Kirov
RU-5: Belarus	RU-10: Sochi

These regions are all located in western Russia or the former Soviet Union. Obviously, with such a small sample from a variety of regions, it is difficult to make any generalizations regarding dialectal differences. Nonetheless, the data from these participants does show a couple of interesting tendencies.

A comparison of perceptual errors on [f s t] was carried out on individuals based on Data Set A. Both dental and alveolar /s t/ were grouped together. These are presented in Figures 4.36-4.39.

First, it is notable that for most of the RU listeners, with the exception perhaps of RU-9, [t-θ] and [t̪-θ] do not appear to be overly problematic. Their data resembles more those of the EF listeners in Figure 4.37 rather than the QF listeners in Figure 4.38, consistent with the predictions for RU at the outset of this chapter.

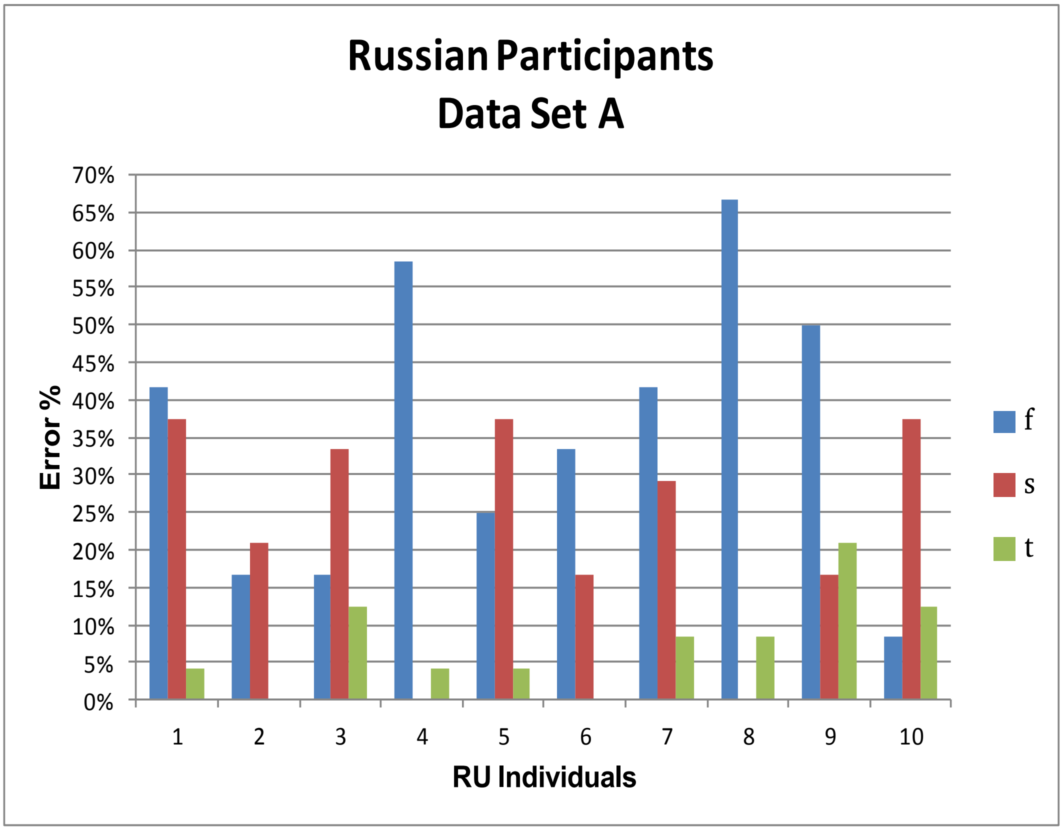

Figure 4.36. Russian Participants. Data Set A (ISI collapsed).
Key: f = [f-θ]; s = [s-θ] and [s̪-θ]; t = [t-θ] and [t̪-θ].

A second noteworthy result in the RU data in Figure 4.36 is that two of the participants (RU-4 and RU-8) have no errors at all for the [s-θ] and [s̪-θ] contrasts; instead, these two listeners have a very high error rate for [f-θ]. This is unusual as they are the only two participants to show this pattern in the test languages depicted in Figures 4.36-4.39. These two participants are from the same region of Russia (Kirov). Participants RU-2 and RU-9 are also from Kirov, and RU-9 also has a very high error rate for [f-θ]. Although tempting to surmise that there might be some dialectal influence for this pattern, it must be pointed out that all four of these participants are also from the same family, which might also explain this finding.

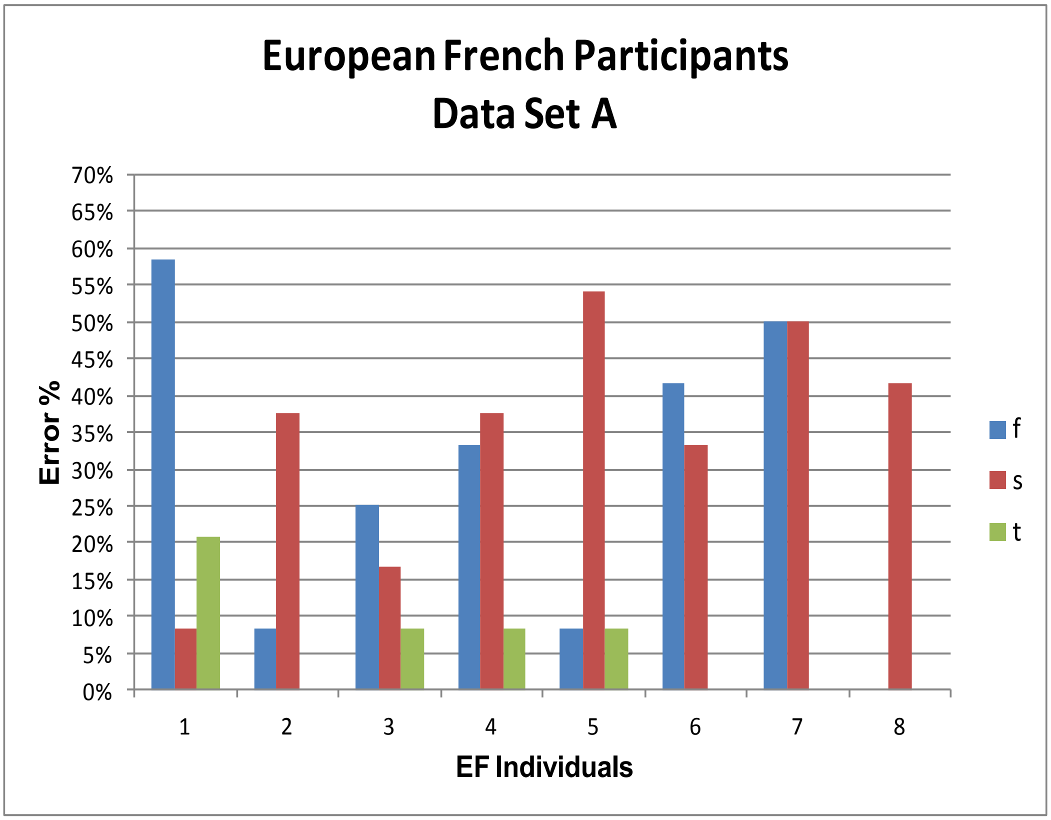

Figure 4.37. European French Participants. Data Set A (ISI collapsed).
Key: f = [f-θ]; s = [s-θ] and [s̺-θ]; t = [t-θ] and [t̺-θ].

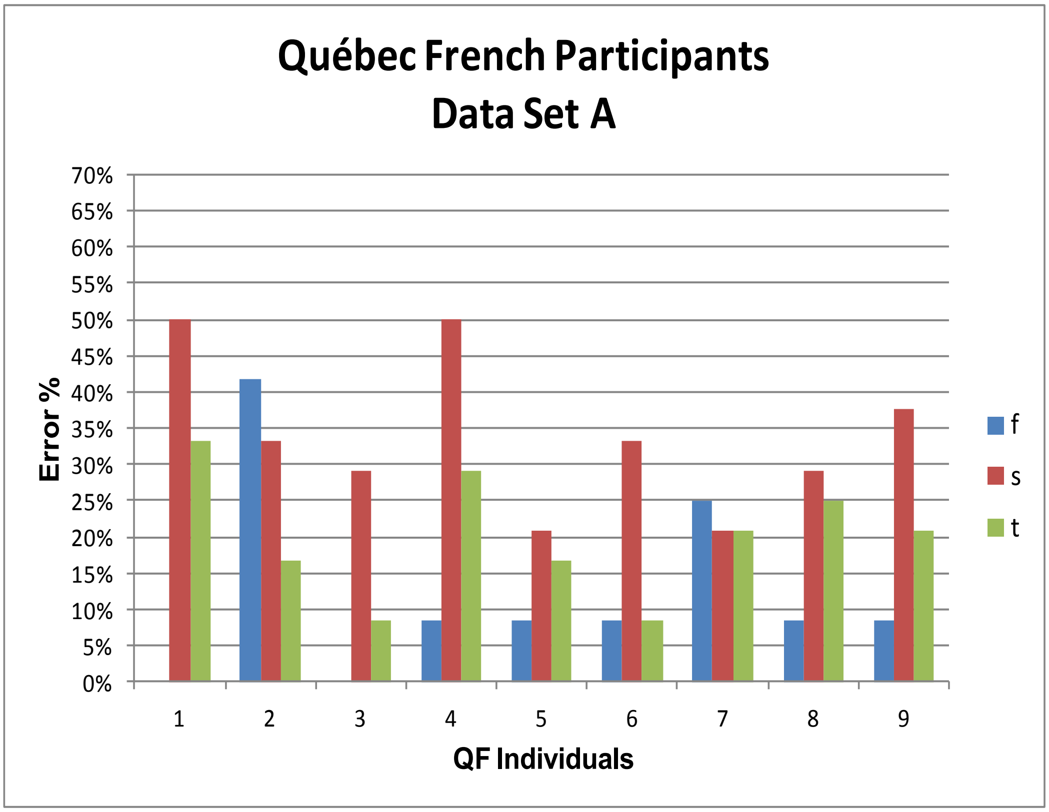

Figure 4.38. Québec French Participants. Data Set A (ISI collapsed).

Key: f = [f-θ]; s = [s-θ] and [s̪-θ]; t = [t-θ] and [t̪-θ].

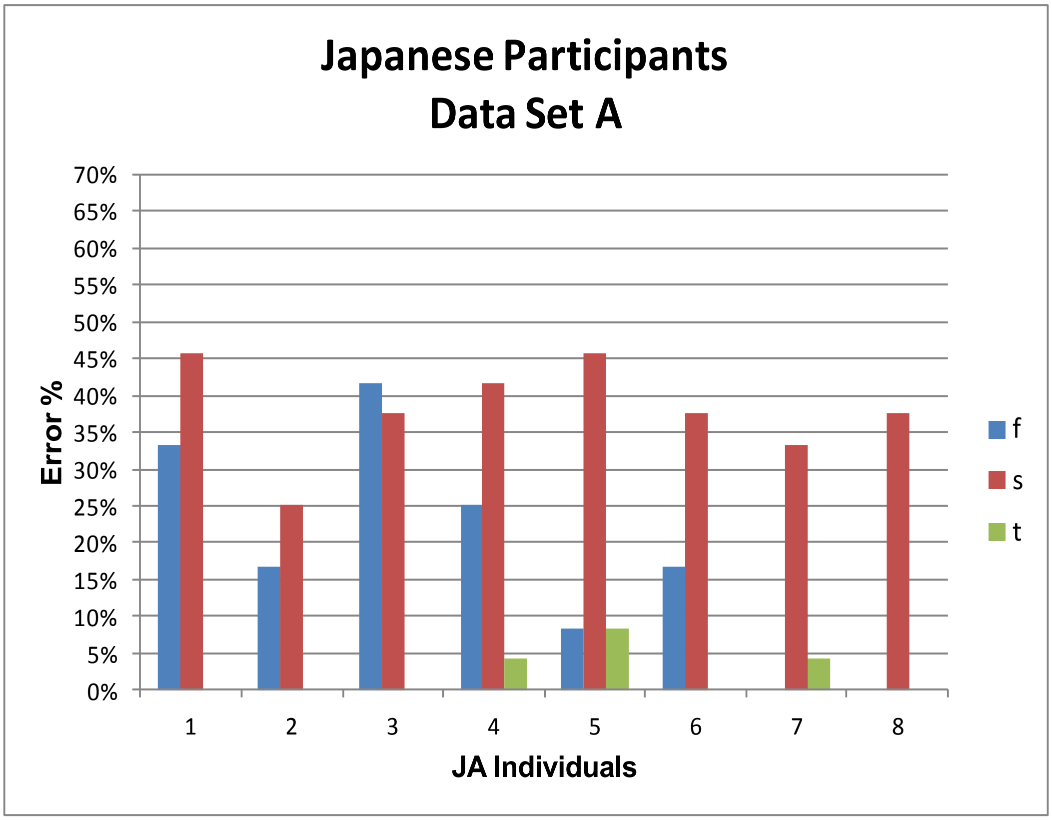

Figure 4.39. Japanese Participants. Data Set A (ISI collapsed).
Key: f = [f-θ]; s = [s-θ] and [s̪-θ]; t = [t-θ] and [t̪-θ].

Perhaps the most pertinent aspect of the data in Figure 4.36 is that a coronal stop is not the principal choice of perceptual substitute for these Russian learners of English. This conflicts with the majority of reports in the literature. This discrepancy may be due to several factors. First, it could be that perception errors do not correspond to production errors.[68] It could also be the case that these reports of stop substitutes are based on speakers from eastern regions of Russia. Then again, it may be that the reports of [t] as the principal substitute are erroneous. Note that most of the reports are based on second-hand information, referring mainly to Weinreich (1953).

[68] A production study was carried out on some of these Russian individuals, and will be reported in Chapter 6. Results from this study indicate that the coronal fricative is the preferred production substitute for these participants.

4.7.8. Results from the Perspective of the Perceptual Assimilation Model

At the beginning of this chapter, predictions were made as to how well each language group would perceive the various contrasts tested in the current study. Phonetic descriptions of each test language's obstruent inventory were given, and feature specifications were proposed.

An outline was given of the Auditory Distance Model incorporating feature weight and auditory distance based on saliency, feature enhancement and muting relations. It was proposed that learners compare the phonetic features and weight of an L2 target sound with those of L1 native phonetic categories. In cases where the L2 target sound does not have an equivalent in the L1 inventory, the closest L1 sound will be selected as a substitute. This model was coupled with Best's Perceptual Assimilation Model (PAM) in order to account for how listeners assess the proximity of pairs of sounds, as in the AXB task. The PAM gives various scenarios for how two sounds may be categorized: as a *Single Category Assimilation* (SC), which would result in poor discrimination; as a *Category Goodness Assimilation* (CG), which would give moderate results, or as a *Two Category Assimilation* (TC), which would yield good discrimination of the pair of sounds.

To determine the type of assimilation for each contrast and language group, an Anova with Bonferroni correction or Wilcoxon tests were conducted on each language group within each data set.[69] In each case, contrast pairs were categorized into either SC, CG, or TC based on significant differences between them. In several instances, pairs could not

[69] If there was no interaction involving two within-subjects factors, then the Anova with Bonferroni correction was used ($p < .05$); if there was an interaction, then the Wilcoxon tests were used ($p < .005$).

be classified, as no significance was found. These were flagged as indeterminate. These data were compiled for all data sets. Then for each language group, I counted how many times a contrast fell into either of the three categories. The category into which a contrast was most often situated was chosen as representing the type of assimilation, either SC, CG, or TC.

The following tables give the predictions as outlined at the beginning of this chapter and the actual results according to the method described above. A separate table is presented for each contrast. Where the actual results agree with either the phonetic or phonemic predictions, the cell is shaded. Factors other than Language and Contrast are conflated in the actual results. The fusion of these factors means that these classifications must be viewed as being rather general. We have seen earlier that vowel context, wordhood, and proficiency have effects on the perception of contrasts, so these categorizations may differ when different factors are analyzed separately.

(21) Predictions vs. Actual Results for [ɸ] vs. [θ]

Language	Phonetic Prediction	Actual
EF	TC	SC or CG
QF	TC	SC or CG
JA	TC	CG
RU	TC	SC or CG
NE	TC	SC or CG

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation.

The results in (21) do not conform to what was predicted for [ɸ-θ].

(22) Predictions vs. Actual Results for [f] vs. [θ]

Language	Phonemic Prediction	Phonetic Prediction	Actual
EF	TC	TC	SC
QF	TC	TC	CG
JA	**CG** or TC	TC	**CG**
RU	**CG** or TC	TC	SC or **CG**
NE	TC	TC	CG

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

The results in (22) were not anticipated for EF, QF, and NE. As for JA and RU, either both [f] and [θ] are heard as poor exemplars of [s], according to the PAM, or it could be that [θ] is perceived as a poor exemplar of [f].

(23) Predictions vs. Actual Results for [t] vs. [θ]

Language	Phonemic Prediction	Phonetic Prediction	Actual
EF	TC	TC	TC
QF	TC	CG	CG
JA	TC	TC	TC
RU	CG or TC	CG or TC	CG or TC
NE	TC	TC	TC

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

For [t-θ] in (23), both the phonemic and phonetic predictions are the same, except for QF, and in this case, the phonetic predictions are supported. While it was anticipated that QF would show a TC assimilation based on CONTINUANT versus STOP, poorer performance was expected in the phonetic condition due to the interplay of phonetic features. Target [θ] is more likely to be associated with QF [t̪] than with [s] because of the availability of the non-contrastive feature STRIDENT on the latter.

(24) Predictions vs. Actual Results for [t̪] vs. [θ]

Language	Phonemic Prediction	Phonetic Prediction	Actual
EF	TC	TC	TC
QF	TC	CG	CG
JA	TC	TC	TC
RU	CG or TC	CG or TC	CG or TC
NE	TC	TC	CG or **TC**

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

In (24) we see that, as with [t-θ], it is the phonetic predictions, not the phonemic, that are supported for QF.

(25) Predictions vs. Actual Results for [s] vs. [θ]

Language	Phonemic Prediction	Phonetic Prediction	Actual
EF	SC	CG	CG
QF	SC	TC	CG
JA	SC	CG	CG
RU	SC or TC	**CG** or TC	SC or **CG**
NE	TC	TC	CG

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

For [s-θ], as seen in (25), the phonetic predictions are borne out for EF and JA. However, the poor performance for NE was not expected.

(26) Predictions vs. Actual Results for [s̪] vs. [θ]

Language	Phonemic Prediction	Phonetic Prediction	Actual
EF	SC	CG	SC
QF	SC	TC	SC
JA	SC	CG or TC	SC
RU	SC or TC	CG or TC	SC or CG
NE	TC	TC	SC

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

In (26), we see that the phonemic condition is the better predictor for [s̪-θ] in the test languages, and again we observe unanticipated poor performance for NE.

(27) Predictions vs. Actual Results for [t̪s̪] vs. [θ]

Language	Phonetic Prediction	Actual
EF	SC	CG
QF	TC	SC or CG
JA	TC	SC or CG
RU	**CG**	SC or **CG**
NE	**CG**	SC or **CG**

TC = Two Category Assimilation; CG = Category-Goodness Assimilation; SC = Single Category Assimilation

For [t̪s̪-θ], in (27), predictions are upheld for RU and NE. For RU, this is consistent with the idea that both sounds are funneled into [s̺], but are poor exemplars of this category. And for NE, this result suggests that [t̺s̺] is interpreted as a poor exemplar of [θ].

The discrepancies between the actual results and the predictions suggest that some enhancement relations require revision; particularly in response to the relatively poor discrimination of [ɸ-θ] and [f-θ] by all language groups. This will be addressed in the discussion below.

4.8. Discussion of Results for AXB-2

This chapter has examined English L2 learners' perception of seven contrasts involving the interdental fricative presented in an AXB task. The languages investigated were European French, Québec French, Japanese, Russian, and a Native English control group. For EF, QF, and JA, two proficiency levels were examined: Beginner/Intermediate and Advanced. The AXB task involved two interstimulus intervals: a Long ISI, intended to induce phonemic processing, and a Short ISI, intended to invoke phonetic processing. As well, the effect of wordhood was examined through real English words versus non-words and the effect of vowel context was investigated by presenting three types of vowels: high front, high back, and non-high.

The AXB-2 experiment was conducted as a follow-up to AXB-1, presented in Chapter 3. Recall that AXB-1 failed to find any difference between the Long and Short ISI conditions. Thus the AXB-2 study incorporated carrier phrases and different talkers within each stimulus item in addition to limiting the Long ISI condition to real words, all in an attempt to evoke phonemic processing.

Results for ISI show that errors increase overall with a longer interstimulus interval. The "across-the-board" nature of this finding indicates that a Long ISI puts a strain on general memory systems, rather than tapping language-specific phonological representations. Otherwise, we would expect different results in each of the ISI conditions depending on Language, because the set of phonologically contrastive features differs across languages. Nonetheless, some inter-language differences were found, although not tied to ISI. As with the AXB-1 study, it appears that the AXB-2 experiment has also tapped a phonetic level of representation.

In AXB-1, differences between the two dialects of French did not clearly emerge: both EF and QF tended to misperceive the interdental fricative as [f], and for EF, equally as [s̪]. It was thought that introducing a lower level of proficiency in the AXB-2 study would bring out such differences between these dialects. As well, a Beginner/Intermediate group of Russian participants was added to examine whether perception matches production reports for this group, and to explore the possibility of inter-dialectal differences in Russian.

Results from this study have shown that there is a difference in proficiency, with an overall improvement in the Advanced learner group. This indicates that the placement test used to classify participants was valid. As well, differences between EF and QF emerged by introducing the lower proficiency groups. The QF-BI group had more errors than the other language groups on the stop-interdental contrast, as was hypothesized.

Interactions involving Proficiency indicate that type of Contrast and Vowel and Wordhood have an influence at different Proficiency levels. One of these results is that the Advanced learners have fewer errors than the BI learners on [f-θ], but the effect is greatest with real words. This may indicate that the improvement that comes with higher proficiency levels stems from increased familiarity with real words, at least for certain contrasts.

Another interesting effect of wordhood concerns the Québec French group, and especially the BI learners, on the stop-interdental contrasts. While in general, we see that the QF participants have more errors than the other language groups on [t-θ] and [t̻-θ], there is an interaction with Wordhood for these contrasts. On the one hand, we observe that QF has more errors for [t̻-θ] in real Words. On the other hand, this same group has more errors for [t-θ] in Non-Words. This may suggest that these participants are more likely to have native-like representations for both [t] and [θ] in real English words. Thus, when presented with [t̻-θ] in real words, the interdental [θ] is correctly represented, but the non-English lamino-dental [t̻] is more likely to be confused with [θ]. If this is correct, it introduces the possibility that non-contrastive information may be available in lexical representations. This will be investigated further in the Chapter 5.

The effect of Wordhood appears largely limited to [f-θ], [t̻-θ], and [t-θ]. It is unclear why Wordhood status would affect only these contrasts. Further research is needed to determine whether these results can be replicated.

Results also reveal differences between languages due to Vowel context. Native language allophonic contexts may influence perceptibility. In general, the data in this chapter have indicated that the Japanese group has more difficulty with the [s-θ] and [s̻-θ] contrasts when they precede a high front vowel. This result is compatible with findings by Muñoz-Sánchez (2003), who found that JA listeners had difficulty with [s-θ] before [i]. She interpreted this as being due to Japanese listeners' inexperience with the sequence [si] (because of Japanese L1 process whereby /s/ → [ɕ]/_[i]), particularly when paired with the non-native sequence [θi].

The existence of allophonic processes in the L1 may have more influence for the Beginner/Intermediate learners than for the Advanced learners, especially for the [ɸ-θ] contrast. For this pair, the QF-BI group has the highest error rate before an HB vowel; whereas the JA-BI group has the lowest error rate in this vowel context. This difference disappears in the NH vowel context and with the advanced groups. Since a high back vowel is the appropriate context for the native JA allophone [ɸ], this may partly explain the good performance for the JA-BI group.

Other interactions with Vowel indicates that, overall, a non-high vowel impedes perception, and this may be due to the interdental. As suggested in Chapter 3, this agrees with previous research (Jongman 1989). Thus, it may be that the fact the one member of each contrast is an interdental fricative results in poorer discrimination before NH vowels. On the other hand, results showed that a HF vowel facilitates perception of these contrasts. I have speculated that better performance before high front vowels as compared to NH vowels may be due to the feature STRIDENT being enhanced on fricatives before a HF vowel.

A high back vowel type also affects discrimination. The [f-θ] and [ɸ-θ] contrasts were easier to distinguish before a HB vowel. It was suggested that the lip rounding which is present on high back vowels aids in discriminating these sounds because for [f], rounding can only begin after the consonant has been released; whereas, for [θ] rounding begins in conjunction with the consonant articulation. We have also seen that [t̪s̪-θ] is easier before a HB vowel; however, this cannot be explained in terms of lip rounding, because rounding can begin during the consonant articulation for both of these sounds.

The data from this experiment demonstrate that it does not suffice to say that certain languages have difficulty with certain contrasts; we must consider the vowel context involved as well. This is consistent with other research that has found fricative-vowel interactions (Soli 1981, Schmidt 1996, Muñoz-Sánchez 2003). It has been suggested elsewhere that fricatives, and especially non-strident fricatives, cannot be defined apart from the vowel that they occur with (Jongman 1989). This seems to receive support from the data presented in this chapter. The majority of the Contrast x Vowel interactions occur with contrasts where [θ] is paired against either a fricative or an affricate. In general, these findings are consistent with the view that processing proceeds in syllable-sized chunks rather than segmental-sized (Peperkamp & Dupoux 2003).

Throughout this chapter, inter-language differences in the perception of certain contrasts have emerged. In some data sets, the European French group registered more errors than QF on the [s-θ] contrast; thus it seems that QF is able to make the distinction between STRIDENT and MELLOW much more effectively.

In addition, we have repeatedly seen that QF does relatively poorly on the [t-θ] and [t̪-θ] contrasts as compared to the other language groups. This finding is consistent with the hypothesis that QF listeners would perceptually confuse [t-θ] and [t̪-θ]. This is compatible with the pattern of errors seen in production data for QF and thus supports the hypothesis that production is determined by perception. Recall that the results of the AXB-1 study indicated that Advanced QF participants were more likely to perceptually substitute [f] for [θ]; thus, the inclusion of Beginner/Intermediate QF listeners in AXB-2 reveals the perceptual origin of this group's stop substitutes. Furthermore, the striking difference between the poor performance of Québec French on the stop-interdental and affricate-interdental contrasts as compared to European French lends support to the idea of a perceptual basis to the differential substitution patterns exhibited by Québec French versus European French.

Another result that is seen throughout this chapter is that JA has considerable difficulty with [s-θ] and [s̪-θ], as compared to the other language groups, and especially as compared to QF.

The results for Russian show that this group generally does relatively poorly on the [s-θ], [s̪-θ], and [f-θ] contrasts and well on the [t-θ] and [t̪-θ] contrasts.

These inter-language differences are again indicative that differential substitution has a perceptual basis. Production reports state that JA and EF tend to substitute [s] for English [θ]. The perceptual data for JA and EF coincide with the production reports. For Russian, as discussed earlier, there is some discrepancy in the production reports with many sources saying that RU speakers produce [t] for [θ], while some others consider the substitute to be [s]. Results from the present experiment are more compatible with the latter. Data from this study reveal that Beginner/Intermediate RU listeners perceptually confuse either [s] or [f] with English [θ]. Thus, for RU, either perception is not at the root of production errors, or the reports that RU speakers substitute [t] are incorrect, or the production reports may be based on Russian dialects other than those tested in the current study.

Individual Russian participants were investigated to determine whether some evidence of inter-dialectal differences exists for this language group, along the lines of the Québec versus European French differences. Inter-individual differences did emerge. Some individuals have the most errors on [f-θ], while for others, the highest error rates are on [s-θ] and [s̪-θ]. However, for all RU participants, errors on [t-θ] and [t̪-θ] were the lowest. Thus there may be some indication of a dichotomy between RU listeners who confuse [f] with [θ], and others who confuse [s] and [s̪] with [θ], and there may be a link between perceptual substitute and regional affiliation. However, there are too few participants in this study to make a generalization about inter-dialectal differences. More research is required to verify these findings.

Another hypothesis behind this study was that listeners make use of non-contrastive features in their assessment of L2 target segments. Statistical analyses here do show that most of the participants in the AXB task are sensitive to differences that involve features that are not contrastive in their native languages. This ability is dependent however on wordhood, vowel context, and proficiency level. The Russian group was alone in failing to show any statistical difference on pairs which manipulated non-contrastive features, i.e. pairs in which [θ] was contrasted with either apico-alveolar or lamino-dental obstruents, viz. [t-θ] vs. [t̪-θ] and [s-θ] vs. [s̪-θ]. Thus EF, QF, and JA, are potentially able to make use of non-distinctive features in their assessment of target segments, at least in the AXB task. This is consistent with the hypothesis that phonetic differences between languages motivate differential substitution.

Despite the inter-language differences observed in this chapter, the overall general trend is towards fricative-interdental contrasts being more difficult than stop-interdental contrasts. Specifically, it was found that for all language groups, the [ɸ-θ] and [f-θ] contrasts were relatively difficult to perceive. Similar results were found in AXB-1. It is well-known that the [f-θ] contrast is a difficult one (Miller & Nicely 1955, Cutler et al. 2004). Cues to Major Articulator (place) features are poorly discerned on mellow fricatives. In Chapter 2 of this thesis, it was argued that Major Articulator features are inherently salient; however, these results suggest that this depends on the type of sound.

It seems that the salience of Major Articulator features is tied to the Turbulence (manner) features. It is therefore suggested that this relationship be formally expressed. Rather than Major Articulator features being inherently salient, I propose that place features on fricatives are muted when these fricatives are non-strident. Thus, I introduce the muting effect MELLOW FRICATIVE MUTES MAJOR PLACE.[70] The consequence is that LABIAL and CORONAL would only have a weight of 1 on [ɸ,β], [f,v] and [θ,ð], making the difference between the labial fricatives and the interdental fricatives perceptually closer. This lack of saliency accounts for the cross-linguistically low frequency of inventories with both labial fricatives and non-strident coronal fricatives.

In the present study, the difficulty with fricative-fricative contrasts extends to the [s-θ] and [s̺-θ] contrasts, particularly so for the latter. This applies to all language groups, including Québec French and, somewhat surprisingly, Native English. Although consonant confusion studies show a propensity for [s] to be heard as [θ] by Native English speakers, we do not observe the bi-directionality of confusion as is seen in [f-θ], i.e. [θ] is more likely to be confused with [f] (Miller & Nicely 1955). In fact, in the data from Cutler et al. (2004), Native English listeners never associated [θ] with [s].

These observations, in conjunction with the lack of significant differences between ISI conditions, leads me to speculate that this AXB task has tapped a mix of phonetic and acoustic processing, similar to findings by Gerrits (2001), who noted a tendency to change from a phonetic mode to an acoustic mode in an AXB task. In the present study, it may be that the greater confusion between the sibilant fricatives and [θ] is a combination of acoustic processing (despite the effort made in the test design to avoid this situation), and language-specific effects. This may be caused by the difficulty inherent in this AXB task, i.e. different talkers and carrier sentences, which has led listeners to sometimes switch into an acoustic mode. In other words, instead of associating each target sound with a

[70] This elevates muting effects to the same formal status as enhancement effects.

native representation, the grammar is bypassed, and the two targets are directly compared based on acoustic similarity. This would also account for the generally higher error rates for [s̺-θ] versus [s-θ]. Lending support to this acoustic explanation is the fact that the inflated error rates for [s-θ] and [s̺-θ] mostly stem from the Short ISI condition: 250msec could favour an assessment based on acoustic traces.

In conclusion, the study presented in this chapter generally supports a perceptual foundation and points to the implication of non-contrastive, phonetic features in differential substitution.

In the next chapter I will present results from a Picture Identification Task. This type of task is generally considered to tap a lexical level of representation, and thus will provide a clearer picture of inter-language perceptual patterns in phonological processing.

CHAPTER 5: PICTURE IDENTIFICATION TASK

5.1. Introduction

In the two previous chapters, results were presented from two AXB discrimination experiments that examined the perception of various contrasts involving the interdental fricative by learners of English as a second language from various native language backgrounds. One of the aims of those tasks was to investigate how learners perceive these contrasts at a phonetic level of representation as opposed to a phonological level of representation. However, neither of these AXB tasks showed a significant difference between the two conditions which manipulated different interstimulus interval lengths, and as such, no clear statement could be made regarding phonological processing.

In this chapter, I present the results from a Picture Identification (PicID) task. This task requires lexical processing; in other words, the participant must access his or her internal phonological representation which corresponds to a picture (Brown 1997,1998). Comparing the results from the Picture Identification task (PicID) with those from the AXB tasks will demonstrate whether different patterns of perception are associated with the different tasks, thereby allowing a comparison between phonological processing, and phonetic or acoustic processing.

5.2. Predictions

The reader will recall the hypotheses behind the research presented in this dissertation:

1. Differential substitution is due to transfer from the native language (L1);

2. Transfer in production is based on transfer in perception;
3. Choice of substitute depends on a comparison of the phonetic properties of the target segment with phonetic properties of segments in the L1 sound system;
4. Transfer of non-contrastive, phonetic features is involved in substitution.

Recall as well that reports from production research have indicated that Japanese and European French speakers tend to substitute [s,z] for English [θ,ð] (e.g. Hancin-Bhatt 1994a,b; Wenk 1979; Brannen 1998), while Québec French speakers tend to substitute [t,d] (e.g. LaCharité & Prévost 1999, Brannen 1998). For Russian speakers, some sources report [t,d] as a production substitute (e.g. Weinberger 1990), while others report [s,z] (Teasdale 1997).

Traditional phonological theory holds that the phonemic (lexical) level of representation is underspecified for non-distinctive features (e.g. Steriade 1987; Calabrese 1988; Mester & Itô 1989; Rice & Avery 1995). Since all languages investigated in this thesis have /t d s z/ in their phonemic inventories, under this view of phonemic representation, it was predicted in Chapters 3 and 4 that EF, QF, and JA would perceptually replace English [θ] with /s/ in phonological processing; whereas for Russian, there are three possible substitutes. These predictions are displayed in the tables below.

(1) Predictions for EF, QF, and JA Based on a Phonemic Assessment

Intake **{θ}**	CONTINUANT CORONAL	
Potential Substitute	*Mismatches*	*# of Featural Conflicts*
☞ /s/		0
/t/	STOP	1
/f/	LABIAL	1
/ʃ/	POST-ALVEOLAR	1

(2) Predictions for RU Based on a Phonemic Assessment

Intake **{θ}**	CONTINUANT CORONAL MELLOW PLAIN	
Potential Substitute	*Mismatches*	*# of Featural Conflicts*
☞ /f/	LABIAL	1
☞ /t/	STOP	1
☞ /s/	STRIDENT	1
/ʃ/	STRIDENT POST-ALVEOLAR	2

Recall that the observation that two languages or dialects sharing the same or similar underlying phonemic inventory display differential substitution led to the hypothesis that learners must make cross-linguistic equations based on more than distinctive features alone. Thus, in conjunction with the hypothesis that production errors reflect perception errors, it was supposed that L2 listeners assess a target sound at a phonetic level of representation, where non-contrastive features are also available. In other words, choice of transfer segment is determined at the phonetic level of representation.

This was tested with both real and nonce words in the AXB experiments detailed in Chapters 3 and 4. Results from these studies revealed inter-language differences in the ability to perceive certain contrasts, including contrasts involving real words. For example, the AXB-2 task showed that Beginner/Intermediate Japanese listeners had higher error rates on real words involving [s-θ] and [s̪-θ] in relation to their Québec French counterparts. Similarly, European French listeners had more errors than Québec French listeners on [s̪-θ] in real words. The QF results allude to the possibility that some non-contrastive features may be specified in lexical representations. The present study is a lexical task, so it will determine whether such cross-linguistic differences emerge in phonological processing.

The contrast of particular interest here is [s,z] versus [θ,ð]. None of the test languages makes a contrast between [s,z] and [θ,ð], which is considered to be based on the features STRIDENT versus MELLOW (e.g. Lahiri & Evers 1991, Hall & Żygis 2010). According to underspecification theory, if these features do not function contrastively, they are absent from the phonological inventory. In such cases, these pairs will be confused in phonological processing.

On the other hand, if some non-contrastive features, such as STRIDENT, *are* available at the phonological level of representation, and if enhancement and muting effects tied to these features are also phonologically active, then the following scenario would be expected.

If the non-contrastive feature STRIDENT is represented phonologically, QF should distinguish [s,z]-[θ,ð]; EF and RU should have more difficulty than the QF group, since EF and RU are considered to have a muted version of STRIDENT (DENTAL MUTES STRIDENT); and JA, which is proposed not to have STRIDENT in either its phonetic or phonological inventory, should have the most difficulty with the [s,z] versus [θ,ð] contrast. The current experiment targets this question.

In this chapter, we will further be able to compare Beginner/Intermediate (BI) learners with Advanced (A) learners. This will enable us to see whether learning takes place.

The PicID also looks at voicing differences. We saw in Chapter 3 that voiced fricative contrasts seemed to facilitate discrimination. The present chapter will further examine this question.

Another variable tested in the PicID task, which was not tested in the other two experiments, is the effect of syllable position on the ability to discriminate features. In addition to onset position, some contrasts are also investigated in coda position. Onset position is generally considered to be the phonologically unmarked position (Clements 1990). As well, cues to fricative identification may be stronger in onset position. Solé (2003) found reduced amplitude of frication for coda as opposed to onset fricatives. For these reasons, it is predicted that simple onset position will have the most facilitative effect on discrimination.

Finally, another contrast was included in the current design: the post-alveolar fricatives were paired with the interdental fricatives. Although these have never been reported as substitutes in either the production or perception literature, they serve as a comparison with anterior [s,z]. The interdental fricatives and [s,z] share the feature

ANTERIOR; whereas, [ʃ,ʒ] is [-anterior] (POST-ALVEOLAR), and as such [ʃ,ʒ]-[θ,ð] should be more easily discriminated than [s,z]-[θ,ð].

5.3. Stimuli

For the PicID Task, only real English words were used. The words used are those which are easily picturable. An effort was made to select a simple, clear picture which was prototypical of the concept that the word depicts. All words chosen constitute a minimal pair with a real English word containing an interdental fricative, e.g. [s] vs. [θ] as in "saw" vs. "thaw". One image was selected for each word, except in cases of inflectional variations on the stem. In these cases, the same image was used for each of the inflectional variants, e.g. the same image was used for the words "tease" and "teasing".

As much as possible, well-known English words were selected. However, good minimal pairs involving the interdental fricatives are quite difficult to come by, so several pairs included some rather infrequent words. However, the participants were trained on these, and results from the training suggest they had internalized the meanings of all stimulus words.

The words chosen were no longer than two syllables; most were monosyllabic. The words used in the PicID task are in (3). The shaded cells indicate conditions for which no picturable minimal pairs could be found.

(3) Stimuli for Picture Identification Task

Vowel/ Position	f-θ	s-θ	t-θ	ʃ-θ
HF-Onset	feces-theses fin-thin	serum-theorem sink-think	tin-thin	shin-thin
NH-Onset	fought-thought four-Thor	sought-thought saw-thaw sigh-thigh sore-Thor sum-thumb	taught-thought tie-thigh tore-Thor torn-thorn tug-thug	Shah-thaw shore-Thor shorn-thorn shot-thought shy-thigh
HF-Coda	reef-wreath			
HB-Coda			boot-booth	
NH-Coda	deaf-death	bass-bath mass-math moss-moth mouse-mouth	bat-bath mat-math rat-wrath	bash-bath rash-wrath
		z-ð	**d-ð**	**ʒ-ð**
HF-Coda		tease-teethe		
NH-Coda		bays-bathe	laid-lathe	beige-bathe

Two talkers (both linguists) each recorded the entire list of words. One of the talkers was male and the other female (the author). Care was taken to pronounce each word with the same falling intonation contour.[71]

The following variables were tested: Language (European French (EF), Québec French (QF), Japanese (JA), Russian (RU), and Native English (NE)), Proficiency level (Beginner-Intermediate (BI) and Advanced (A)),

[71] Words were not produced within a carrier sentence.

Syllable position (Simple Onset, Simple Coda), and Voicing (Voiceless and Voiced), Vowel (High Front (HF), Non-High (NH), and High Back (HB)).

The following tables graphically present these test variables. The table on the left shows between-subjects factors, and the table on the right displays within-subjects factors. Shaded cells indicate that no stimuli were created for that condition.

(4) Test variables

Between-Subjects	
Language	**Proficiency**
EF	BI
	A
QF	BI
	A
JA	BI
	A
RU	BI
NE	n/a

Within-Subjects			
Position	**Voicing**	**Contrast**	**Vowel**
Onset	Voiceless	f-θ	HF
			HB
			NH
		s-θ	HF
			HB
			NH
		t-θ	HF
			HB
			NH
		ʃ-θ	HF
			HB
			NH
Coda	Voiced	d-ð	HF
			HB
			NH
		z-ð	HF
			HB
			NH
		ʒ-ð	HF
			HB
			NH
	Voiceless	f-θ	HF
			HB
			NH
		s-θ	HF
			HB
			NH
		t-θ	HF
			HB
			NH
		ʃ-θ	HF
			HB
			NH

5.4. Participants

There were 73 participants in the PicID task. Learners were classified either as Beginner-Intermediate or Advanced according to the aural comprehension portion of Michigan English Placement Test. Forty-six of these participants also participated in the AXB-2 experiment.

The Beginner-Intermediate group consisted of native speakers of European French (EF) (N = 11), Québec French (QF) (N = 9), Japanese (JA) (N = 8), Russian (RU) (N = 8), as well as a native English control group (NE) (N = 9). The Advanced group consisted of native speakers of EF (N = 9), QF (N = 11), JA (N = 8), and the same NE control group. It was difficult to locate an adequate number of Russian speakers who were at an advanced proficiency level in English, so it was decided not to include an advanced RU group.

Participants were mainly recruited through Montréal universities, and thus were students in their 20s and 30s. However, some participants were not university students (this is particularly true of the QF, JA, and RU groups).

5.5. Training

Participants were initially trained on each of the image/word items. First, the meaning of word depicted in each picture was explained. Then, to further familiarize the participant with the picture-word association, each image was presented with its corresponding aural stimulus. The participant's task was to simply watch and listen. Next, in the active training session, the participant was presented with a pair of pictures presented simultaneously, side by side on a computer screen; then one word was presented aurally to the participant through headphones. The talker's voice in the aural presentation was random, i.e. sometimes the male and sometimes the female. The participant had to select the picture which corresponded to the aurally presented word by pressing the appropriate key on the computer keyboard, the key labelled “L” (image on the left) or the key labelled “R” (image on the right).[72] If the participant made an error, a beep would sound. For example, the participant would be presented with a picture of a *fin* on the left and a picture of a *mouse* on the right, then would hear the word *mouse*. If the participant pressed the "L" key, he or she would hear a beep, signalling an erroneous choice. The image stayed on the screen until the participant responded.

For the training session, it was ensured that the pairs of words corresponding to the two pictures for each item were not minimal or near-minimal pairs. Care was taken to make the present pairs of images whose associated words were as phonologically distinct as possible.

Two blocks of stimuli were presented. All of the stimuli were presented in each block. Within each block, the target image (i.e. the image that matches the aural stimulus) was on the left for half the items (List A) and on the right for half the items (List B). The pairs of images were the same in Block 1 and Block 2; however, the target was reversed.

All participants did well on the training session, with low error rates - ranging between 0% and 5%. When questioned, participants reported that they felt they had internalized the sound-picture association. Many reported that any errors they did make were due to slips in their responses, i.e. mistakenly hitting the wrong key, and not to a failure to associate a word with its picture.

---

[72] The labels were affixed to the actual "A" and "L" on the keyboard.

5.6. The PicID Experiment

After the training session and a 15 minute pause, participants continued with the test component of the task. The PicID experiment is similar to the training component; however, in this test component, participants were presented with three images, side by side. They heard one word, and were to select the picture that corresponded to the word that was aurally presented. To do this, they had to hit one key, either the key labelled “L” (left picture), “M” (middle picture), or “R” (right picture). There was a 500 msec delay between the visual presentation and the aural presentation. Accuracy and Response times were recorded, but only accuracy will be reported.[73]

Each item consisted of one minimal pair, e.g. “saw - thaw” and one distracter item, e.g. “bash”.[74] Care was taken to make the distracter as phonologically distinct from the minimal pair as possible.

For the PicID task, there were six blocks of items. In each block, there were 47 items for a total of 282 items. The items within each block were randomly presented. Participants had a short rest (maximum five minutes) between blocks.

Participants were classified as outliers and their data removed if their errors on foils and fillers were more than two standard deviations above the average for their language group and level. An error on a foil is where the participant wrongly chose the foil when presented with a test target, e.g. picking “bash” when presented with the target “thaw” in the item “thaw-saw-bash”. An error on a filler is where the participant wrongly

---

[73] Response times were not analyzed due to time constraints; however, casual observation of the long response times shows that these correspond to items which had a high error rate for each participant.

[74] The distracter item actually functions as part of a minimal pair in another item.

chose a test item when presented with a foil target, e.g. picking “thaw” when presented with “bash” in the item “thaw-saw-bash”. Two participants were removed on this basis, one from the EF group, and the other from the RU group, reducing the total number of participants to 73.

5.7. Results

Results are divided into three sections: Both Proficiencies, Beginner-Intermediate, and Advanced. For the “both proficiencies” section, only EF, QF, and JA were analyzed, since RU and NE had only one proficiency level. In the “beginner-intermediate” section, EF, QF, JA, RU, and NE are compared. In the “advanced” section, EF, QF, JA, and NE are compared, since there was no advanced RU group.

As with AXB-2, the number of levels for each factor was not always even. Therefore, the data were cut up in different ways, yielding a total of 18 data sets. The following is a breakdown of the individual data sets analyzed.[75]

And again, the data were not homogeneous; therefore, both Anovas and Non-Parametric tests or Anovas with Bonferroni correction were run on the data. In order to be considered significant, both of these types of tests had to have shown significance, the Anova at the .05 level and the Non-Parametrics at .005 in cases of multiple comparisons.

---

[75] For the PicID task, the data sets are numbered 1-18. This is to keep them distinct from the data sets in the AXB-2 task, where they are lettered A-R.

5.7.1. Proficiency

The first six data sets examine the difference between Beginner/Intermediate learners and Advanced learners; thus the Russian group and Native English controls are omitted from the analysis, since they had only one level of proficiency.

Data Set 1. **(**EF/QF/JA; BI/A; O/C; Vclss; /f s t ʃ/; NH). For this first data set, Onset (O) and Coda (C) position are examined for contrasts /f-θ/, /s-θ/, /t-θ/, and /ʃ-θ/. An Anova found a significant main effect for Proficiency (Anova, $F(1,50), p < .001$); significance also emerges in the two-tailed Mann-Whitney test ($X^2(1), p = .001$). Figure 5.1 shows that the Beginner/Intermediate listeners make more errors than do the Advanced listeners. There is a general overall improvement, since there are no interactions with other variables.

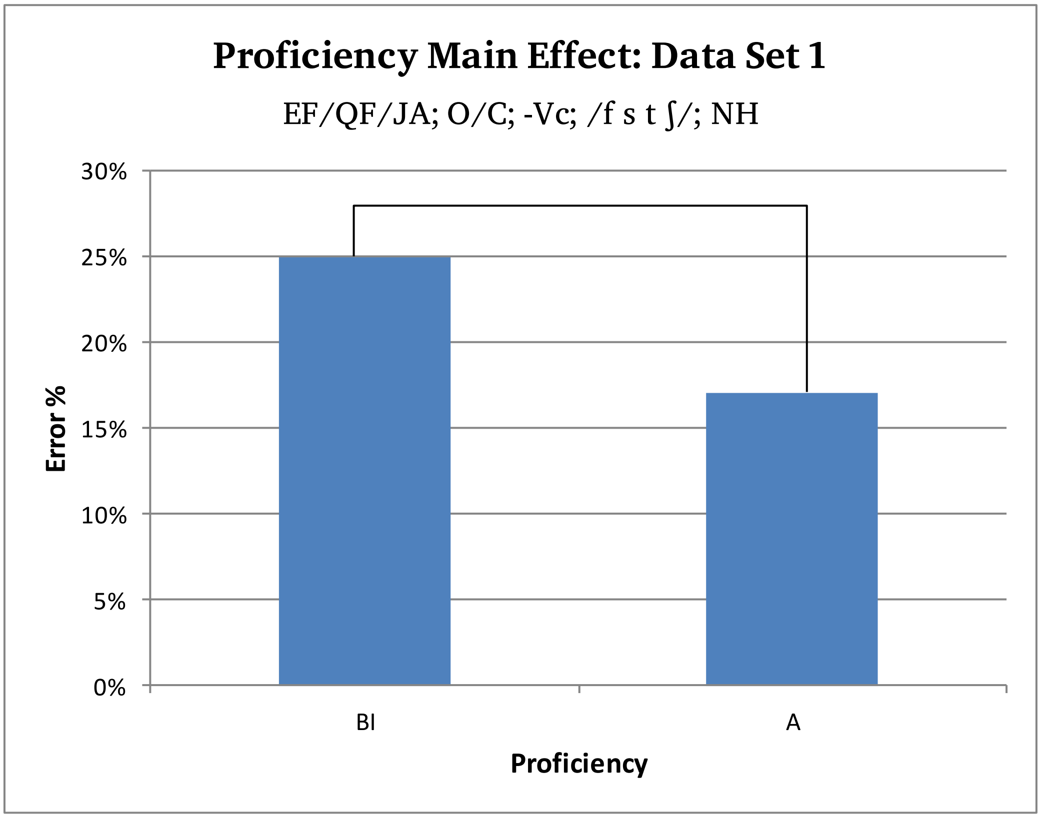

Figure 5.1. Data Set 1. Proficiency Main Effect.
Significant differences are indicated with a connector line.

Data Set 2. (EF/QF/JA; BI/A; O/C; Vclss; /f/; HF/NH). This analysis looks at the contrast /f-θ/ alone, but this time in two vowel contexts: HIGH FRONT and NON-HIGH in both Onset and Coda position.

Again we see that there is an improvement with increased proficiency, (Anova $F(1,50), p < .05$); two-tailed Mann-Whitney ($X^2(1), p = .009$). There were no interactions with Proficiency.

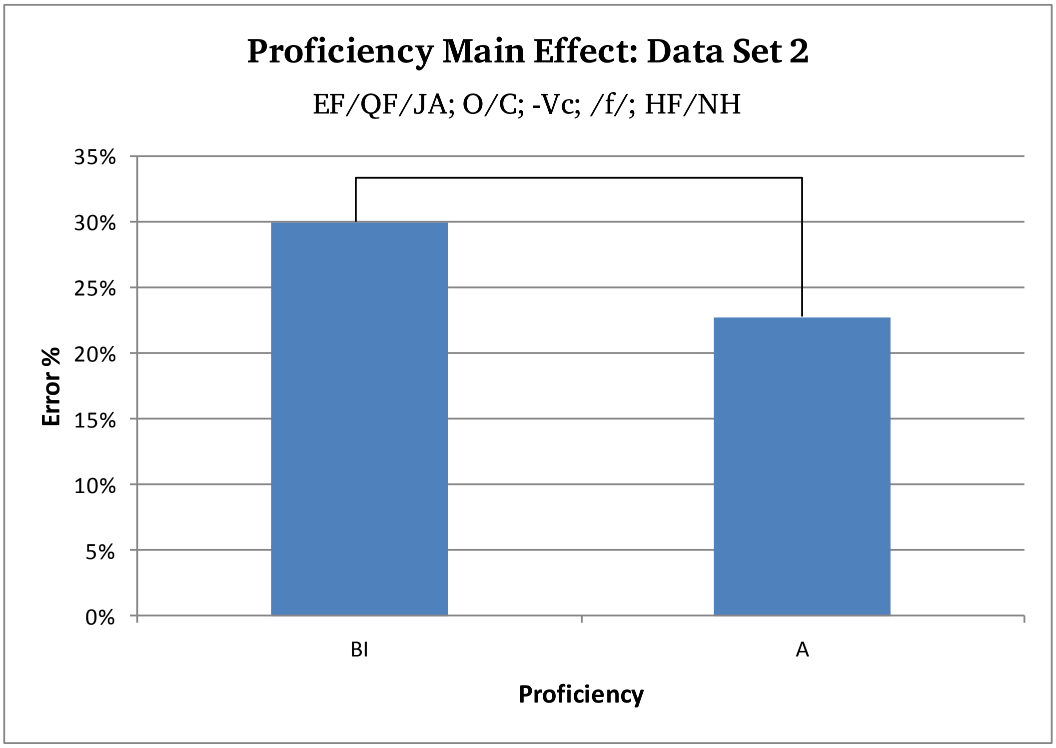

Figure 5.2. Data Set 2. Proficiency Main Effect.

Significant differences are indicated with a connector line.

Data Set 3. **(**EF/QF/JA; BI/A; O; Vclss; /f s t ʃ/; HF/NH). This data set is similar to Data Set 1, except here two Vowel conditions are examined: HIGH FRONT versus NON-HIGH. We see in Figure 5.3 that Proficiency interacts with Vowel and Language (Anova F(2,150),p=.03). This means that differences across Proficiency levels depend on the Language and Vowel involved. A Mann-Whitney test found the difference to stem from the QF group, with a higher error rate for the QF-BI group as compared to the QF-A group, but only significantly so in the context of a HF vowel.

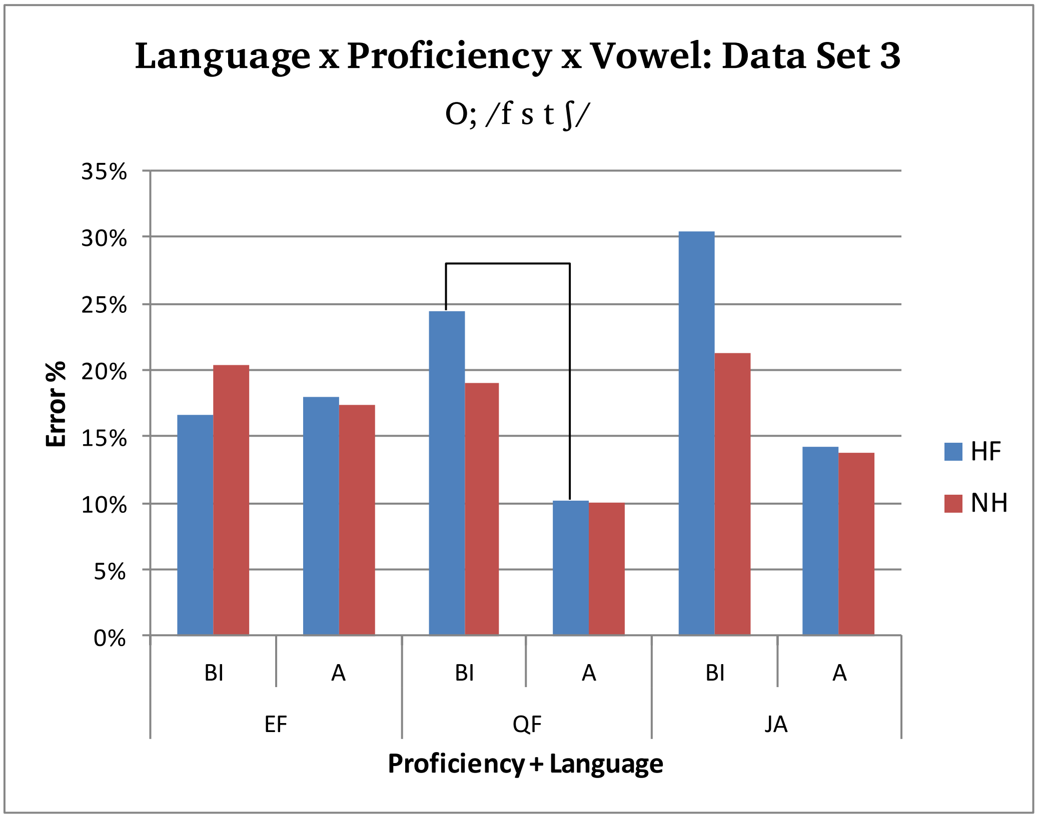

Figure 5.3. Data Set 3. Language x Proficiency x Vowel.
Significant differences are indicated with a connector line.

Data Set 4: **(**EF/QF/JA; BI/A; C; Vcd/Vclss; /t,d s,z ʃ,ʒ/; NH). The previous data sets included Onset position; the remaining data sets in this section examine Coda position alone.

For Data Set 4, both voiced and voiceless contrasts are examined after a non-high vowel: /d-ð/, /t-θ/, /z-ð/, /s-θ/, /ʒ-ð/, and /ʃ-θ/. As shown in Figure 5.4, a main effect for Proficiency emerges, but there are no interactions, Anova $F(1,5), p<.001$; Mann-Whitney ($X^2(1), p=.001$).

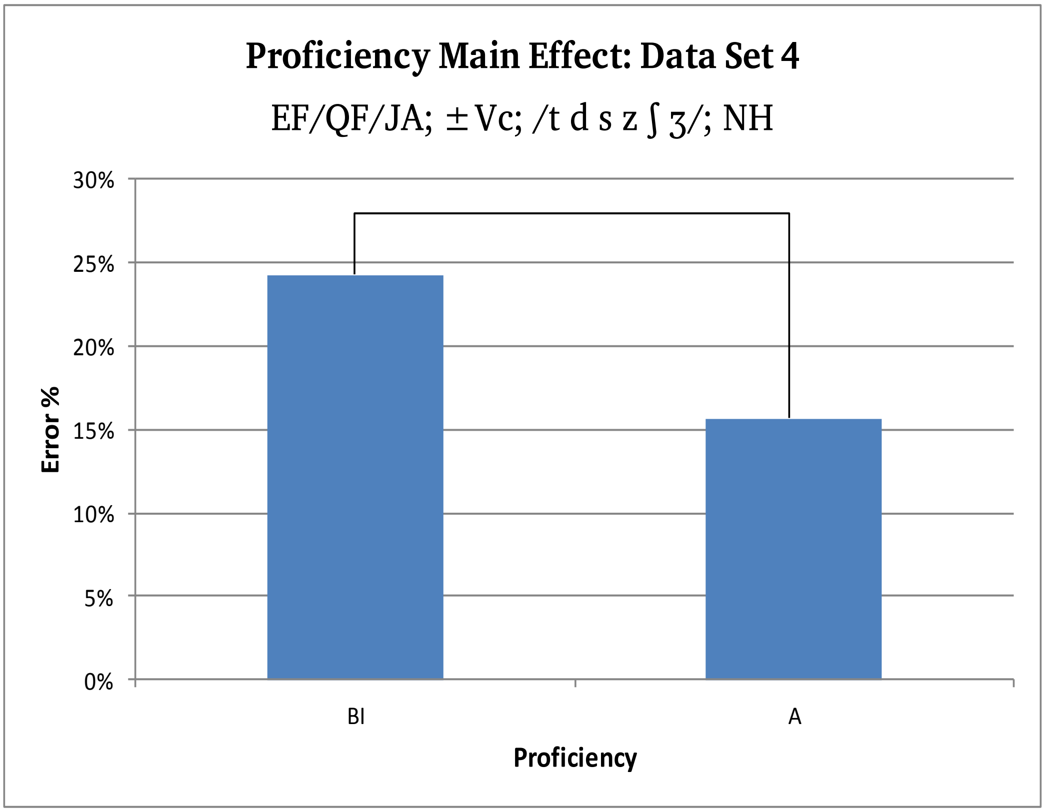

Figure 5.4. Data Set 4. Proficiency Main Effect.
Significant differences are indicated with a connector line.

Figure 5.4 shows an overall improvement for the Advanced learners, as expected.

Data Set 5: **(**EF/QF/JA; BI/A; C; Vcd; /z/; HF/NH). Here, the voiced /z-ð/ contrast alone is examined in Coda position this time, in the context of two vowels: HIGH FRONT vs. NON-HIGH. The statistical analysis did not show any significant differences between Proficiency levels.

Data Set 6: **(**EF/QF/JA; BI/A; C; Vclss; /t/; HB/NH). This is the final data set that investigates Proficiency level. It investigates the Contrast /t-θ/ alone after the two vowels: HIGH BACK and NON-HIGH. Results show a main effect for Proficiency, Anova F(1,50),p=.007, Mann-Whitney

(X^2(1),p = .004). Again we see improvement in the Advanced groups, as shown in Figure 5.5.

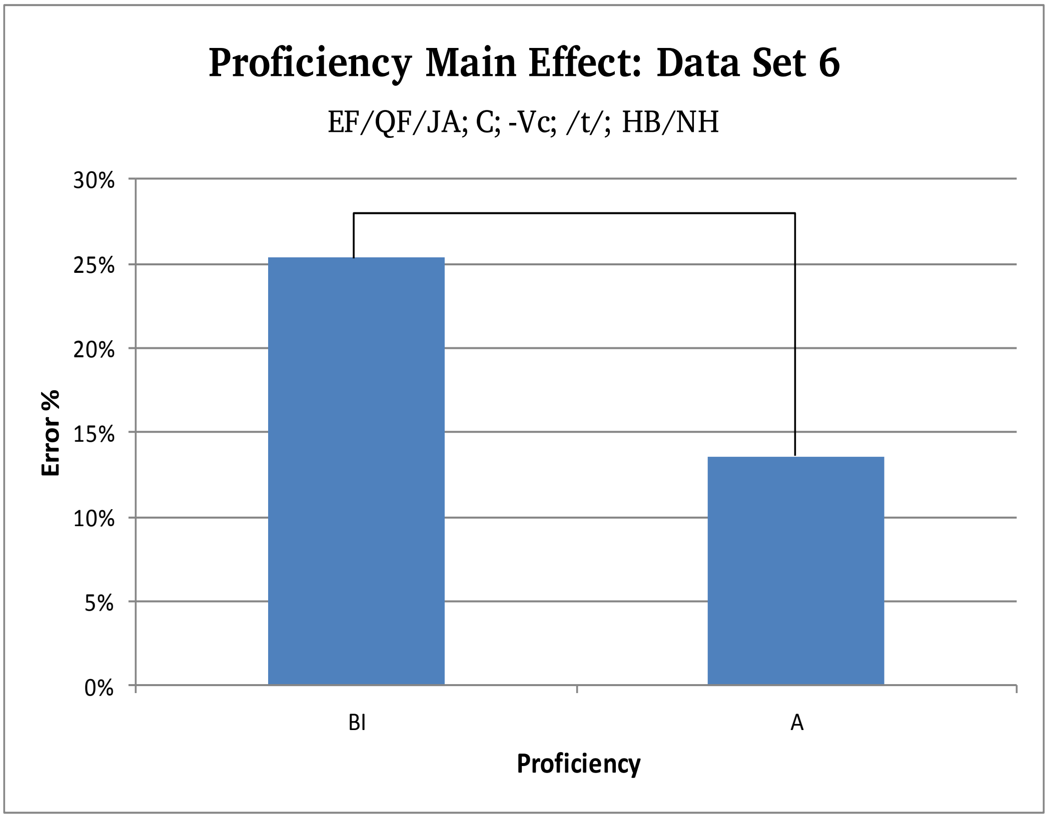

Figure 5.5. Data Set 6. Proficiency Main Effect.
Significant differences are indicated with a connector line.

5.7.1.1. Summary of Results on Proficiency

With the exception of Data Set 5, which looked at /z-ð/ in coda position, all the data sets investigating the effect of proficiency level on the ability to correctly identify the interdental fricative show an improvement for Advanced learners over Beginner/Intermediate learners of English. For Data Set 3 (Figure 5.3), which is the only data set to investigate /f s t ʃ/ versus /θ/ in Onset position alone, results indicate that a contributing factor to this improvement is the context of a high front vowel, especially for the Québec French group, but to a certain extent for the JA group as well. The Beginner/Intermediate learners have more difficulty than the Advanced learners, but especially so with before a high front vowel compared to a non-high vowel. The general finding of improvement in the Advanced groups is an expected result, and indicates that the interdental fricative can be acquired in perception.

5.7.2. Statistical Results for Vowel

This section reports on those data sets which included a comparison between vowel contexts. The one data set in which there was an interaction between Vowel and Proficiency was reported above in Figure 5.3.

5.7.2.1. Vowel Effects in the Beginner/Intermediate Group

Data Sets 8, 9, 11, and 12 examine the influence of different vowel contexts for the Beginner/Intermediate groups. The parameters investigated in these analyses correspond to Data Sets 2, 3, 5, 6 (both proficiency levels). Unlike previous data sets, these include the Native English control group.

Data Set 8: **(**EF/QF/JA/RU/NE; BI; O/C; Vclss; /f/; HF/NH). This analysis, which looks at /f-θ/ alone, reveals a Position x Vowel x Language interaction, Anova $F(4,40), p < .001$; a Kruskal-Wallis test ($X^2(4), p \leq .001$) finds the significance to stem from cross-language differences in Onset before a HF vowel and in Coda before a NH vowel. Two-tailed Mann-Whitney tests show that /f-θ/ is more difficult for the test languages as compared to the English control group when before a high front vowel in

Onset position, but in Coda position, it is the preceding non-high vowel that causes problems.

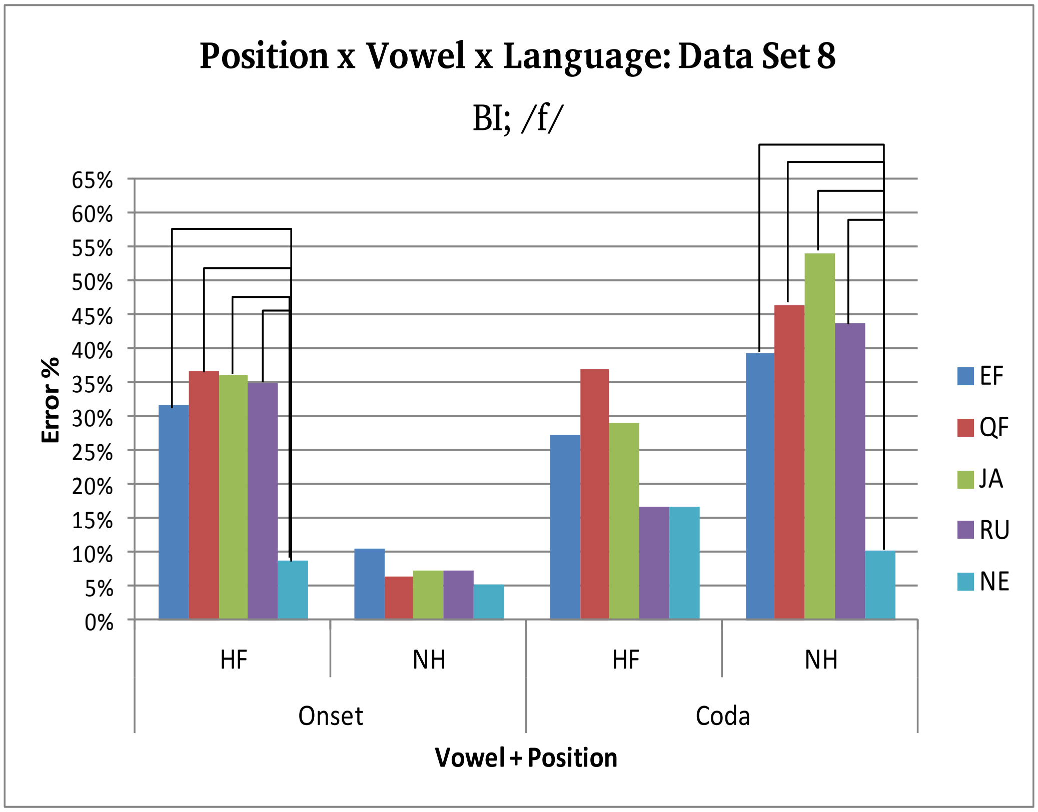

Figure 5.6. Data Set 8. Position x Vowel x Language interaction. Significant differences are indicated with a connector line.

Data Set 9: (EF/QF/JA/RU/NE; BI; O; Vclss; /f s t ʃ/; HF/NH). There was a Language x Vowel interaction here, Anova $F(4,40), p = .001$; Kruskal-Wallis ($X^2(4), p < .001$). Two-tailed Mann-Whitney tests show that the test languages had more difficulty than the NE group in both Vowel conditions, but also that JA has more errors than EF in the HF condition alone.

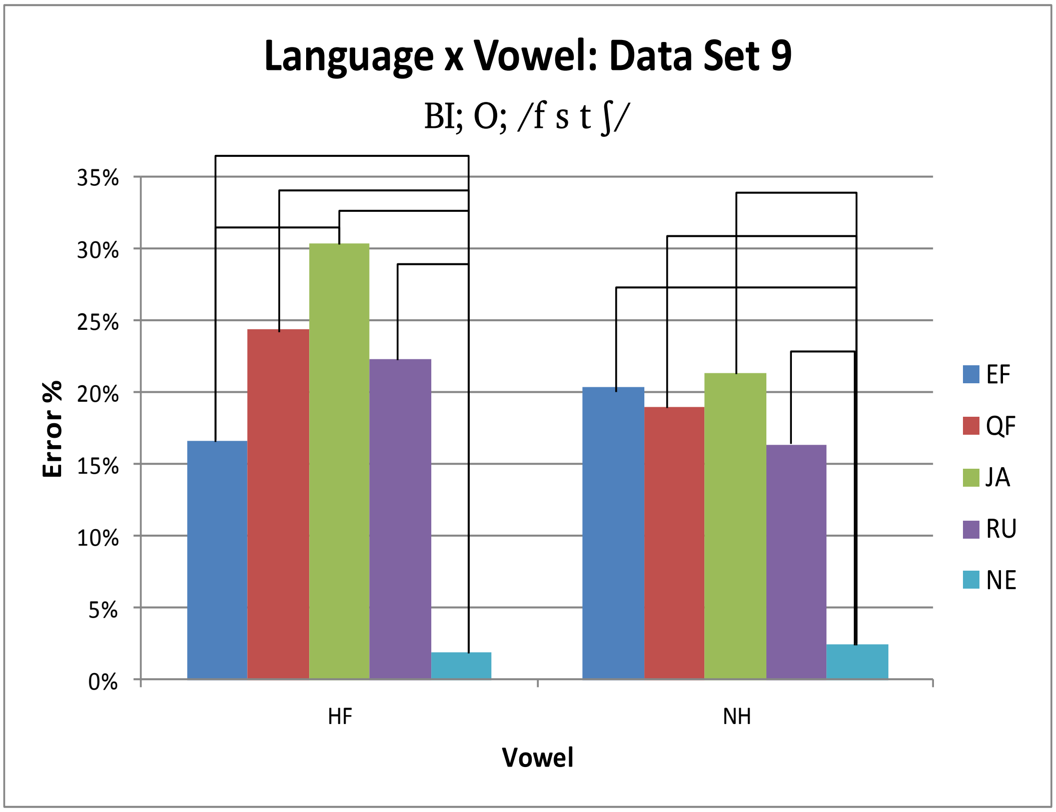

Figure 5.7. Data Set 9. Language x Vowel interaction.
Significant differences are indicated with a connector line.

There was also a Contrast x Vowel interaction, Anova F(3,120),$p<.001$; Friedman ($X^2(7)$,$p<.001$). Two-tailed Wilcoxon tests show that /f-θ/ is more problematic before a HF vowel.

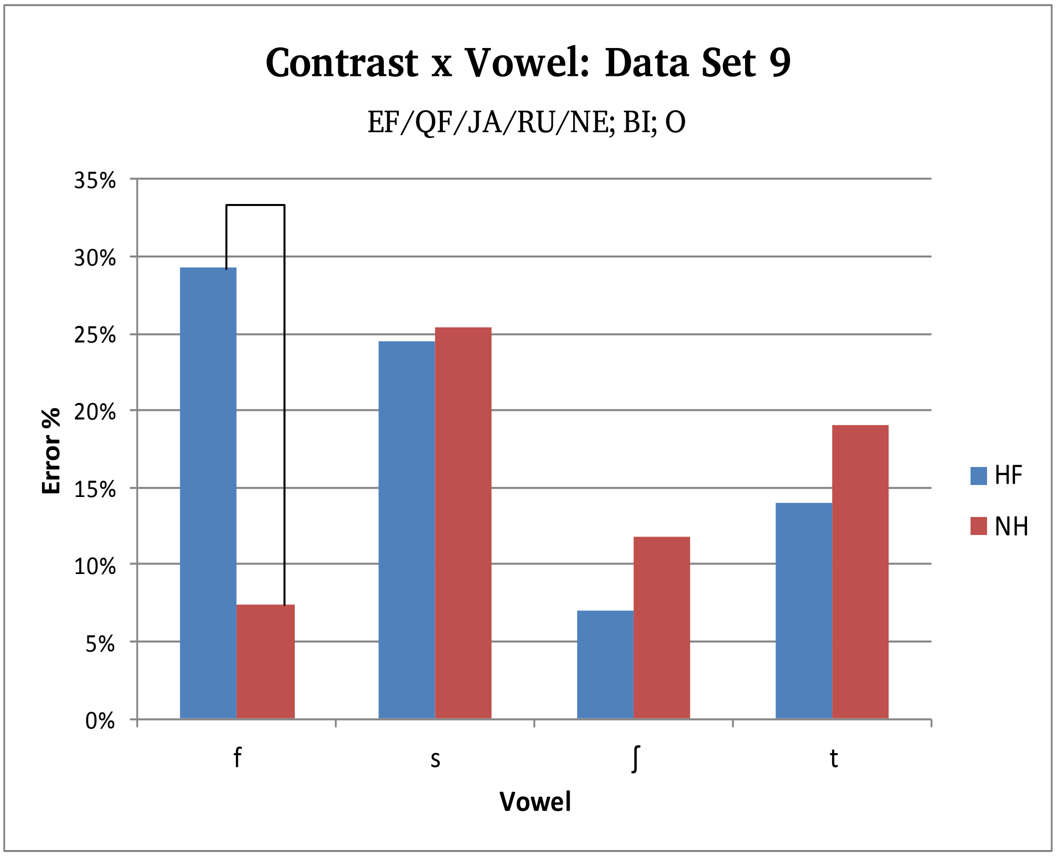

Figure 5.8. Data Set 9. Contrast x Vowel interaction.
Significant differences are indicated with a connector line.

Data Set 11: (EF/QF/JA/RU/NE; BI; C; Vcd; /z/; HF/NH). This analysis examined /z-ð/ in Coda position, and found a significant main effect for Vowel, Anova F(1,40),p = .02. A preceding HF vowel makes /z-ð/ more difficult to discriminate than a preceding NH vowel, as can be seen in Figure 5.9.

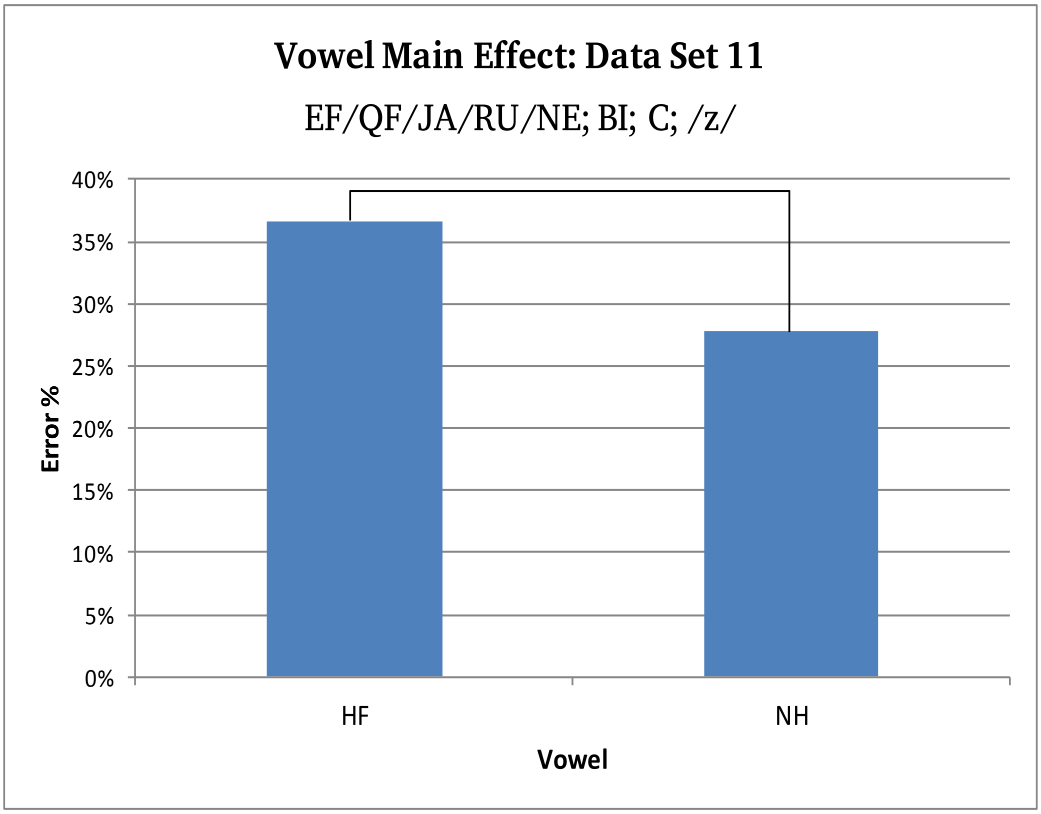

Figure 5.9. Data Set 11. Vowel main effect.
Significant differences are indicated with a connector line.

Data Set 12: (EF/QF/JA/RU/NE; BI; C; Vclss; /t/; HB/NH). No effect or interactions for Vowel emerged in this data set.

5.7.2.2. Vowel Effects in the Advanced Group

Data Sets 14, 15, 17, and 18 examine the Advanced groups. These correspond to Data Sets 2, 3, 5, and 6 (both proficiency levels). The Native English control group is included in these analyses.

Data Set 14: (EF/QF/JA/NE; A; O/C; Vclss; /f/; HF/NH). This investigation of /f-θ/ reveals a Position x Vowel x Language interaction, Anova $F(3,33), p < .05$. Two-tailed Mann-Whitney tests show this interaction to arise from EF having significantly more errors than NE in Onset position before a HF vowel, but in Coda position after a NH vowel.

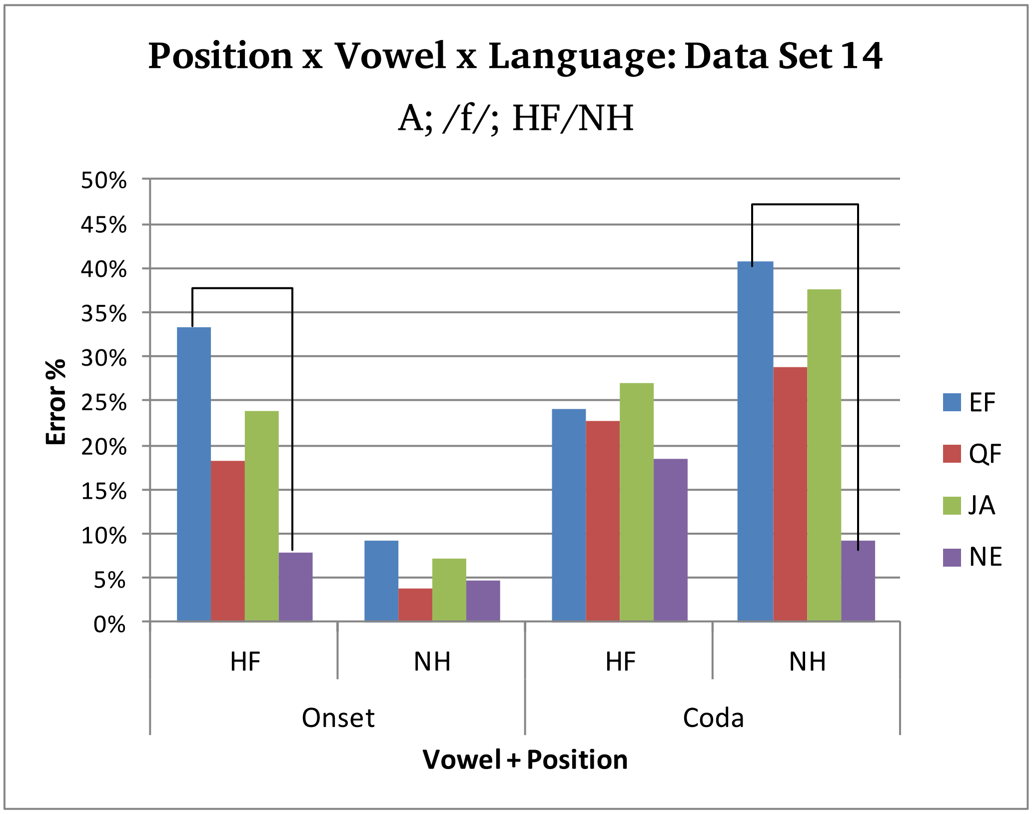

Figure 5.10. Data Set 14. Position x Vowel x Language interaction. Significant differences are indicated with a connector line.

Data Set 15: (EF/QF/JA/NE; A; O; Vclss; /f s t ʃ/; HF/NH)**.** There was a Contrast x Vowel x Language interaction here, Anova F(9,99),p=.003, Kruskal-Wallis (X^2(3),p<.001). Two-tailed Mann-Whitney tests show significant differences on the /f-θ/, /s-θ/, and /t-θ/ contrasts. For /f-θ/, EF has significantly more errors than NE, but only preceding a HF vowel, as we saw above in Figure 5.10. A separate graph is presented below for /s-θ/ and /t-θ/.

For the /s-θ/ contrast, JA has more errors than QF and NE in both vowel conditions. Figure 5.11 shows that EF has a higher error rate than NE in both conditions, and higher than QF before a NH vowel.

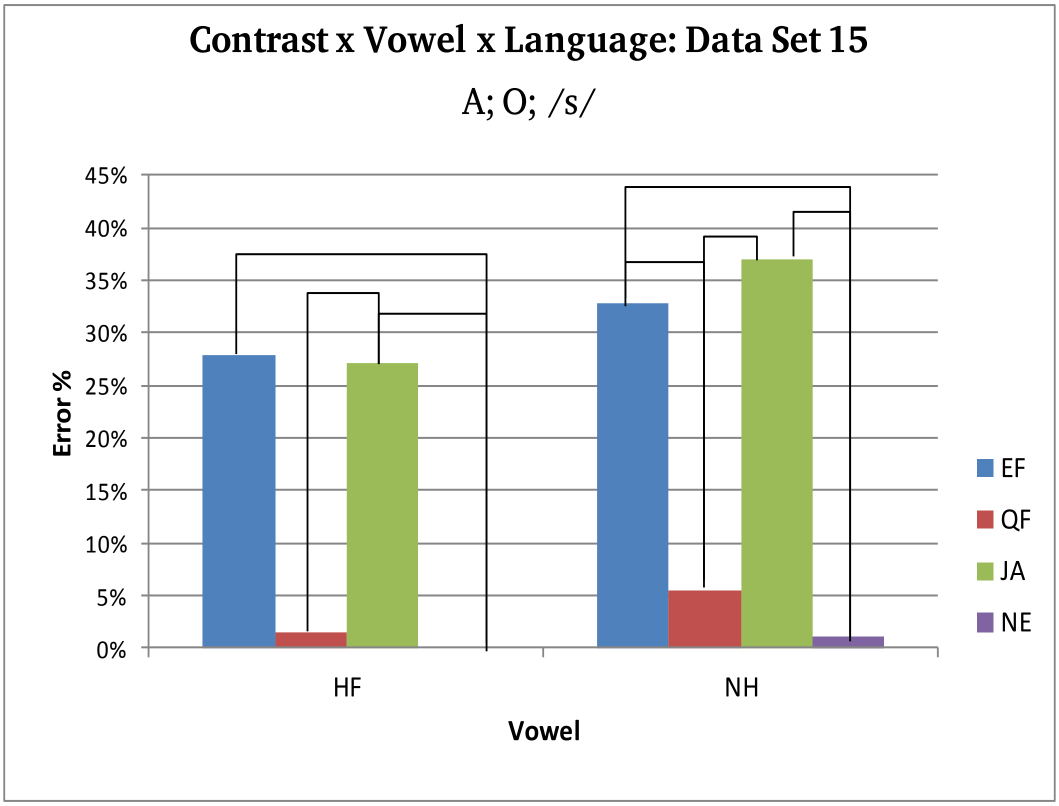

Figure 5.11. Data Set 15. Contrast x Vowel x Language interaction. Significant differences are indicated with a connector line.

For the /t-θ/ contrast, EF and QF have more errors than JA and NE, but only in the NH condition. This is shown in Figure 5.12.

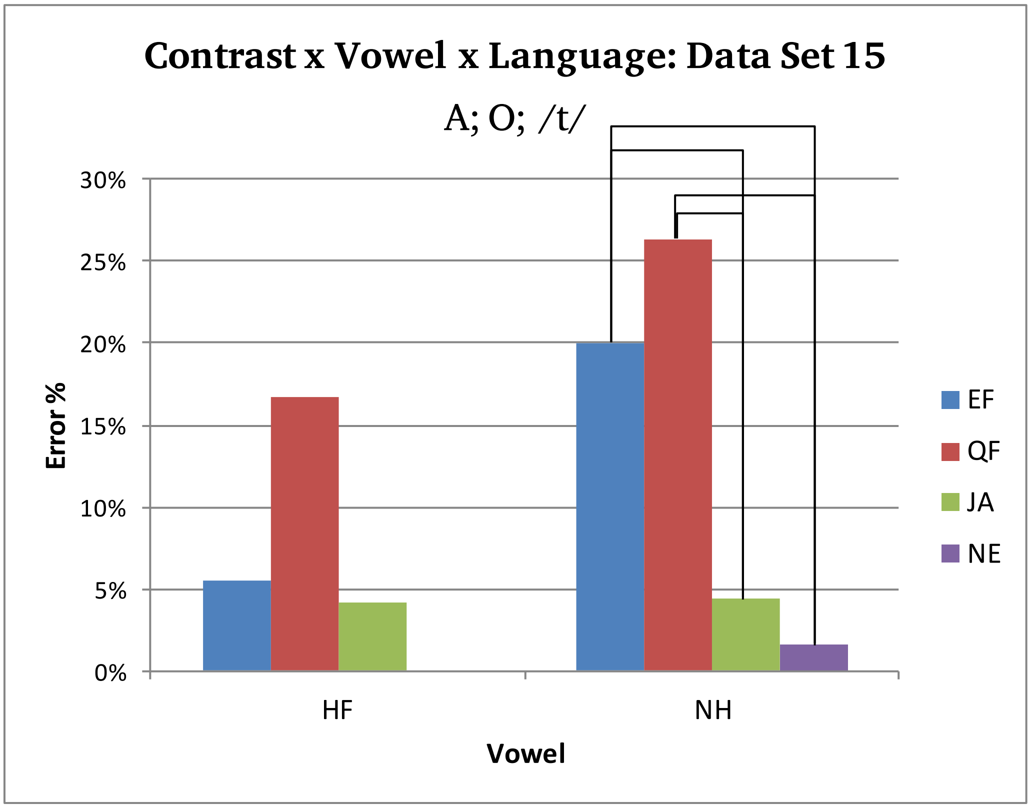

Figure 5.12. Data Set 15. Contrast x Vowel x Language interaction. Significant differences are indicated with a connector line.

There were no significant differences between languages for the /ʃ-θ/ contrast, with all languages having a relatively low error rate (under 8%).

A line graph is provided in Figure 5.13 below to better show the significant interactions from this data set.

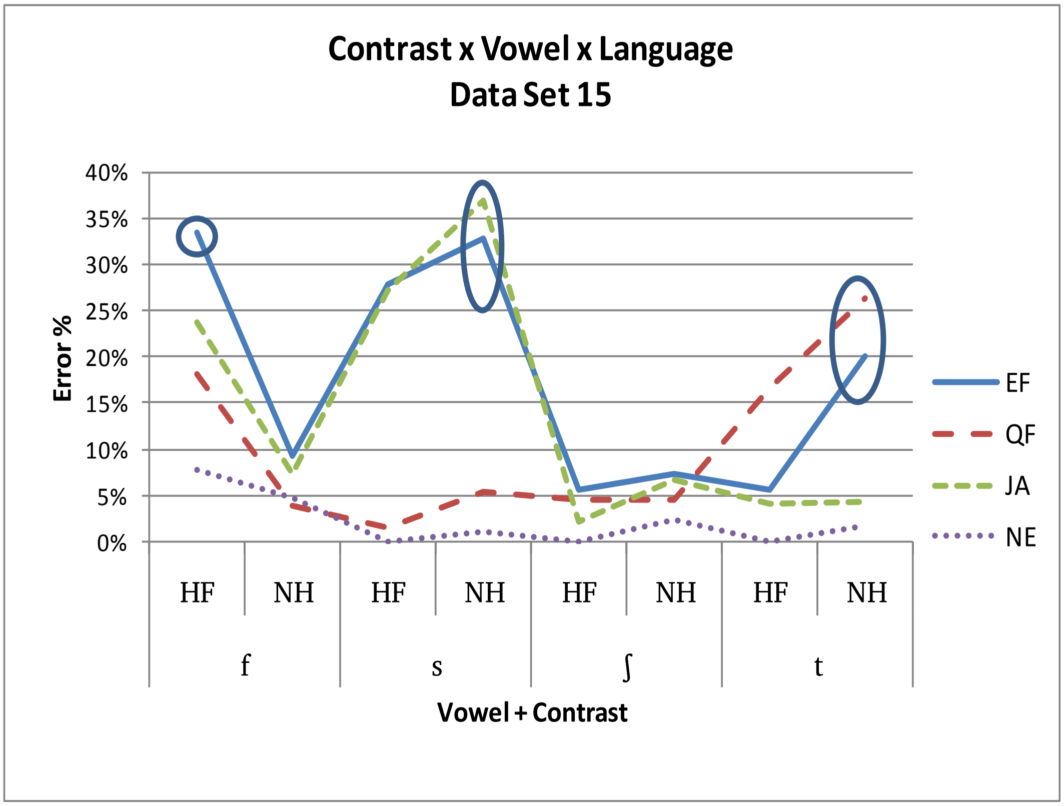

Figure 5.13. Data Set 15. Contrast x Vowel x Language interaction. The source of the interaction is circled.

Data Set 17: (EF/QF/JA/NE; A; C; Vcd; /z/; HF/NH). There was no main effect or interactions for Vowel in this data set.

Data Set 18: (EF/QF/JA/NE; A; C; Vclss; /t/; HB/NH). Again, no significant differences emerged for this data set.

5.7.2.3. Summary of Results for Vowel

This section has examined the influence of vowel context on the ability to perceive contrasts involving an interdental fricative.

What emerges most prominently is the difficulty learners have with the /f-θ/ contrast preceding a high front vowel (/i/ and /ɪ/) as compared to a non-high vowel (/ɑ/ and /o/). This is especially true of the Advanced European French group. As was suggested in Chapter 4, this may be due to the spread lip configuration of the vowel. The labiodental fricative is produced with slight lip rounding. However, before a high front vowel, this rounding may be attenuated, thus removing an auditory cue that helps differentiate /f/ from /θ/. The fact that Native English listeners do not seem to be affected by the vowel when /f-θ/ is in onset suggests that they do not need cues from the vowel to identify these sounds, and that they are able to compensate for any influence of the vowel on the fricative.[76]

In coda position however, the opposite scenario emerges for /f-θ/: For the test groups, this contrast is more difficult after a non-high vowel (/ɛ/) as compared to a high front vowel (/i/). It is unknown why this occurs, although I speculate that it may be related to vowel length, with perhaps more information available on the long vowel (/i/).

Finally, this section has shown that the /t-θ/ contrast is more difficult before a non-high vowel /ʌ aɪ o ɑ/ than it is before a high front vowel /ɪ/. This is due to the Advanced EF and QF groups having more difficulty than the Advanced JA and Native English groups. In fact, an adjacent high front vowel is the conditioning environment for affrication in QF. In Chapter 4, I discussed results indicating that when a sound that has allophonic status in the L1 appears in a context which is incompatible with the L1 conditioning environment, this causes difficulties for listeners from that L1 background. Thus, in the present case, we might have expected QF to have more problems with /t/ in the context of a high vowel, as this is the position where [ts] occurs in QF.

---

[76] These findings for Native English do not coincide with those of Jongman (1989). He found that the interdental fricative was more poorly identified before [a] than before [i u].

5.7.3. Statistical Results for Contrast

This section gives results involving the Contrast factor. Any interactions of Contrast with Proficiency or Vowel were reported in previous sections.

5.7.3.1. Contrast Effects for Beginner/Intermediate Groups

Data Set 7: **(**EF/QF/JA RU/NE; BI; O/C; Vclss; /f s t ʃ/; NH). In this data set, there is a Contrast x Position x Language interaction, Anova $F(12,120), p<.001$, Kruskal-Wallis ($X^2(4), p<.003$). This analysis shows that /f-θ/ in Coda presents particular difficulty for all learner groups. This was seen in Figure 5.6 in §5.7.2.1.

For /s-θ/, though, there is no difference between positions: in Onset and Coda position, all language groups except QF have significantly more errors than the NE control group. Also, JA surpasses QF in error rate.

Results for the /ʃ-θ/ contrast are opposite to those of /f-θ/ in that Onset position is most difficult for EF, QF, and JA.

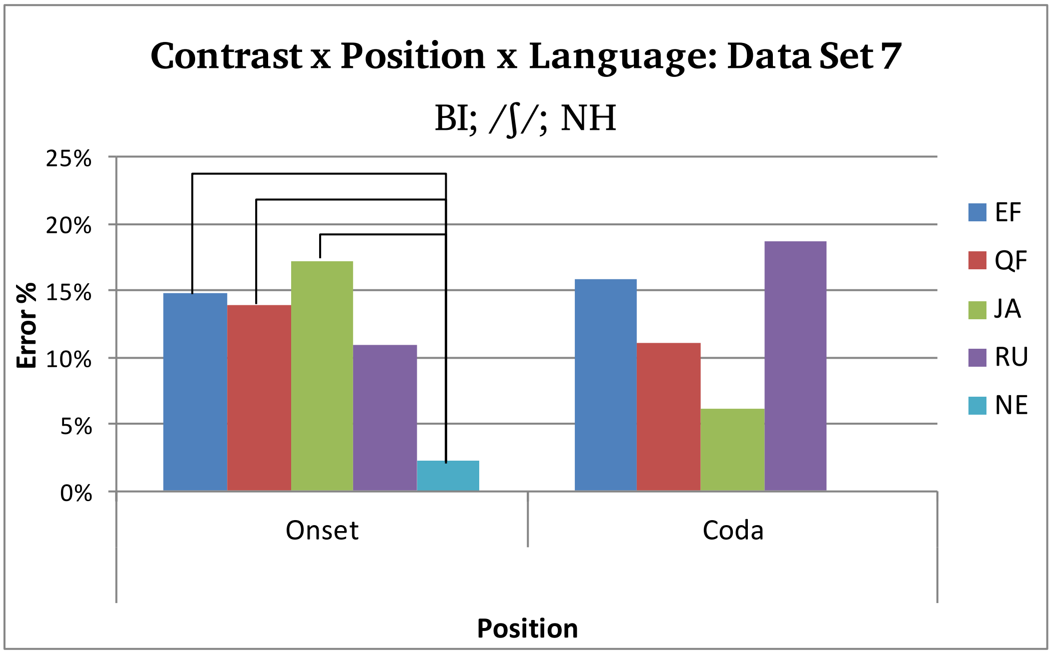

Figure 5.14. Data Set 7. Contrast x Position x Language interaction. Significant differences are indicated with a connector line.

For /t-θ/, as with /s-θ/, position plays no role; thus, there is only a Contrast x Language interaction. All languages except JA have more errors than NE on /t-θ/, and QF experiences significantly more difficulty than EF and JA.

A line chart of this interaction is provided in Figure 5.15.

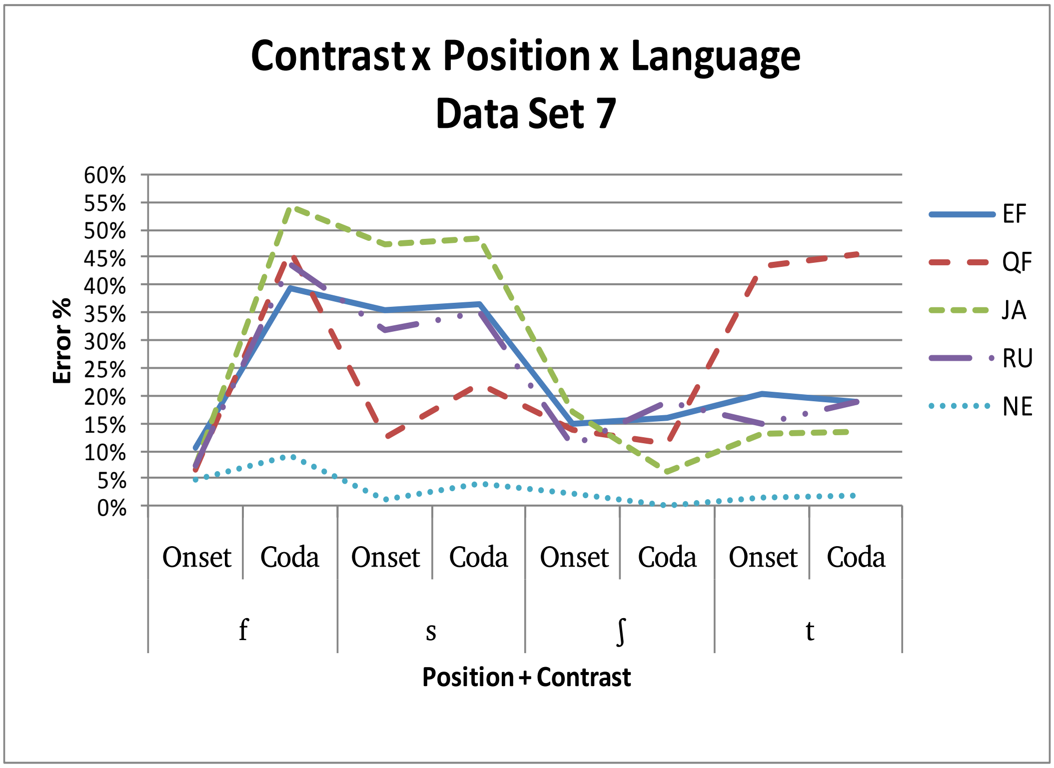

Figure 5.15. Data Set 7. Contrast x Position x Language interaction.

Data Set 9: (EF/QF/JA/RU/NE; BI; O; Vclss; f/s/t/θ; HF/NH). The Anova results indicate a Language x Contrast interaction for this data set, Anova F(12,120),p < .001; Wilcoxon ($X^2(4)$,p < .003). Two-tailed Mann-Whitney tests show the following significant differences. For /f-θ/, all learner language groups have more errors than the Native English control group. For /s-θ/, European French, JA, and RU all have more errors than the NE group; additionally, European French and JA have more errors than QF. For /ʃ-θ/, all learner groups except RU have more errors than NE; additionally, Japanese has more errors than EF. For /t-θ/, Québec French has more errors than all other language groups, while European French and RU have more errors than NE.

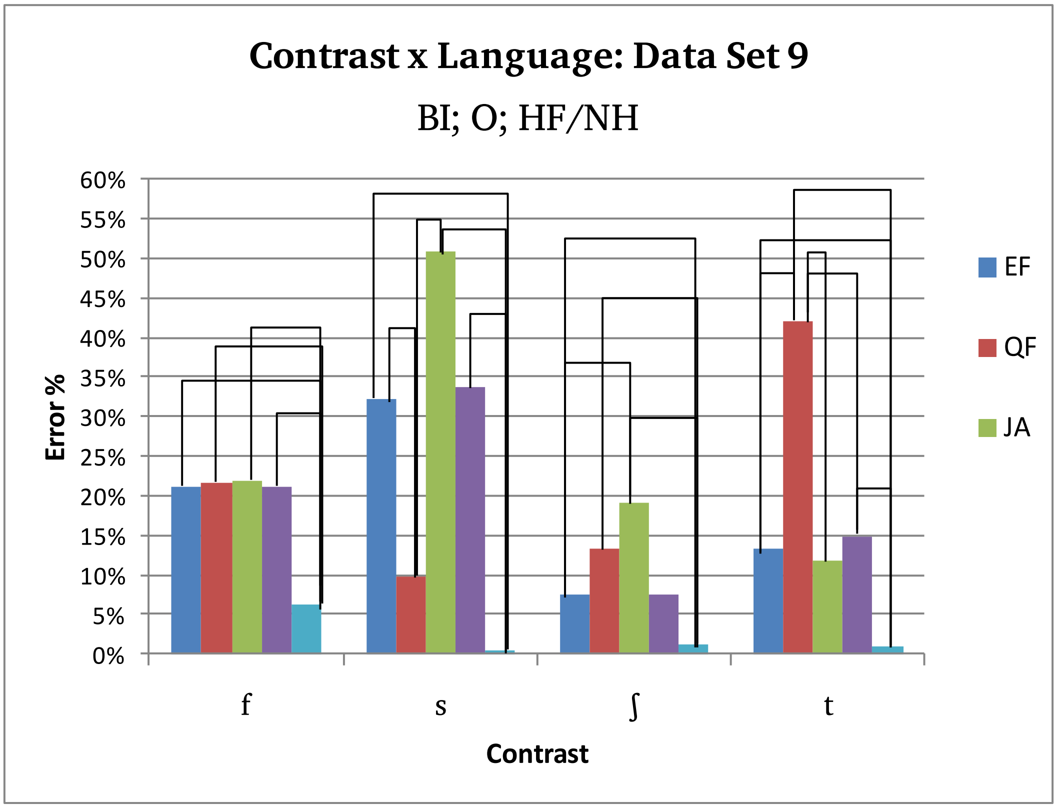

Figure 5.16. Data Set 7. Contrast x Language interaction.
Significant differences are indicated with a connector line.

Data Set 10: (EF/QF/JA/RU/NE; BI; C; Vcd/Vclss; t/d s/z ʃ/ʒ; NH). This data set looks at voicing contrasts in Coda. Statistical tests reveal a Contrast x Language interaction here, Anova F(8,80),p<.001; Friedman (X^2(4),p<.001). Two-tailed Mann-Whitney tests show the following significant differences. For /t-θ/ and /d-ð/, all test languages have more errors than NE. Additionally, Québec French has more errors than EF and RU. For /s-θ/ and /z-ð/, all test languages have more errors than NE; and, for /ʃ-θ/ and /ʒ-ð/, Québec French has more errors than NE.

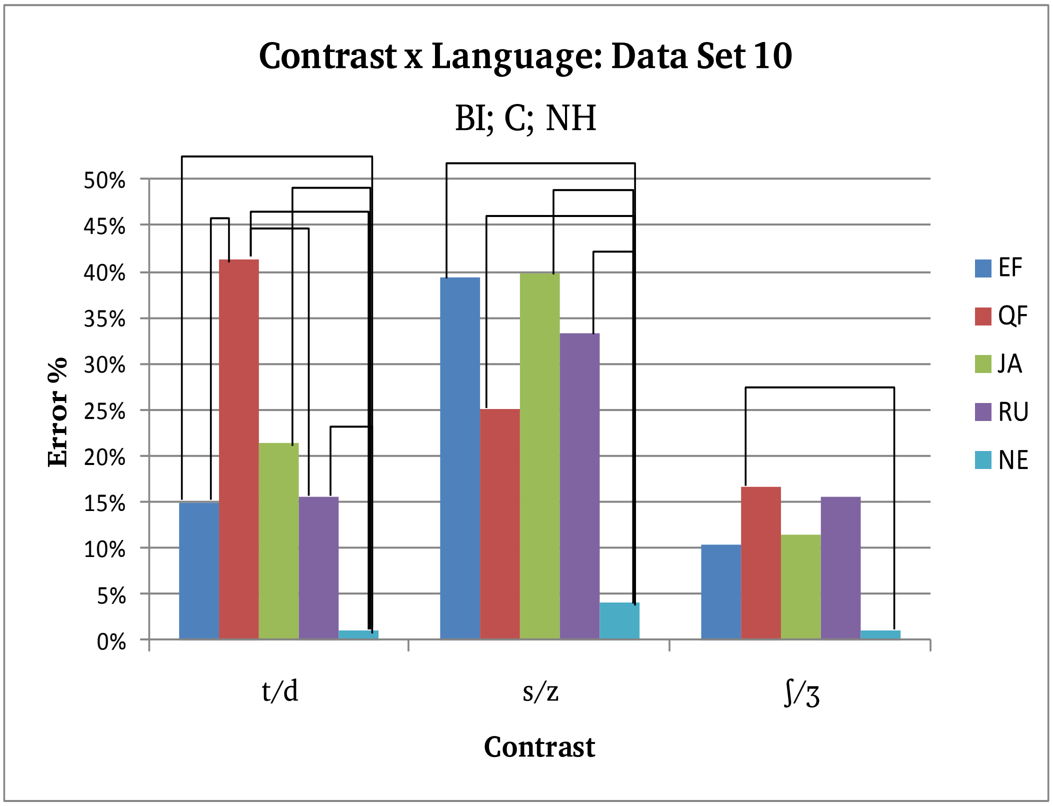

Figure 5.17. Data Set 10. Contrast x Language interaction.
Significant differences are indicated with a connector line.

5.7.3.2. Contrast Effects for Advanced Groups

Data Set 13: EF/QF/JA/NE; A; O/C; Vclss; /f s t ʃ/; NH. A Contrast x Position x Language interaction emerges from this analysis, Anova F(9,99),$p<.05$; Kruskal-Wallis ($X^2(3)$,$p<.001$). The significant differences revealed by two-tailed Mann-Whitney tests are shown in Figures 5.18 and 5.19 below.

For /f-θ/ and /ʃ-θ/, EF has significantly more errors than NE, but only in Coda position.

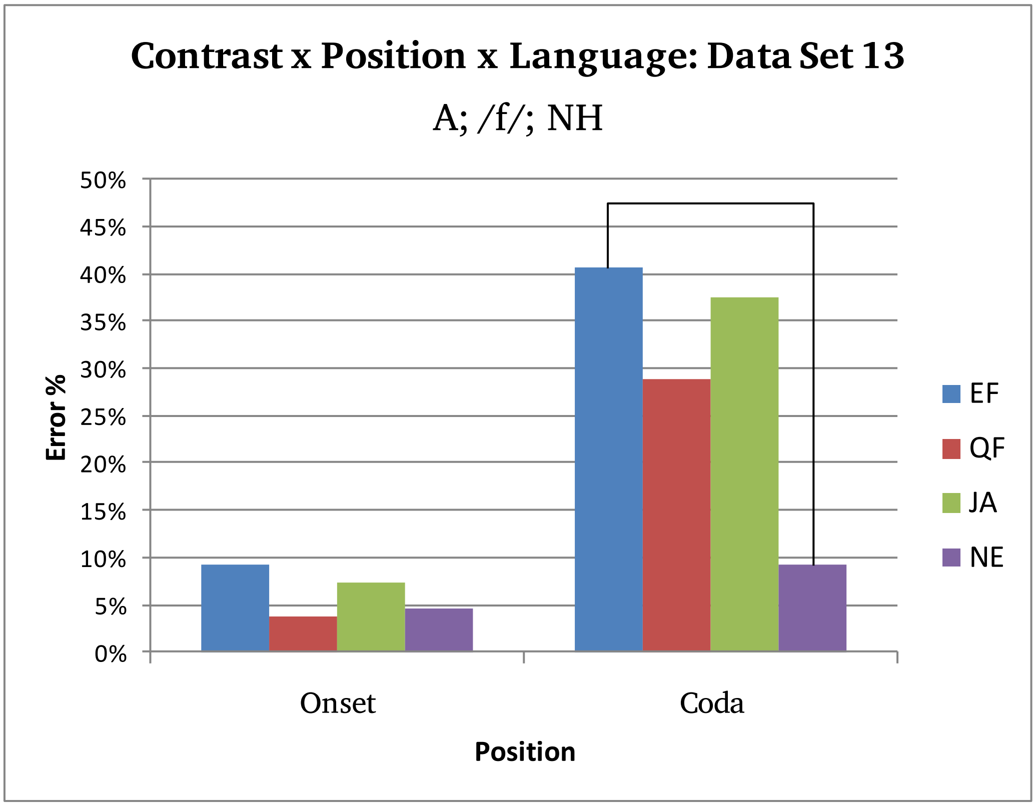

Figure 5.18. Data Set 13. Contrast x Position x Language interaction. Significant differences are indicated with a connector line.

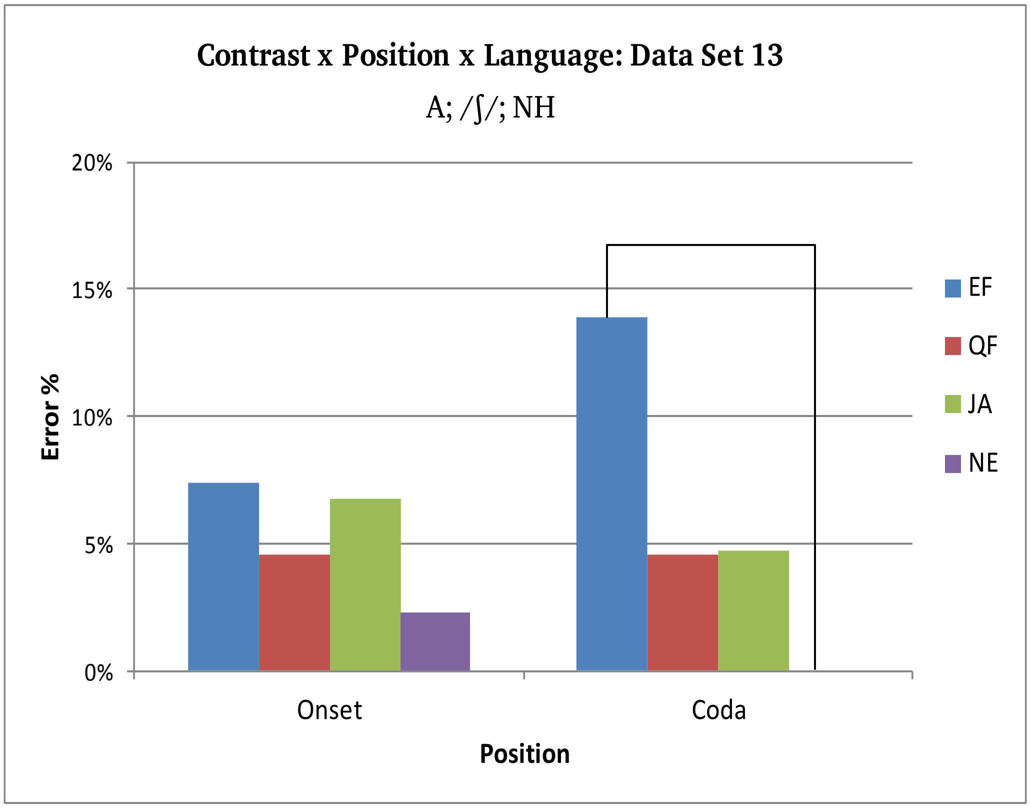

Figure 5.19. Data Set 13. Contrast x Position x Language interaction. Significant differences are indicated with a connector line.

For /s-θ/, JA has more errors than both QF and NE in both positions. European French also has significantly more errors than QF and NE in Onset, but more than NE alone in Coda position.

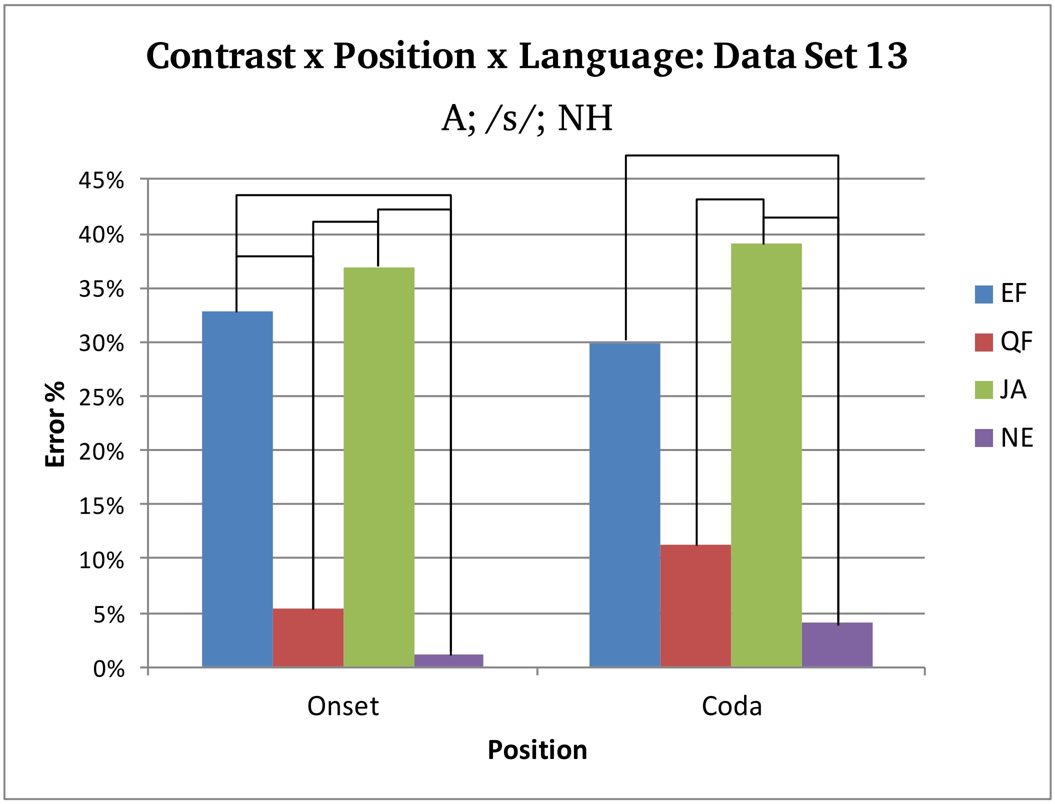

Figure 5.20. Data Set 13. Contrast x Position x Language interaction. Significant differences are indicated with a connector line.

For /t-θ/, there is no interaction with position, but EF and QF both have more errors than JA and NE.

The Contrast x Position x Language interaction is shown in Figure 5.21 below.

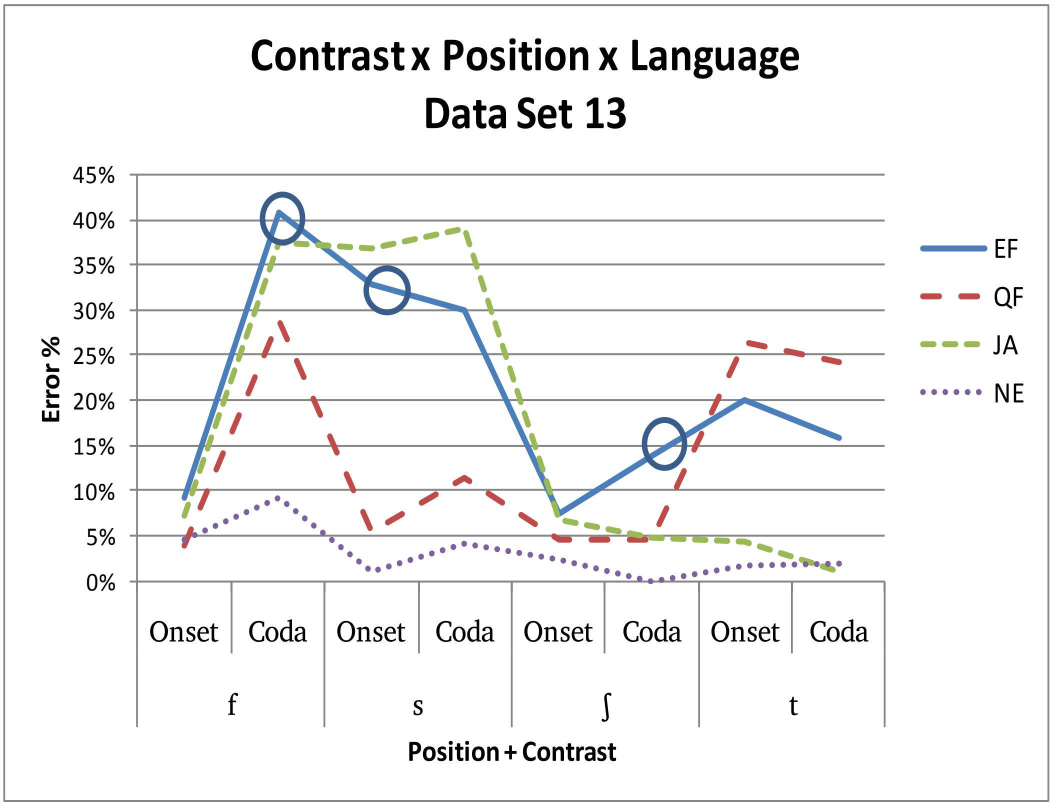

Figure 5.21. Data Set 13. Contrast x Position x Language interaction.

Data Set 16: (EF/QF/JA/NE; A; C; Vcd/Vclss; /t,d s,z ʃ,ʒ/; NH). Again here, we observe a Contrast x Language interaction, Anova F(6,66),p<.001, Kruskal-Wallis (X^2(3),p=.003). Two-tailed Mann-Whitney tests show that both EF and QF have more errors than NE on /t-θ/. For /s-θ/, JA has more errors than both QF and NE, and EF has more errors than NE. This is seen in Figure 5.22 below.

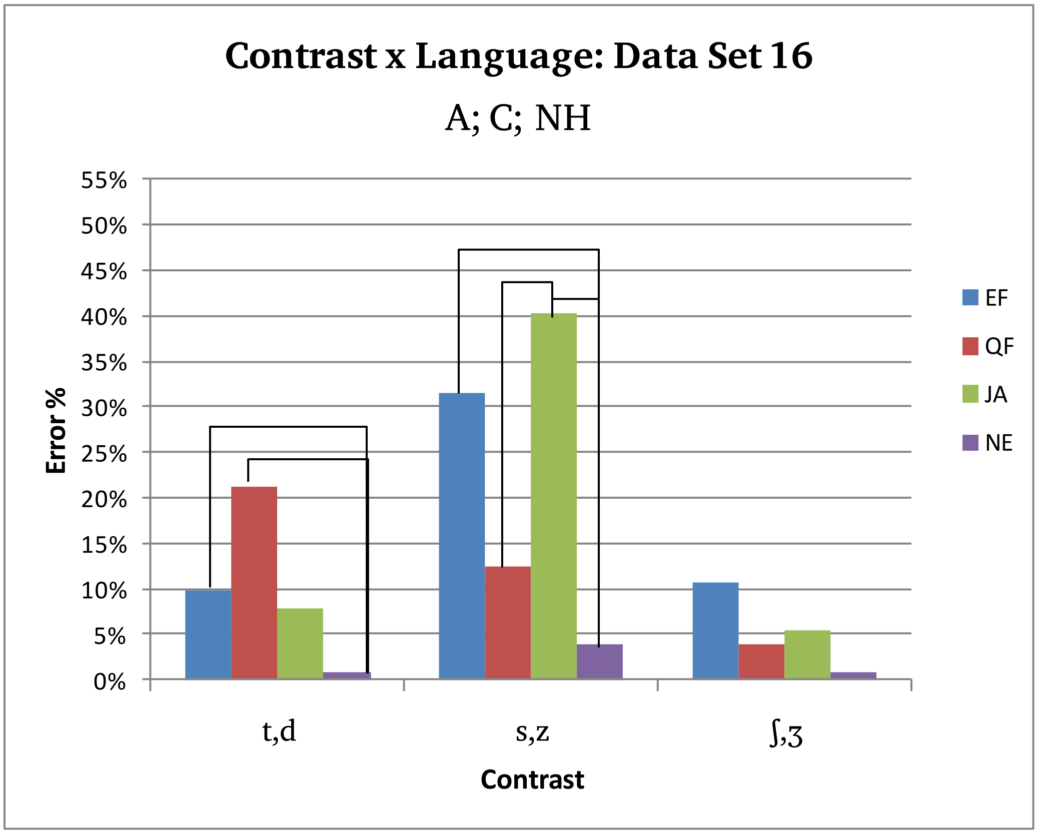

Figure 5.22. Data Set 16. Contrast x Position x Language interaction. Significant differences are indicated with a connector line.

5.7.3.3. Summary of Results on Contrast

The results presented in this section have addressed the ability of listeners to perceive contrasts involving interdental fricatives. In this summary, I present the main findings which have emerged.

First, regarding the contrast /s-θ/, we have seen that Québec French listeners (especially the Beginner/Intermediate listeners) perform significantly better than European French and Japanese listeners. The Russian group patterns with EF and JA, although this does not quite reach significance. All test groups exhibited more errors than the Native English control group, as expected.

This result does not correspond to the predictions for Québec French which rest upon the assumption that non-contrastive features are unavailable at the lexical level of representation. The feature which distinguishes this pair is STRIDENT/MELLOW, which has been argued to be unnecessary in the contrastive phonology of Québec French. On the other hand, this result *is* in accord with the phonetic predictions set out at the beginning of Chapter 4, in which the feature STRIDENT is argued to be particularly salient in QF.

Conversely, for the /t-θ/ contrast, the Québec French group performs more poorly than the European French, Japanese, and Russian groups. Again, the source of this difference largely stems from the Beginner/Intermediate listeners, although the interaction with Proficiency did not reach significance.

This finding is not expected if only contrastive features are available in lexical representations; however, it is consistent with the phonetic predictions set out in Chapter 4 for EF, QF, and JA. With respect to Russian, the phonetic predictions were that [s̪] and [t̪] would be equal contenders as substitutes for [θ]; but the results here show [s̪] to be more likely; in fact, the RU group displayed a very similar pattern to the EF group on both /s-θ/ and /t-θ/.

Other findings that emerge from the results in this section concern the /f-θ/ and /ʃ-θ/ contrasts. For /f-θ/, all test groups have more difficulty than the control group, especially before a high front vowel or in coda position. This is consistent with the phonetic predictions, where /f-θ/ was figured to be the second choice for all learner languages, but it is not consistent with the phonemic predictions, in which good performance was expected for everyone based on the distinctive features available to all languages tested, i.e. LABIAL and CORONAL. Recall from Chapter 4, however, that a muting effect was introduced to explain the poor performance on [f-θ] in the AXB-2 study, viz. MELLOW FRICATIVE MUTES MAJOR PLACE. This phonetic muting effect does not appear to perturb the Native English group, for whom the /f-θ/ contrast is well-established. This may be an indication that second language learners are more likely to be influenced by phonetic factors even in a lexical task.

For /ʃ-θ/, while registering the least amount of errors overall, we have nonetheless seen that EF, QF, and JA experience more difficulty with this contrast than the NE control group. The Russian listeners pattern with the other learner groups, but this trend falls short of significance.

The data presented in this section have shown that speakers of QF are quite good at discriminating STRIDENT versus MELLOW as compared to EF and JA. On the other hand, they are quite poor at distinguishing the STOP versus CONTINUANT contrast in /t-θ/ as compared to the other language groups.

We also see that position plays a role in the discrimination of the LABIAL-CORONAL contrast for the fricatives /f-θ/, and this is seen across the test language groups. This suggests that formant transitions into the following vowel may be necessary for fricative discrimination for the L2 learners, in particular for these mellow fricatives.

In general, the data on /s-θ/ and /t-θ/ in this section support the hypothesis that production errors are based on perception errors.

5.7.4. Statistical Results for Voicing

Data Set 4: **(**EF/QF/JA; BI/A; C; Vcd/Vclss; /t d, s z, ʃ ʒ/; NH). Voicing was only compared in coda position, and no main effects were found. A Contrast x Voicing x Language interaction emerges in this analysis, Anova $F(4,100), p<.001$. Two-tailed Mann-Whitney tests show significant differences emerge only in the Voiceless condition. Québec French has more errors for voiceless /t-θ/ than for voiced /d-ð/, and fewer errors for voiceless /s-θ/ compared to voiced /z-ð/. This scenario is reversed for the JA group, leading to significant differences between QF and JA. It seems that the contrasts that are problematic for these language groups are made even more ambiguous when voiceless, at least in coda position.

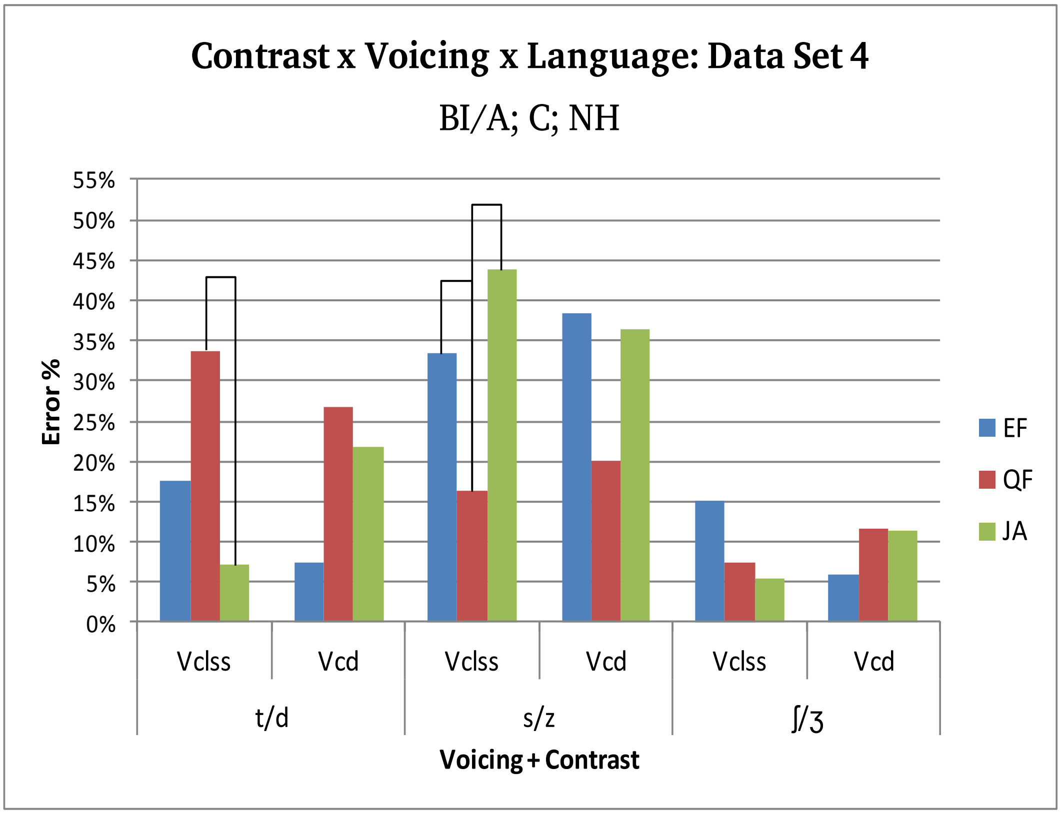

Figure 5.23. Data Set 4. Contrast x Voicing x Language interaction. Significant differences are indicated with a connector line.

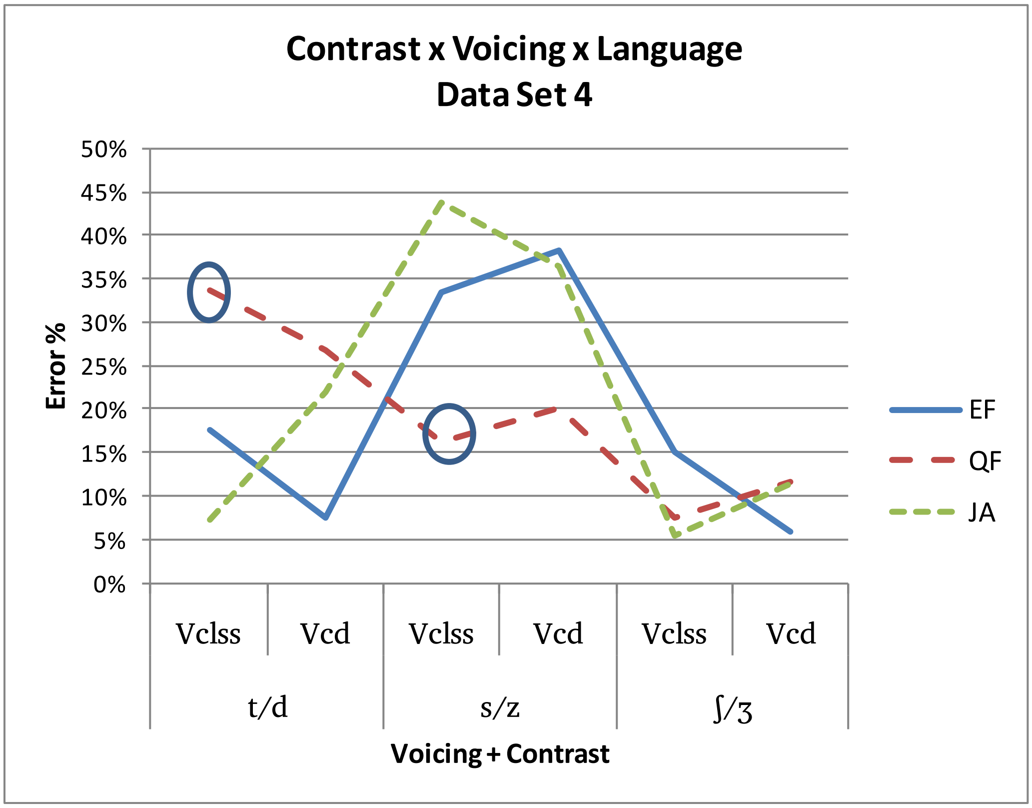

Figure 5.24. Data Set 4. Contrast x Voicing x Language interaction. The source of the interaction is circled.

5.7.5. Statistical Results for Language

This section gives statistical main effects for the Language factor. All interactions involving Language have been presented in previous sections.

Data Set 2: (EF/QF/JA; BI/A; O/C; Vclss; /f/; HF/NH). There was no main effect and no interactions for Language.

Data Set 6: **(**EF/QF/JA; BI/A; C; Vclss; /t/; HB/NH). This analysis of /t-θ/ in Coda gives a significant main effect for Language, Anova F(2,5),p<.001), Kruskal-Wallis (X^2(2),p<.001). Two-tailed Wilcoxon tests show that this stems from higher error rates for QF versus the other language groups.

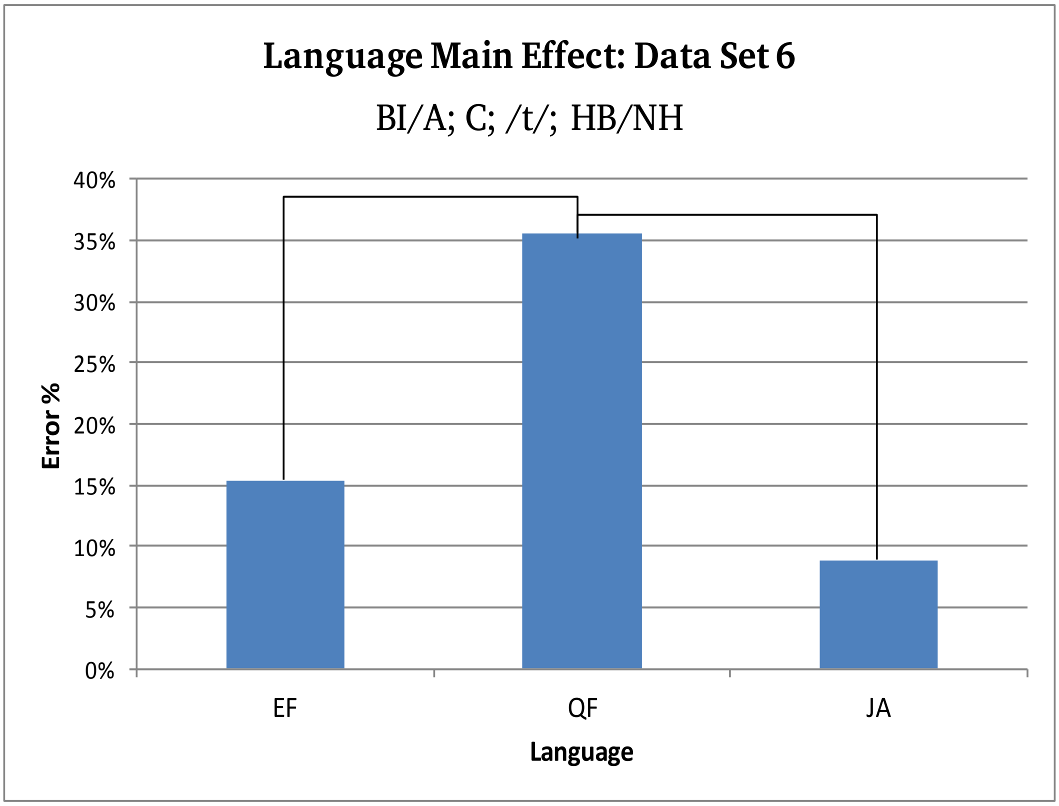

Figure 5.25. Data Set 6. Language main effect.

Significant differences are indicated with a connector line.

Data Set 11: (EF/QF/JA/RU/NE; BI; C; Vcd; /z/; HF/NH). This data set examines /z-ð/ in Coda, and also shows a main effect for Language, Anova $F(4,40), p<.001$; Kruskal-Wallis ($X^2(4), p=.001$). Post-hoc Scheffé and Mann-Whitney tests show that all the learner groups have significantly more errors than the Native English control group, as seen in Figure 5.26.

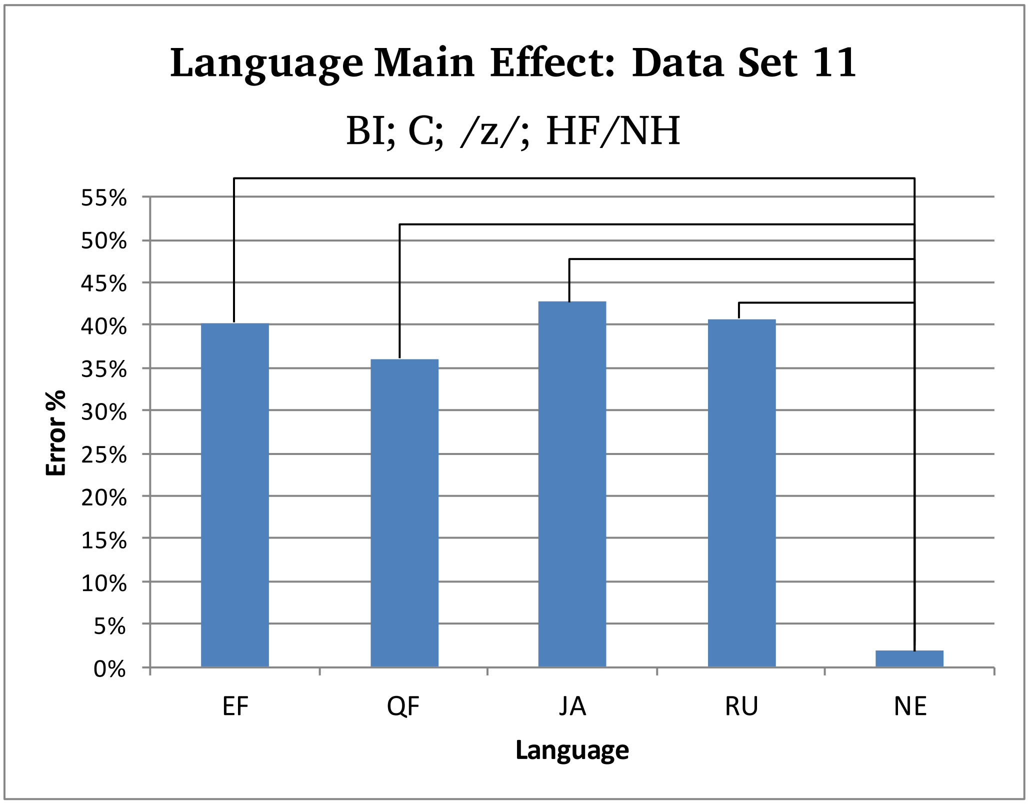

Figure 5.26. Data Set 11. Language main effect.
Significant differences are indicated with a connector line.

Data Set 12: (EF/QF/JA/RU/NE; BI; C; Vclss; /t/; HB/NH). As seen for both Proficiency levels in Data Set 6, here we observe that, for the BI group alone, there is also a main effect for Language, Anova F(4,40),p < .001; Kruskal-Wallis (X^2(4),p = .002). Scheffé post-hoc and Mann-Whitney tests show that QF has more errors than EF, JA, and NE.

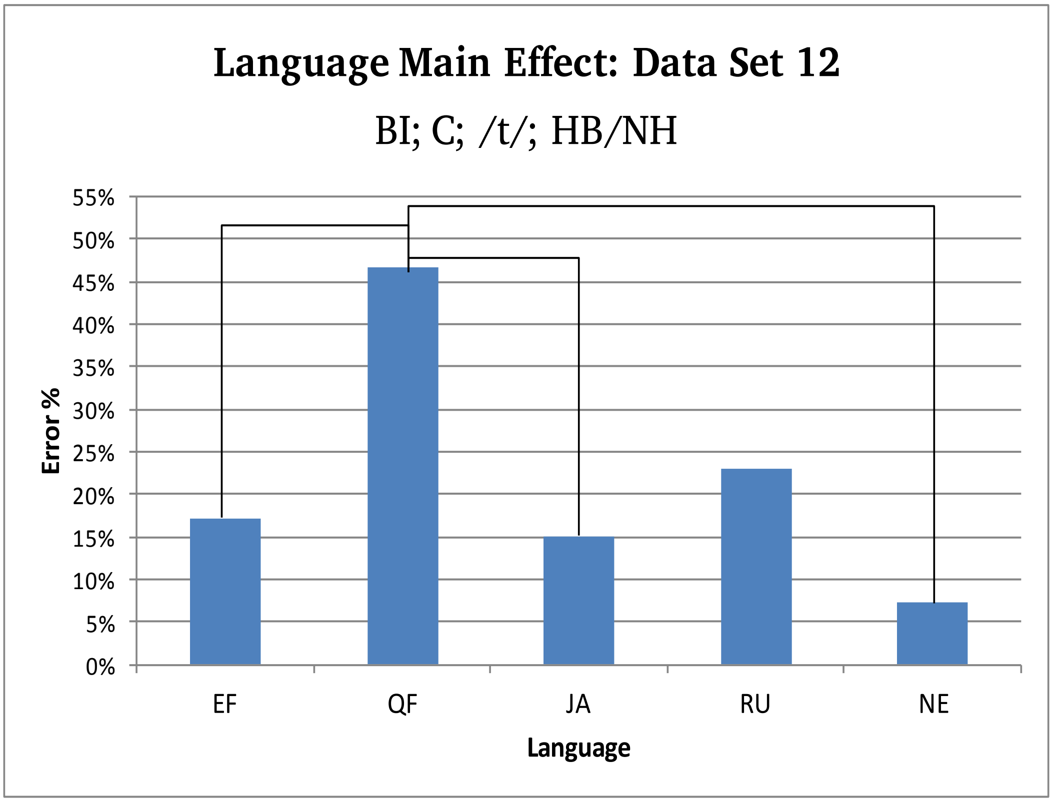

Figure 5.27. Data Set 12. Language main effect.
Significant differences are indicated with a connector line.

Data Set 17: (EF/QF/JA/NE; A; C; Vcd; /z/; HF/NH). As with the BI group in Data Set 11 (Figure 5.26), the Advanced group shows a main effect for language on this /z-ð/ contrast, Anova $F(3,33), p<.001$, Kruskal-Wallis ($X^2(3), p<.001$). Two-tailed Mann-Whitney tests show that EF and JA have more errors than NE.

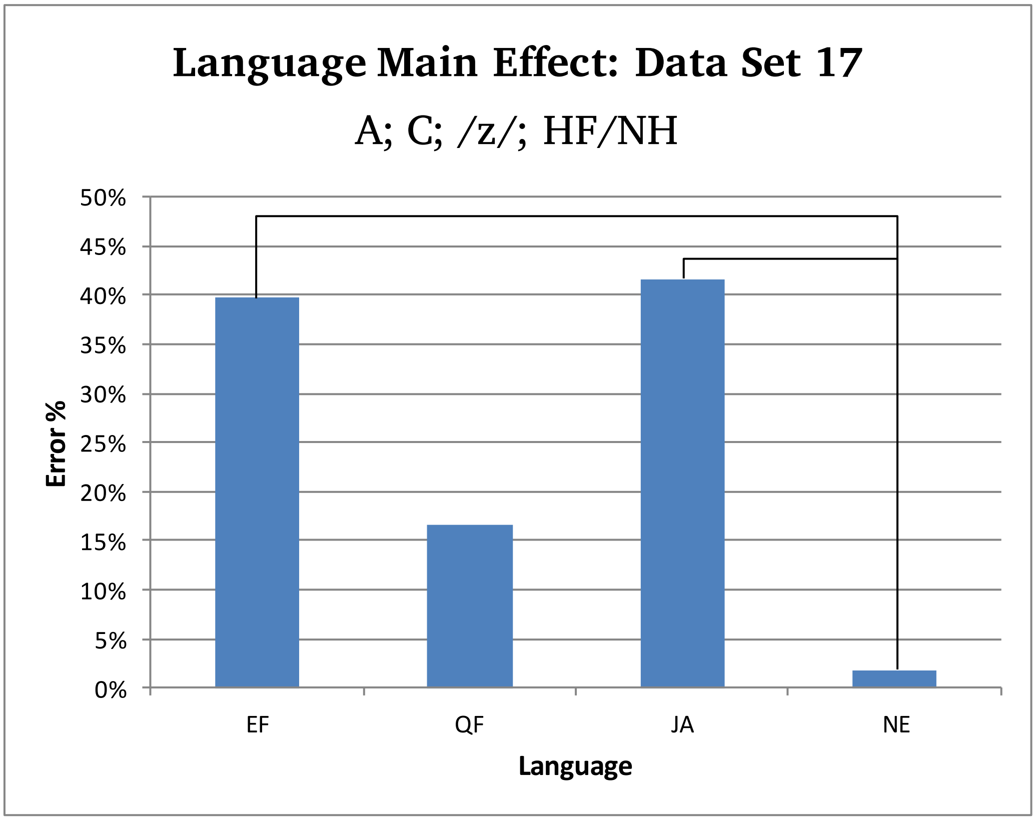

Figure 5.28. Data Set 17. Language main effect.
Significant differences are indicated with a connector line.

There were no other significant differences.

Data Set 18: (EF/QF/JA/NE; A; C; Vclss; /t/; HB/NH). The Anova for this data set shows a main effect for Language (Anova F(3,33),p=.005) with QF having more errors than JA and NE; however, the non-parametric Kruskal-Wallis test did not attain the .005 level set for significance (X^2(3),p=.01).

5.7.5.1. Summary of Results on Language

This section has focussed on the main effect of Language. The data sets which show such a main effect did not demonstrate any interactions involving Language. For all of these significant analyses, there could be no interaction with Contrast or Position, since they involve a single contrast in coda. Vowel quality was manipulated, but the lack of interaction with Vowel indicates that the type of vowel does not influence perceptibility in these cases.

The findings in this section again highlight the difficulty that the Québec French listeners have with /t-θ/ as compared to European French and Japanese listeners. As for /z-ð/, we have seen that all learner groups have more errors than the Native English control group; however, the QF group has the least errors amongst the learner languages, and this is especially true of the Advanced QF group.

The findings on /t-θ/ here are again consistent with the phonetic predictions for QF, but not with the phonemic predictions. On the other hand, the results for /z-ð/ lean somewhat more in the direction of the phonemic predictions in that QF listeners exhibit a fairly high error rate; nonetheless, the rate of errors remains lower than that for the /t-θ/ contrast. It seems that coda position increases the likelihood of a /z-ð/ merger.

5.7.6. Statistical Results for Position

There were no main effects for Position, and all interactions involving Position have been reported in previous sections (see Figures 5.6, 5.14-5.15, 5.17-5.21). However, I will give an overview of these interactions here as well. Firstly, most interactions with Position involved the /f-θ/ contrast. This contrast is better perceived in onset position than in coda. Furthermore, in onset, /f-θ/ is better perceived before a non-high vowel as compared to a high front vowel; while in coda, this is reversed.

A second contrast which interacts with position is /ʃ-θ/. For QF, and JA in particular, this contrast presents more difficulty in onset position. And for the Advanced EF group, it is more problematic in coda position.

Finally, for /s-θ/, for the Advanced learners, European French has slightly more difficulty in onset position; whereas, Québec French finds this contrast easier in onset.

5.8. Summary and Discussion of PicID

In this chapter, I have presented results from a Picture Identification task. This type of task elicits lexical processing; in order to complete it, one must have a mental representation of a word, and evoke its phonological representation in order to match an aural stimulus with a picture. Unlike the AXB task, in the PicID task, the risk of tapping an acoustic (non-linguistic) processing mode is avoided.

Two levels of proficiency were examined: Beginner/Intermediate versus Advanced. Only European French, Québec French, and Japanese groups were compared on proficiency. In general, and as expected, the Advanced learners in these language groups show improvement on all contrasts in all positions, in all vowel contexts, and in both voicing conditions tested. This is encouraging news for learners -- perceptual acuity with respect to the interdental fricatives of English can indeed be acquired.

Although the data from this study indicate that learners can acquire new phonological features or recombine features to form new phonemic categories, at the beginning of acquisition of a second language and throughout the acquisition process, learners do not necessarily possess all target features and phonemic categories. And this incongruence leads to misperception and hence foreign accent in production, as learners must associate a target form with a form that is available in their current interlanguage inventory, which at the outset of acquisition, is deemed to be equivalent to their L1 inventory.

At the beginning of this chapter, it was predicted that if only contrastive features are available at the phonological level of representation, then European French, Québec French, and Japanese listeners should have associated the target interdental fricative /θ/ with their native coronal sibilant /s/, and Russian listeners should have exhibited variation between /f, t, s/ substitutes. Results from the PicID study support these predictions for European French and Japanese, but not for Québec French, who display relatively low error rates on the /s-θ/ contrast. Russian tends to pattern with European French. This indicates that Québec French has a feature in its phonology that enables these listeners to make this distinction.

As discussed at the beginning of this chapter, I consider the feature in question to be STRIDENT. It was argued in Chapters 2 and 3 that the /s/ of EF, QF, and RU is specified for STRIDENT in the phonetic component; whereas, for JA it is not. It was further argued that only in QF is STRIDENT enhanced by virtue of its being produced at the alveolar place of articulation, in comparison to EF and RU, where [s̪] has a dental place of articulation. Thus, for QF, stridency should be more salient than it is for EF and RU, and especially more than JA.

The results from the current chapter are consistent with this analysis, but situate it in the phonological component. We have seen that JA has the most difficulty with /s-θ/, followed by EF and RU, then by QF and NE. However, despite the tendency for the elevated error rates for JA which make this group stand apart from EF and RU, this did not quite reach statistical significance. In other words, statistical analyses group JA, EF, and RU together on the /s-θ/ contrast. Thus, the data from this study suggest two possibilities: either the /s/ of JA, EF, and RU is not specified for STRIDENT in lexical representations or in EF and RU it is specified with a muted, weak version of STRIDENT, and JA /s/ is not specified at all for STRIDENT. The latter scenario would imply that enhancement and muting has an influence on L2 learners in the phonological component.[77]

In contrast to Russian, where the feature STRIDENT functions to make a phonemic distinction between the affricate /ts/ and the stop /t/, I have argued that STRIDENT does not serve a contrastive function in EF, QF, or JA. Why then would QF have STRIDENT in its lexical representations, while EF would not (or would have only a weak STRIDENT feature)? I contend that it is due to the phonetic characteristics of their anterior sibilants. As I have argued in previous chapters, Québec French alveolar [s] is more strident than European French and Russian dental [s̪].

Another explanation that has been offered for the presence of STRIDENT in QF and its absence in EF rests on the process of affrication in Québec French. Recall that in QF, /t,d/ → [ts,dz] before high front vowels. Kim (2001) argues that the reason for this is that the turbulence created in the long transition between the occlusive and following high vowel is interpreted as [+strident] in the post-lexical phonology in QF. Although a reasonable argument, this cannot explain the discrepancy between QF and JA. Japanese also has affrication of stops before high vowels, and Kim considers that JA also inserts [+strident] post-lexically. Thus this late insertion of [+strident] does not explain the results from the current study. It must be that STRIDENT is available earlier to QF, and later or not at all to

[77] This would differ from Keyser & Stevens (2006) since they state that enhancement occurs in the phonetic component. Note however that theirs is a production model.

Japanese. In fact, the results from AXB-1 and AXB-2, which tapped a phonetic level of representation, suggest that STRIDENT is not available at all to Japanese. So either the affrication process in Japanese is different from QF, and does not involve the insertion of STRIDENT, or JA somehow cannot make use of STRIDENT in their assessment of target interdental fricatives.

It could be argued that affrication provides sufficient evidence for QF children to postulate different underlying representations for their consonant inventory compared to EF children. Since QF only has minimal distinctions between stops and continuants at the labial and coronal places of articulation, the existence of affrication may induce L1 learners to use MELLOW-STRIDENT to make the STOP-CONTINUANT contrast, i.e. to distinguish /p,b/ from /f,v/ and /t,d/ from /s,z/. Thus, one could say that QF is a "MELLOW-STRIDENT" language; whereas, EF is a "STOP-CONTINUANT" language. I would like to argue against this.

I contend that the source of MELLOW-STRIDENT in QF is not because these features are used to distinguish stops from continuants in QF. Both EF and QF have a rule of vowel lengthening before the consonants [r v z ʒ] (Delattre 1951, Côté 2005). This is a natural class of voiced, continuant, non-lateral consonants. We must appeal to the feature CONTINUANT here; neither MELLOW nor STRIDENT function to delimit this class. Since the feature CONTINUANT is needed elsewhere in the language, then STRIDENT is redundant. Yet QF seems to have access to STRIDENT, even at the lexical level, as shown by the Picture Identification task. As previously discussed, I argue that this evidence comes from the phonetic (non-contrastive) specification of their anterior coronal fricatives.

Let us return now to the other results from the PicID study. Looking at the /t-θ/ contrast, we have seen that Québec French has more difficulty than EF, JA, RU, and of course Native English. These results fit more with the predictions outlined for a phonetic scenario in Chapters 3 and 4 than with what was anticipated in a phonological analysis.

Again for this /t-θ/ pair, we have seen that EF and RU pattern together, suggesting that their representations for /t/ (and /θ/) are similar. Note that error rates on /t-θ/ for Beginner/Intermediate EF, JA, and RU groups are still statistically higher than the NE group; although they fall considerably for the Advanced JA group, but actually rise for the Advanced EF and RU groups. In other words, while consistently higher in the /s-θ/ condition, error rates between /s-θ/ and /t-θ/ are more evenly distributed for the EF and RU groups.

In general, these findings are consistent with the idea that perception determines production, especially for EF, QF, and JA. As discussed earlier, data on production indicates that QF speakers tend to substitute stops for the target interdental fricatives; whereas, European French and JA tend to substitute their anterior coronal fricative. Results from the PicID task follow this direction. For Russian, the tendency is the same as for EF, with /s-θ/ tending more towards perceptual merger.

However, the PicID study has also revealed that findings must be nuanced somewhat. Position and vowel quality influence perceptibility. We have seen that there seems to be a tendency for all non-native language groups to confuse /θ/ with /f/, particularly in coda position. In addition, vowel quality influences perception of this contrast, and this is dependent upon position. For /f-θ/, discrimination is facilitated before a non-high vowel, but inhibited after a non-high vowel. For /t-θ/ though, perceptibility is diminished before a non-high vowel.

It is suggested that since high front vowels involve a spread lip configuration, this would eliminate the front cavity present in the articulation of [f] that is present in the context of other vowels. The lack of cavity may cause [f] to resemble [θ] much more closely, especially since both sounds are of low intensity. Poor perception in coda is expected, given the markedness of this position. This is also consistent with the suggestion by Solé (2003) that the amplitude on fricatives is reduced in coda position. This may be consequential for the already precarious /f-θ/ contrast, and the open oral cavity of a preceding non-high vowel may further exacerbate the problem. The length of the preceding vowel may also play a role here, with a longer vowel encoding more information as to the identity of the final fricative.

The high error rates for /f-θ/ do not correspond with the production literature, with the possible exception of European French. However, as discussed in Chapter 2, the labiodental fricative may be avoided in production due to visual cues on the interdental fricative.

In general, these trends do not extend to the NE group. This is mostly because the control group is performing at ceiling (although their error rates did edge up to 9% on /f-θ/ before a HF vowel). This suggests that native controls likely have firm phonological categories; whereas, learners' categories are less well established; and, therefore, they may rely more on various non-contrastive phonetic cues.

That learners have less solid phonological categories is seen even in the results from the contrast that was most easily discriminated: /ʃ-θ/. Error rates were lowest on this pair, yet in general the learner groups still had significantly more errors than the control group.

In addition to the influence of position and vowel on perceptibility for these L2 learners, there was some evidence that voicing also plays a role, at least in coda position. For QF and JA, the voiceless condition seems to aggravate an already problematic contrast: for QF, /t-θ/ is worse than /d-ð/ in coda, and for JA, /s-θ/ is worse than /z-ð/. This may be related to vowel length: in English vowels are longer before voiced consonants (Dauer 1993). As mentioned above, a longer vowel may encode more information as to the identity of the following consonant. Recall that in the AXB-1 study reported in Chapter 3, [z̪-ð] in onset position of nonce words as also better perceived than voiceless [s̪-θ]. It was suggested there that weak formant transitions on the voiced fricative might have aided perception.

The results from the PicID task are generally consistent with those found in the AXB tasks. In both types of task, there is a tendency towards higher errors on /s-θ/ for EF, JA, and RU compared to QF, and an opposite tendency for /t-θ/. However, in the AXB-2 task, we observed that error rates on [s-θ], and especially on [s̪-θ], were generally more elevated for all language groups, including the native English control group. As previously discussed in Chapter 4, this may be due to the heavy processing load imposed by the AXB-2 task.

In contrast to the AXB-1 study, where each triad consisted of isolated words, in the AXB-2 experiment, listeners heard three short carrier phrases which incorporated the target words. Within these triads, the carrier phrase differed, as did the talker. With such variation occurring, the ability to focus on the target segments is particularly taxing.

I suggest that such demands may occasionally cause the listeners to switch into an acoustic processing mode, in which the phonetic/phonological systems are bypassed (Gerrits 2001), and the target sounds are directly compared rather than being associated to a phonetic or phonological representation, as is the case in the Perceptual Assimilation Model. This implies that acoustically, [s,s̪] and [θ] are more similar than [t,t̪] and [θ]. The former are acoustically close, but they are not necessarily phonetically or phonologically close because of the particular configuration of linguistic features which are subject to enhancement or muting effects.

However, perhaps an even more acoustically fragile contrast is [f-θ]. We have seen in both the AXB tasks and the PicID task that this contrast presents problems for the test groups. Interestingly, this contrast is also the one that is most subject to influences from neighbouring vowels in the PicID task. This influence does not appear to be phonological in nature, but rather phonetic; in other words, it is not due to feature spreading, but rather to gradient degrees of stridency and/or differences in formant transitions on the vowel.

Thus, this may be further evidence that phonetic information is encoded in the phonological representations. I would like to suggest that the phonological component is not necessarily free of redundancy. Some non-contrastive features may be specified, such as STRIDENT, in the case of Québec French. As well, I speculate that for language learners, new phonological representations may initially encode non-contrastive phonetic information. As development proceeds, some of this phonetic information may be pruned from phonological representations, and the result is an increasingly solid phonological representation, impervious to vowel influence, eventually approaching the stability exhibited by the native English control group in this Picture Identification study.

In general, we see that the perceptual patterns found in this experiment mirror those found in the production literature. This supports a perceptual basis for cross-linguistic production patterns. The next chapter will investigate the production aspect of interdental substitution. It examines the phonetic quality of /s/ in European French, Québec French, Japanese, and Russian, to determine whether there are differences in stridency amongst these languages. Also, an investigation of the production of English words containing an interdental fricative is reported. Results from individuals who participated in the perception experiments are compared on their production to further test the hypothesis that perception underlies production.

CHAPTER 6: PRODUCTION

6.1. Introduction

In this chapter, I report on two production studies. Both investigate data from native speakers of European French, Québec French, Japanese, and Russian. The first study is a word production task, and the second a spectrographic analysis of the native /s/ produced by these speakers.

The purpose of the Word Production task is two-fold. First, it will serve to confirm or disconfirm other production reports. Second, and more importantly, it will provide a direct comparison of perceptual substitutes and production substitutes. Thus, its aim is to test whether there is a correlation between perception and production. A positive correlation would be consistent with the main hypothesis of this thesis that perception underlies production.

The spectrographic analysis tests the hypothesis that the quality of /s/ in the native language of the learner is a determining factor in the instantiation of interdental substitute observed for that language. Specifically, it tests the degree of stridency of the /s/ in European French, Québec French, Japanese, and Russian.

The next section describes the Word Production task.

6.2. Word Production Task

6.2.1. Participants

The participants retained for this study were those who also participated in either the AXB-2 task and/or the PicID task discussed in Chapters 4 and 5 respectively.

Only participants classified as Beginner/Intermediate (BI) are reported in this study. This was done for two reasons: First, only the BI group includes Russian participants. Recall that there was no Advanced Russian group in either the AXB tasks or Picture Identification task. Secondly, it was thought that the BI group, rather than the Advanced group, would give a clearer picture of the types of production errors made, as the Advanced group is more likely to have mastered production of [θ].

6.2.2. Procedure

The Word Production task was designed in such a way as to divert the participants from the purpose of the study, viz. their production of the English "th" sounds. They were told that they would be presented with two words on the computer screen. One of the words would be in a larger font than the other. Participants were instructed to pronounce the word that was in a larger font. Both words contained the target [θ] sound in the same position, viz. onset or coda; thus, a sample was collected irrespective of which word was pronounced.

The experiment was constructed using PsyScope (Cohen et al. 1993) on a MacIntosh Powerbook. Words stayed on the screen for 2 seconds, after which time, the next pair of words would appear. This put some pressure on participants to respond quickly, thereby discouraging recourse to metalinguistic knowledge. Responses were recorded in analogue and later digitized with Kay Elemetrics Computer Speech Lab at 22 kHz.

There were 55 pairs of words, of which 22 were test items, and the rest fillers. These were presented in random order. Test items incorporated either [θ] or [ð] in simple onset, simple coda, or intervocalically, but only the voiceless targets were analyzed for this study, as they provided the best comparison with the perception tasks. The words were presented in orthographic form since several of them are not easily picturable. Only voiceless items in simple onset and coda were retained for analysis. Again, this was done in order to provide a good comparison with the perception data, which did not include intervocalic position, and did not have many voiced types in coda position.

The onset and coda test stimuli consisted of 12 pairs -- 6 onset, 6 coda. Each pair was presented twice, in opposite order; for example, once as "thirst-think", then again as "think-thirst".[78] A list of the stimuli is in Appendix F.

Participants' responses were narrowly transcribed by the author, and these transcriptions were verified by another trained linguist, who was instructed to pay attention to fine phonetic detail. In cases of disagreement, the author listened again to the recording, and made the final decision. In the few cases where there was disparity, it was usually with respect to the fine phonetic detail. Non-target responses were coded into four categories: /f s t ʃ/. Both voiced and voiceless renditions were included in these categories.[79] Category /f/ included instances where the participant misproduced [θ] as either [f] or [ɸ]; category /s/ included both dental and

[78] It sometimes happened that a subject would fail to pronounce a stimulus item or pronounce the same item twice. In the latter case, both instances were analyzed.

[79] There were only three instances where the participant produced a voiced segment in place of the target.

alveolar tokens; /ʃ/ included both [ʃ] and [ɕ]; and [t] included both dental and alveolar renditions. All other non-target tokens were excluded from analysis including affricates, glottals ([ʔ] and [h]), and sonorants. There were very few of these other renditions; for the most part, if the participant did not produce a sound that fit into /f s t ʃ/, it was instead the target interdental fricative.

6.2.3. Results

Results from the Word Production task were calculated and compared to the results from the AXB-2 perception task in the first analysis and the PicID task in the second analysis. Participants were matched for comparison in both analyses. Some participants did both the AXB-2 and PicID tasks, while others did only one; therefore the analyses for Word Production vs. AXB-2 and Word Production vs. PicID are not directly comparable. However, there is no reason to suspect that the participants are not representative of their language group and proficiency level.

For these comparisons, in the Perception data reported below, the factor of Vowel was collapsed, as was minor place of articulation -- [s] with [s̪] and [t] with [t̪]. This was done in order to coincide with the Production data, for which Vowel was also collapsed due to insufficient tokens in each vowel category, and for which dental and alveolar renditions of /s/ and /t/ were grouped.

Correlations were calculated between the percentage of perceptual errors involving /f s t/ for real words alone in the AXB-2 task, for both Long and Short Interstimulus Interval (ISI) conditions, and for /f s t ʃ/ in the PicID task, for Onset and Coda conditions. Any /ʃ/ production errors were excluded from comparison with the AXB-2 task, since this sound was not tested in that perception task. The Kendall's tau_b correlation coefficient was selected instead of the Pearson correlation due to the small sample size and since normality of distribution cannot be assumed for these data.

Figure 6.1 shows that for the AXB-2 task, Long ISI, there is a significant positive correlation between perceptual errors and production errors involving /t/, $\tau = 0.495, N = 27, p < .01$, but not for the other types of substitutions, for all language groups. This means that perceptual errors where [θ] was misperceived as /t/ were proportionately the same as production errors where [θ] was misproduced as /t/. No such correlation was found for /f/ and /s/.

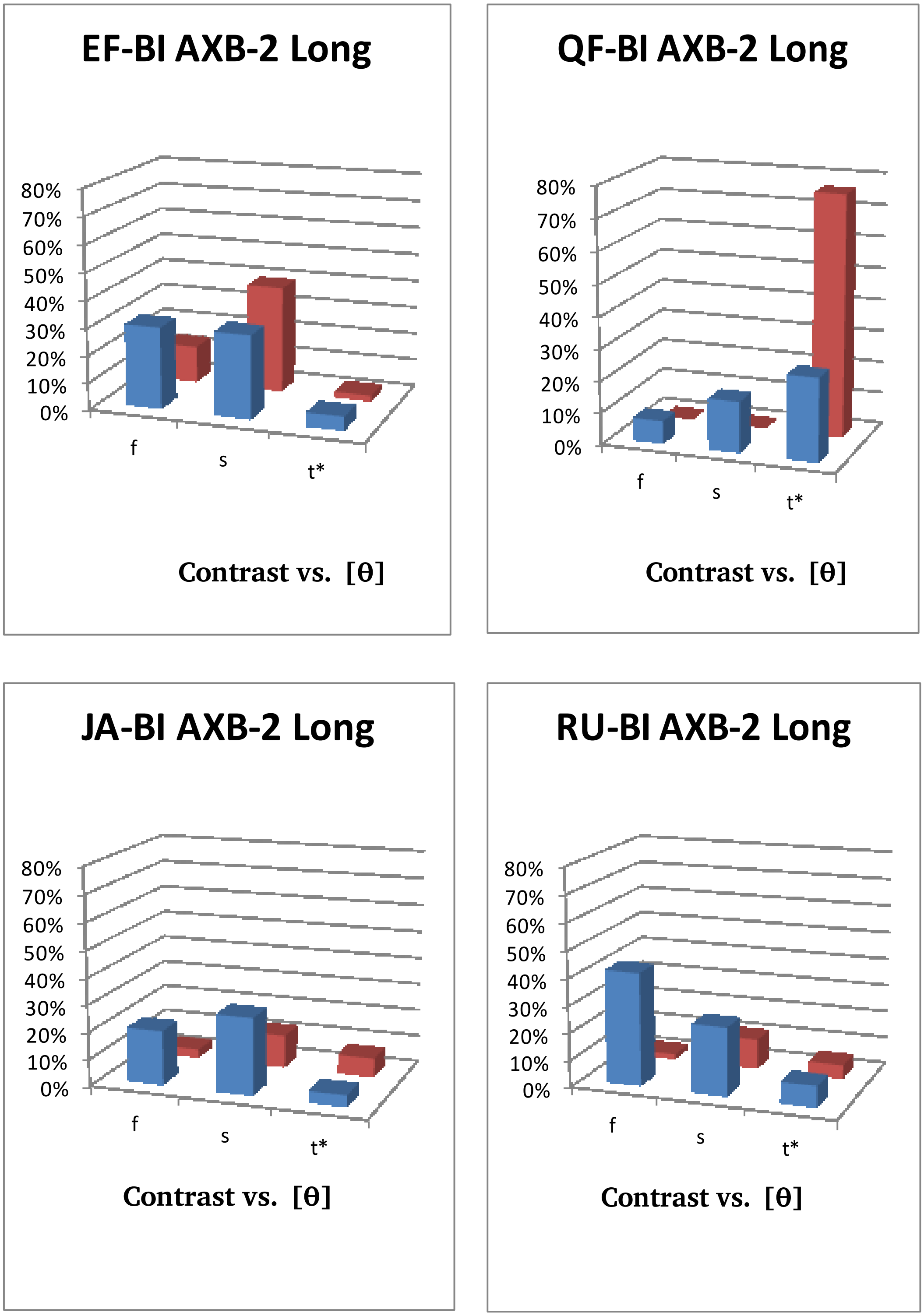

Figure 6.1. AXB-2 Long-ISI Perception Task vs. Word Production Task. Perception error %: front row (blue); Production error %: back row (red). The asterisk indicates significance.

Table 6.1 below summarizes the error rates for the AXB-2 Long-ISI Perception task versus the Word Production task. We see that for EF, the high percentage of production substitutes involving /s/ corresponds to a comparably high rate of perceptual errors implicating /s/. However, there is an equally high rate of /f/ errors in perception, but in production, the percentage of /f/ errors is only a third of those as /s/. For QF, we see a correspondence between perception and production, in that /t/ errors are highest; however, while we observe some instances of /f/ and /s/ errors in perception, in production there were no /f/ and /s/ errors whatsoever. Japanese shows that the most errors in perception and production involve /s/. In perception, /f/ is next, and the fewest perceptual errors are involve /t/; while in production, /t/ errors are second highest, followed by /f/. For Russian, we observe the most perceptual errors on /f/, followed by /s/, then /t/; but in production, the most errors involve /s/, followed by /t/, with few /f/ errors.

	Perception Errors (AXB-2 Long)				**Production Errors**			
	f	**s**	**t**	**Order**	**f**	**s**	**t**	**Order**
EF	30%	31%	5%	s>f>t	13%	39%	2%	s>f>t
QF	7%	16%	26%	t>s>f	0%	0%	76%	t> f,s
JA	20%	28%	4%	s>f>t	3%	12%	7%	s>t>f
RU	42%	25%	8%	f>s>t	2%	11%	5%	s>t>f

Table 6.1. Comparison of AXB-2 Long-ISI Perception task and Word Production task. Error rate percentages for Beginner/Intermediate group. The order of errors is based on raw scores, and is ranked from greatest to least for both tasks.

For the AXB-2 task, Short ISI, there is a correlation for all language groups between perception and production errors for /s/ and /t/, but not for /f/. For /s/, $\tau = .35(24), p < .05$; for /t/, $\tau = .375(24), p < .05$.[80]

[80] The correlations between perception and production for all tasks are relatively weak, with the AXB-2 Short ISI showing the weakest correlation. This may be partly due to the small sample size.

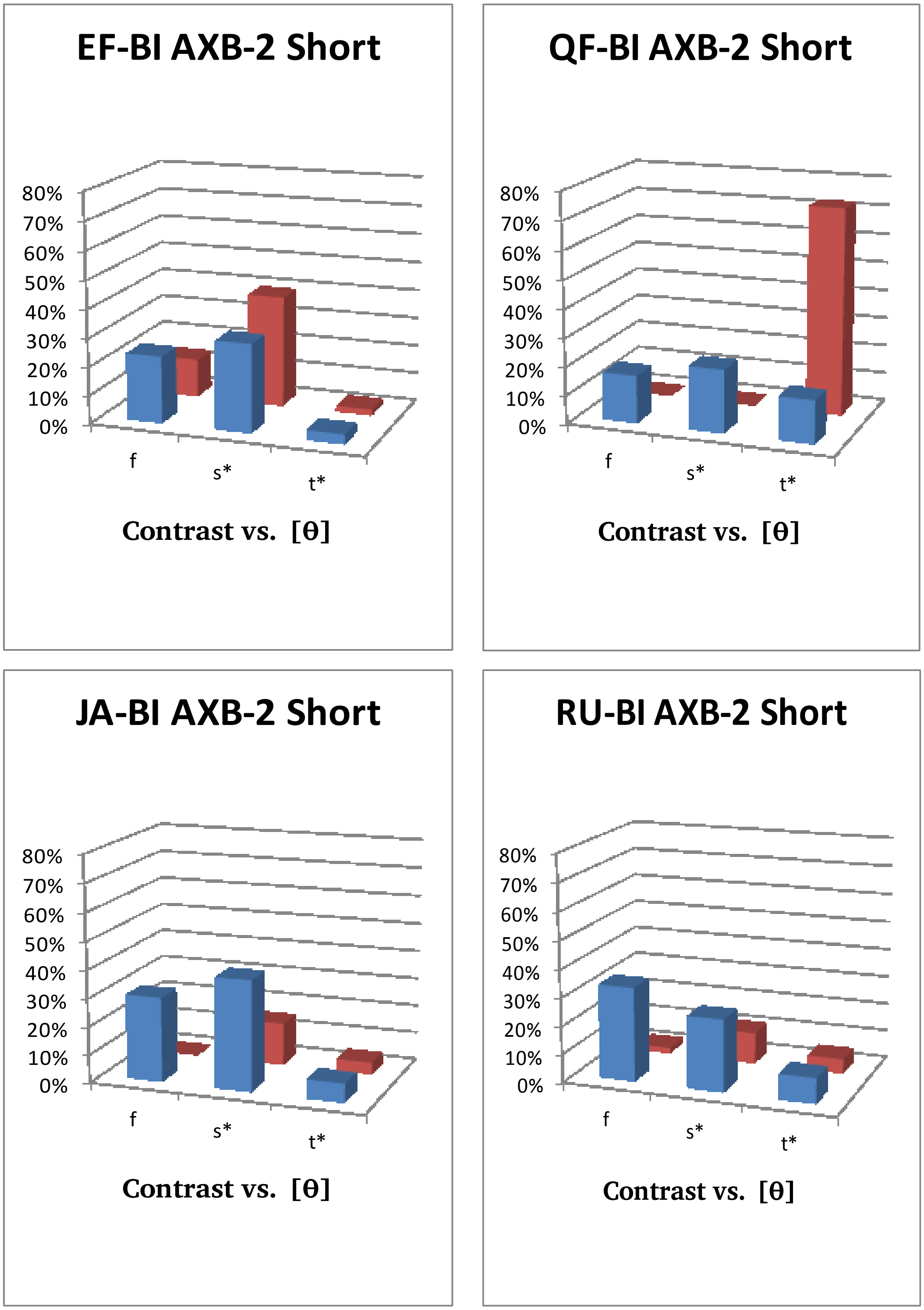

Figure 6.2. AXB-2 Short-ISI Perception Task vs. Word Production Task. Perception errors %: front row (blue); Production error %: back row (red). The asterisk indicates significance.

Table 6.2 summarizes the error rates for the AXB-2 Short-ISI Perception task versus the Word Production task. We see again that for EF, the order of perceptual errors corresponds to the order of production errors, with more errors involving /s/. For the Short-ISI, QF shows highest perceptual errors on /s/, followed by /f/, then /t/. This does not correspond well with their production errors where /t/ is the only type of error observed. Japanese again shows /s/ to be the most likely error in perception and production. In perception, this is followed by /f/, then /t/; in production, /t/ errors are next highest, with no instances of /f/. For Russian, again we observe the most perceptual errors on /f/, followed by /s/, then /t/; but in production, the most errors involve /s/, followed by /t/, with no /f/ errors at all.

	Perception Errors (AXB-2 Short)				**Production Errors**			
	f	**s**	**t**	**Order**	**f**	**s**	**t**	**Order**
EF	23%	31%	3%	s>f>t	13%	39%	2%	s>f>t
QF	17%	22%	15%	s>f>t	0%	0%	76%	t>f,s
JA	30%	39%	7%	s>f>t	0%	15%	4%	s>t>f
RU	33%	26%	9%	f>s>t	2%	11%	5%	s>t>f

Table 6.2. Comparison of AXB-2 Short-ISI Perception task and Word Production task. Error rate percentages for Beginner/Intermediate group. The order of errors is based on raw scores, and is ranked from greatest to least for both tasks.

For the PicID task, Onset position, only the relationship between perception and production errors for /s/ was significant, $\tau = .438(27), p < 01$.

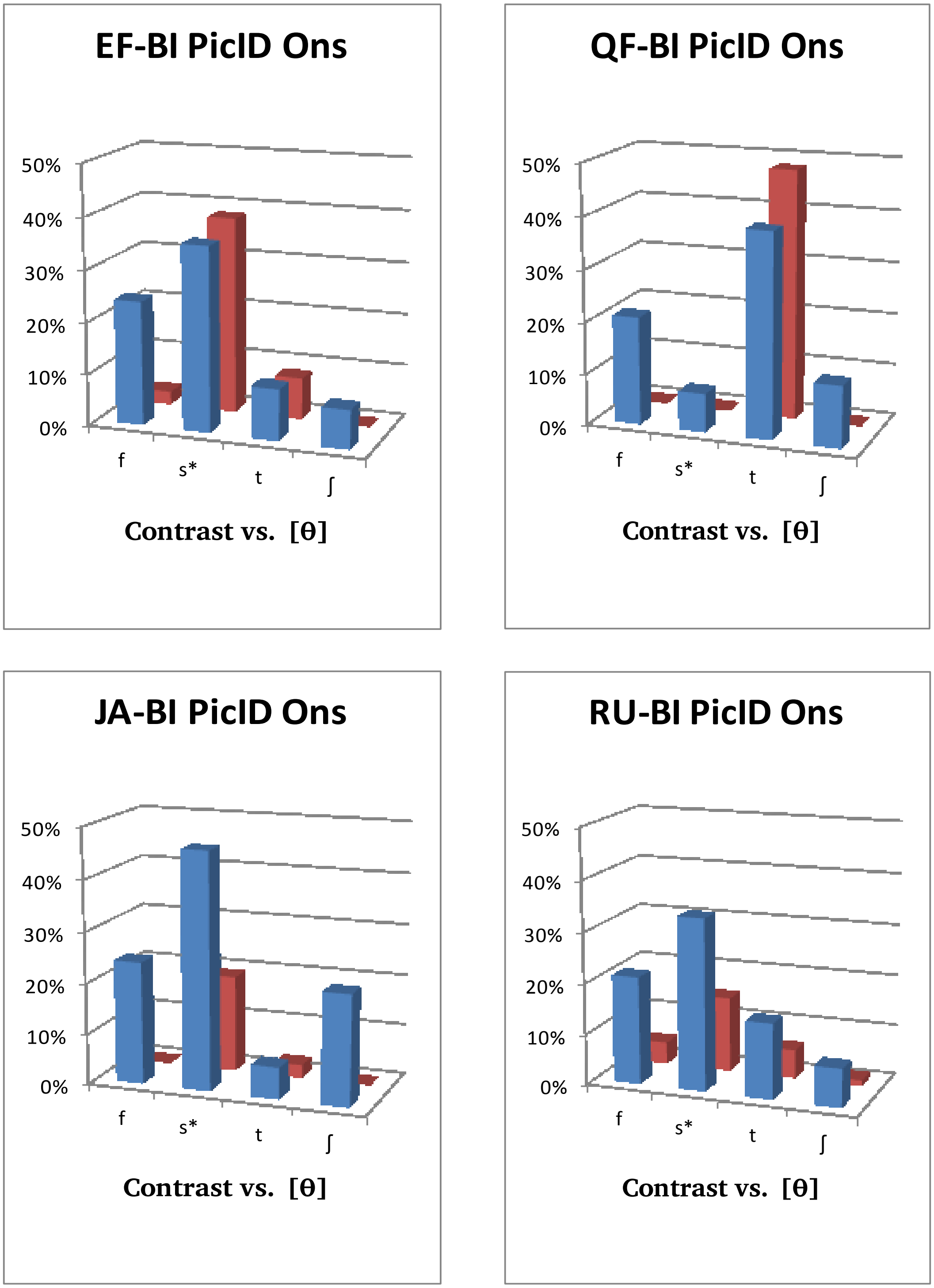

Figure 6.3. PicID Onset Perception Task vs. Word Production Task. Perception errors: front row (blue); Production errors: back row (red). The asterisk indicates significance.

The table below summarizes the error rates for the PicID Onset Perception task versus the Word Production task. For EF, the order of perceptual errors roughly corresponds to the order of production errors, with most errors involving /s/, and the least errors implicating /ʃ/; errors on /t/ are largely the same in perception and production, but for /f/, we see a discrepancy, with quite high perceptual errors compared to low production errors. Québec French shows highest perceptual errors on /t/, followed by /f/, then /ʃ/, with /s/ lowest. In production, again we see that /t/ is the only type of error. Japanese registers the highest errors on /s/ in both perception and production; however, the similarities end there. In perception, /f/ and /ʃ/ are next highest, followed by /t/; but there are virtually no errors on these three types in production. Russian shows the same pattern as European French.

	Perception Errors (PicID Onset)					Production Errors				
	f	**s**	**t**	**ʃ**	**Order**	**f**	**s**	**t**	**ʃ**	**Order**
EF	24%	36%	10%	7%	s>f>t>ʃ	2%	38%	8%	0%	s>t>f>ʃ
QF	21%	7%	31%	12%	t>f>ʃ>s	0%	0%	48%	0%	t>f,s,ʃ
JA	24%	46%	6%	22%	s>f>ʃ>t	0%	19%	3%	0%	s>t>f,ʃ
RU	21%	34%	15%	8%	s>f>t>ʃ	4%	15%	6%	1%	s>t>f>ʃ

Table 6.3. Comparison of PicID Onset Perception task and Word Production task. Error rate percentages for Beginner/Intermediate group. The order of errors is based on raw scores, and is ranked from greatest to least for both tasks.

For PicID, Coda position, only /t/ showed a significant correlation between perception and production, $\tau=.478(27), p<.05$.

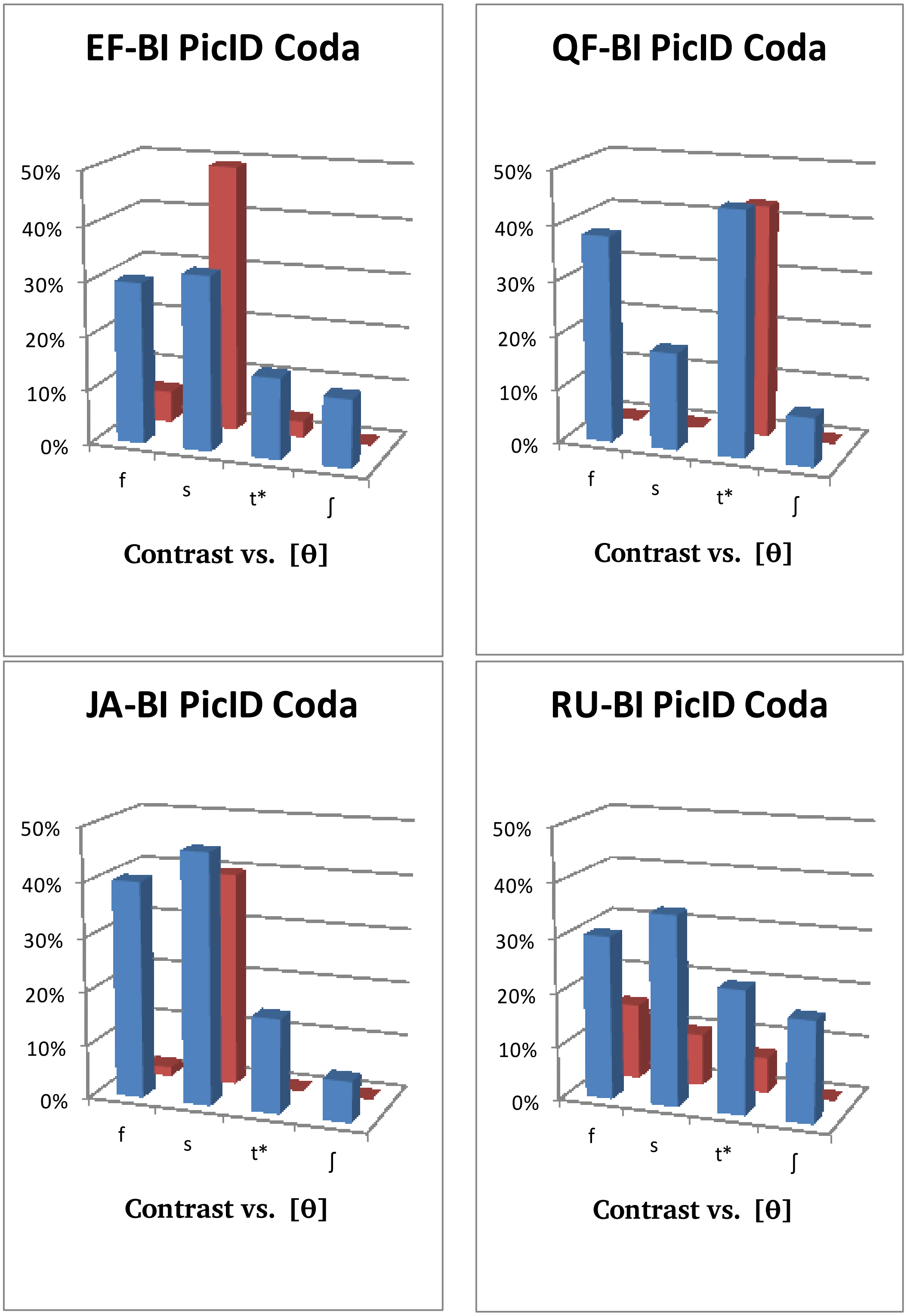

Figure 6.4. PicID Coda Perception Task vs. Word Production Task. Perception errors: front row (blue); Production errors: back row (red). The asterisk indicates significance.

Table 6.4 below summarizes the error rates for the PicID Coda Perception task versus the Word Production task. For EF, the order of perceptual errors corresponds to the order of production errors, with most errors involving /s/; however, production errors on /f t ʃ/ are much lower than perception errors. Québec French shows highest perceptual errors on /t/, followed by /f/, then /s/, with /ʃ/ lowest. In production, again we see that /t/ is the only error. Japanese has roughly the same order for perception and production errors, with the highest errors on /s/. However, like EF, the remaining types show virtually no production errors. Russian has relatively fewer errors on /t/ and /ʃ/, in both production and perception; but for /f/ and /s/, there are more errors, especially in perception.

	Perception Errors (PicID Coda)					Production Errors				
	f	**s**	**t**	**ʃ**	**Order**	**f**	**s**	**t**	**ʃ**	**Order**
EF	30%	32%	15%	13%	s>f>t>ʃ	6%	49%	3%	0%	s>f>t>ʃ
QF	38%	18%	45%	9%	t>f>s>ʃ	0%	0%	43%	0%	t>f,s,ʃ
JA	40%	46%	18%	8%	s>f>t>ʃ	2%	40%	0%	0%	s>f>t,ʃ
RU	30%	35%	23%	19%	s>f>t>ʃ	14%	10%	7%	0%	f>s>t>ʃ

Table 6.4. Comparison of PicID Coda Perception task and Word Production task. Error rate percentages for Beginner/Intermediate group. The order of errors is based on raw scores, and is ranked from greatest to least for both tasks.

6.2.4. Summary and Discussion of Production and Perception Tasks

This section has reported on a Word Production Task by Beginner/Intermediate learners of English from European French, Québec French, Japanese, and Russian native language backgrounds. The production data from these participants was compared with perception data from these same participants. Results do show a correlation between contrasts which cause problems in perception and errors in production. The L1 groups who most often confused [θ] with /s/ in perception (EF, JA, RU), were also those to most often produce /s/ in place of the interdental, and likewise for /t-θ/ (QF). These were the substitutes that were predicted by the Auditory Distance Model for phonetic processing.

However, there are also discrepancies between the two modalities. For /f-θ/ and /ʃ-θ/, we have observed more errors in perception than in production, and this disparity is especially noticeable for /f-θ/. Perceptual confusion of /f/ and /θ/ has been documented in several studies subsequent to the initial ground-breaking analysis by Miller & Nicely (1955). In Chapter 4, this was formally expressed as a muting effect, MELLOW FRICATIVE MUTES MAJOR PLACE.

Why perceptual errors involving /f/ and /ʃ/ are not carried over into production may be due to the visual information available in the articulation of these sounds. The labiodental gesture of [f] and lip rounding on [ʃ] provide visible cues which may be encoded in representations. Hardison (2005) found that identification of /f/ by Japanese and Korean ESL learners in a gating experiment was facilitated following training involving both auditory and visual modalities compared to auditory training alone. This and other studies on the role of visual input and its interaction with auditory input have revealed what is known as the McGurk effect (e.g. McGurk & MacDonald 1976), in which conflictual cues from each of these modalities can result in the perception of a sound intermediate between the input from each modality. For example, when presented with an auditory stimulus [ba] dubbed onto a visual stimulus [ga], native speakers of English often report the hearing [da]. As discussed in Hardison (2003), these findings suggest that perceptual representations integrate both auditory and visual information. Thus auditory input alone may not provide sufficient information for the L2 learner to distinguish between [f] and [θ] for example. But in production, the learner may draw upon stored visual information, and thus be able to either correctly produce the target sound or at least avoid [f] and [ʃ]. Visual cues are not as robust in distinguishing [s] and [t] from [θ], especially since the latter is often produced as dental rather than interdental in Canadian and British English (Ladefoged & Maddieson 1996), and thus visual cues may fail to have a facilitating effect in the production of these contrasts.

In the present study, we have seen that EF and RU display more production errors involving /f/ than do QF and JA. In Chapter 3, §3.4.4., it was argued that the reason QF rarely substitutes /f/ in production is due to these learners having more visual exposure to English articulation as compared to EF and RU; however, this explanation does not seem to extend to JA.

Another factor that could affect production is orthography. Orthography may contribute to correct pronunciation of the target (see e.g. Taft 2006). However, if this were the case, we might expect lower production error rates overall. Conversely, given that orthographic "th" is pronounced as [t] in French, spelling should have a negative effect for both European and Québec French. While this could conceivably account for some of the QF /t/ production substitutes, as we have seen in the graphs above, production errors involving /t/ are very low for the EF group. Thus it does not appear that orthography is a determining factor here.

Although the QF ESL learners compared in this chapter have more perception errors implicating /t/ across all tasks except AXB-2 Short-ISI, these are not the only contrasts that present perceptual confusion. There are also some problems with the other contrasts as well. However, their production errors are categorical: always /t/, and never /f s ʃ/. The other language groups produce more variants.

I suggest that part of this has to do with the precision required in order to produce a fricative versus a stop consonant, and especially a non-strident interdental fricative (e.g. Hardcastle 1976); the latter being particularly difficult for ESL learners. Even for native English speakers, there is a tendency to articulate a stop or an affricate (e.g. [dð]) when targeting an interdental fricative (more commonly it seems in the unstressed voiced variants). In other words, ease of articulation contributes to a certain propensity for all speakers to produce a stop in place of the interdental fricative; thus, for ESL learners who misperceive [θ] as /s/, production will show a grammar-driven variant, [s], along with performance-driven [t], but for ESL learners who misperceive [θ] as [t], both linguistic and extra-linguistic factors conspire to yield [t] in production. Thus, for QF ESL learners, there may be three sources of difficulty reinforcing [t] production: perceptual confusion, orthographic influences, and ease of articulation. It is possible that the combination of these factors leads to fossilization in production.

The findings presented in the section generally support the hypothesis that production errors are due to perception errors. However, there are discrepancies between perception and production. It was argued that some of these disparities may be because production is subject to various influences to which perception is not, such as visual cues encoded in lexical representations, orthography, and pressure towards ease of articulation. All in all, the production substitutes in this study correspond to the perceptual substitutes that were predicted by the Auditory Distance Model for phonetic processing. A major ingredient in these predictions concerning differential substitution was that languages with a non-strident or less strident /s/ would tend to replace target theta with their native /s/. The next section in this chapter addresses this question.

6.3. Spectrographic Analysis

In this section, I examine the phonetic properties of the fricative /s/ particular to four language groups: European French, Québec French, Japanese, and Russian. As discussed in previous chapters, it is hypothesized that cross-linguistic differences in the perception of the English interdental fricatives are due to their phonetic proximity with native language (L1) sounds. In particular, it is hypothesized that languages whose coronal sibilant /s/ is phonetically MELLOW (non-strident) will perceptually associate this sound with the interdental fricative [θ]. On the other hand, languages whose coronal sibilant is phonetically STRIDENT will not make this association, and will instead link the target [θ] with their coronal stop /t/. The current analysis focuses mainly on two hypotheses: First, that the choice of substitute depends on evidence from the L1 sound system; and second, that the transfer of non-contrastive, phonetic features is involved in interdental substitution.

Let us recall the motivation behind these hypotheses. It is thought that there is something about the native language that drives substitution patterns. In other words, the choice of substitute for an L2 sound is not accidental -- it is based on internal evidence from the L1. European French and Québec French have identical consonant inventories at the phonemic level, yet these two dialects demonstrate differential substitution of the English interdental fricatives. This discrepancy suggests that evidence for the substituted sound cannot be attributed to contrastive feature specifications. Instead, it is hypothesized that non-contrastive phonetic specifications that differ between EF and QF constitute the evidence that accounts for their differential substitution patterns (as well, phonetic differences between QF and JA, RU also explain differential substitution.)

As I have pointed out in Chapters 2 and 3, it is proposed that the phonetic evidence lies in the different articulatory manifestations of the coronal /s/ across languages. Specifically, I have proposed that EF and RU have a dental, and therefore less strident, [s̪]; whereas, QF has an alveolar, and therefore more strident, [s]. I have also proposed that Japanese has a mellow (non-strident) [s̠].

This chapter attempts to identify phonetic differences in the cross-linguistic expressions of /s/ by looking at spectrographic information from speakers of EF, QF, RU, and JA.

6.3.1. Stimuli

For each of the test languages, the stimuli were real words beginning with /s/ before each of five vowels, /i e a o u/. The material for each test language was developed in collaboration with a linguist who is a native speaker of that language. The criteria used in the choice of target words were as follows: The word was monosyllabic. As much as possible, content words were used. If possible, a word of CV shape was chosen; however, if no appropriate word existed, then a CVC word was selected. No words with complex codas were used. If a word with a coda was used, an attempt was made to find a word with an "unmarked" coda, i.e. a cross-linguistically frequent coda consonant. The first type of unmarked coda searched for was a nasal coda, preferably /n/. If this was not possible, then a coda with a velar stop was preferred.

The list of words chosen for each language is provided in (1):

(1) French: "si, C, sac, saut, sous" [si se sak so su]

("yes/ok/if, the letter "C", bag, jump, penny/beneath")

Japanese: "sin, se, san, son, su" [çin se san son su][81]

("core, height, three, loss, vinegar")

Russian: "sip, sev, sad, sok, suk" [sip sef sat sok suk]

("hoarseness, sowing, garden, juice, branch")

Each word was embedded in a carrier sentence. For each language's carrier sentence, the word preceding the target terminated in a nasal consonant. The carrier sentences are shown in (2):

(2) French: "Ils apprennent ___ encore." ("They (masc.) learn ___ again.")

Japanese: "Kare wa totsuzen ___ to itta." ("He suddenly said ___.")

Russian: "Skazal on ___ op'at'." ("He said ___ again.")

6.3.2. Participants

For the European French group, there were three participants, all female. One of these participants had also done the Picture Identification Task; the others were newly recruited for the Spectrographic Task. In the Québec French group, there were five participants: four females and one male. All were newly recruited for this task. The Japanese group consisted of three participants: two female and one male. One had participated in the AXB2 and Picture ID Task, while the other two were newly recruited. There were six participants in the Russian group: two females and four

---

[81] Recall that in Japanese /s/ → [ç] / _ [i].

males. One had participated in the AXB2 and Picture ID tasks, while the other five were newly recruited.[82]

6.3.3. Procedure

Participants read each group of five sentences three times at a normal rate of speech. They were recorded on a PC using Kay Elemetrics Computer Speech Lab software at a 22kHz sampling rate. Each of the three tokens for every target word was excised from the carrier sentence and analyzed using Praat software (Boersma & Weenink 2010).

There are several attributes which distinguish between fricative consonants (e.g. Behrens & Blumstein 1988; Nartey 1982; Strevens 1960; Hedrick & Ohde 1993; Jongman et al. 2000). In this experiment, three different measures were taken: Spectral mean, intensity difference between the fricative and following vowel, and length difference between the fricative and following vowel.

1. **Spectral Mean (Centre of Gravity).** The spectral mean of a fricative refers to the average concentration of energy at a particular frequency (see Jongman et al. 2000). This measure is thought to be indicative of the frontness of the place of articulation. A higher spectral mean signifies a more fronted articulation. Thus, we might expect a dental [s̪] to have a higher spectral mean than an alveolar [s] (see Boersma & Hamann 2007).

For analysis, the fricative was isolated, and the spectral mean was calculated using a Fast Fourier Transform (FFT).[83]

---

[82] The uneven group sizes are due to difficulty in recruitment due to time constraints. Ideally, monolingual speakers would be preferred for this type of task, as it is known that phonetic adjustments to L1 articulation occur in bilinguals (e.g. Flege 1995 and other work). Future research should take this into consideration.

2. **Intensity.** The phonetic/phonological features STRIDENT and MELLOW are acoustically based on the degree of intensity or amplitude (Jakobson, Fant, & Halle 1963/1969). Sibilant fricatives are relatively high in amplitude as compared to non-sibilants. Recall from earlier chapters that Japanese [s̪] has been described as being less strident than English [s]. It was argued that this would cause Japanese learners of English to associate [θ] with their native coronal fricative.

In addition to fricative amplitude, vowel intensity was also measured, for two reasons. First, Stevens (1985) found that for English listeners, when the amplitude of a fricative was higher at certain formant frequencies in comparison to the following vowel, then a [s] was perceived. Conversely, if the amplitude of the fricative was lower than the vowel at these frequencies, then a [θ] was identified. Second, comparing fricative and vowel amplitude controls for individual variation in amplitude: some speakers may simply speak louder than others. Therefore, this measure permits a relative evaluation of fricative intensity.

Amplitude (in decibels (dB)) was averaged across the entire length of the fricative (see Behrens & Blumstein 1988). The average intensity of the vowel was measured from the onset of voicing (first glottal pulse) to the end of the vowel.

3. **Length.** The length of noise duration associated with fricatives has been shown to distinguish [s] from [θ] (e.g. Behrens & Blumstein 1988). It was therefore decided to measure relative length in case this cue is attended to by ESL learners. While it is hypothesized that amplitude is a correlate of

---

[83] It was found that there was some low frequency energy in the stimuli (perhaps from voicing leak or due to the computer fan); therefore, before subjecting the fricatives to the spectral mean analysis, they were pass band filtered below 2000 Hz.

stridency, it is also possible that the feature STRIDENT could manifest itself in length relations at the phonetic surface, and that learners whose first language has a short coronal fricative might tend to associate the relatively short [θ] with their native /s/.

Vowel length was measured in order to get a relative measure of fricative vs. vowel length. This was done because there are individual differences in rate of speech; therefore absolute duration may vary as a function of speaking rate, such that some speakers produce long fricatives and vowels; whereas, others produce short fricatives and vowels (Jongman et al. 2000). Thus, this measure permits a relative evaluation of fricative length.

The fricative was measured from the onset of aperiodic noise to its offset, i.e. where the waveform started to take on a discernable periodic shape. Vowel length was measured from the first glottal pulse associated with the onset of voicing to the end of the vowel in CV words, or to the beginning of the transition into the coda consonant in CVC words.[84]

6.3.4. Results

Spectral Mean, the difference in intensity between the fricative and vowel (Intensity F-V), and the difference in length between the fricative and vowel (Length F-V) were analyzed in separate two-way ANOVAs (Language x Vowel).

[84] Vowels are often shorter in CVC syllables than in CV syllables, and the cross-language stimuli did not match in syllable form (see (44) above). However, as we shall see later, it appears that syllable shape does not affect vowel length in these test languages.

6.3.4.1. Spectral Mean

Results on Spectral Mean showed a significant main effect for Vowel ($F(4,172)=9.36, p<.001$), Language ($F(3,43)=8.73, p<.001$), and a significant Language x Vowel interaction ($F(12,172)=4.68, p<.001$). Table 6.5 gives the Spectral Mean and Standard Deviation measured in Hertz (Hz) for each native language group (European French, Québec French, Japanese, Russian) in each vowel context (/i e a o u/).

Spectral Mean (Hz)		**i**	**e**	**a**	**o**	**u**	**Mean Total**
MEAN	**EF**	6830	6859	7102	6671	6671	6826.63
	QF	7311	7347	7332	7158	7016	7232.62
	JA	5328	6425	6664	5992	6063	6094.42
	RU	6163	6299	6394	6263	6712	6366.38
StDev	**EF**	614	692	592	1141	1433	
	QF	431	493	541	945	857	
	JA	152	447	656	492	480	
	RU	753	456	847	577	578	

Table 6.5. Spectral Mean for native /s/ in Hertz.
Means and Standard Deviations are provided for each vowel context.

Non-parametric Mann-Whitney U tests were conducted to determine the source of significance in the interaction. These revealed a significantly higher spectral mean for EF, QF, RU compared to JA on /i/, for QF compared to JA, RU on /i e/, and for QF compared to RU on /a/ ($p<.005$).[85] The difference between QF and JA, RU on /o/ and between QF and JA on /u/ was close to significant ($p<.01$).

[85] As with the AXB-2 and PicID statistical reports, a level of .005 is used for the non-parametric tests.

The lower spectral mean for the JA speakers is expected as Japanese /s/ becomes [ɕ] before [i]. This allophonic variant is an alveolo-palatal fricative, whose posterior place of articulation explains the low values. These data also suggest that JA and RU /s/ before /e/ is articulated farther back than it is in QF.

The results on Spectral Mean do not appear to lend support to the expectation that EF and RU have a dental [s̪] as compared to QF having an alveolar [s]. In fact, although not significant, the Spectral Mean for QF is actually higher than that for EF and RU, suggesting that the /s/ of QF is more fronted than it is in EF and RU. However, although it is generally considered that spectral mean is indicative of place of articulation, Fuchs & Toda (2010) argue that it can also be lowered by a wider constriction area (formally expressed by the feature LAMINAL in this thesis). They found that German speakers had a wider constriction in their articulation of [s], similar to the characteristics of English [θ], to which they attributed lowered spectral means for German [s]. This would be an interesting avenue to pursue in further research: if EF, RU, or JA has a relatively wide constriction as compared to QF, this may be yet another factor contributing to the auditory similarity between their native /s/ and English [θ].

6.3.4.2. Intensity Fricative-Vowel

Analysis of the relative intensity of the fricative compared to the following vowel revealed a significant main effect for Vowel ($F(4,164) = 6.6, p = .001$), and a significant interaction for Language x Vowel ($F(12,164) = 8.4, p < .001$). Table 6.6 shows the means and standard deviations of the difference in amplitude between the fricative and vowel in decibels (dB).

Intensity F-V (dB)		i	e	a	o	u	Mean Total
MEAN	EF[86]	-7.12	-7.26	-7.60	-12.09	-10.38	-8.89
	QF	-11.34	-12.97	-11.34	-12.97	-15.00	-12.72
	JA	-13.56	-12.34	-13.33	-14.30	-6.71	-12.05
	RU	-16.24	-15.93	-16.19	-16.07	-17.43	-16.37
StDev	EF	11.00	9.47	8.74	7.24	10.16	
	QF	2.64	2.36	2.64	2.36	2.75	
	JA	2.96	3.82	3.12	1.75	8.04	
	RU	2.87	2.90	2.83	2.97	2.89	

Table 6.6. Fricative-Vowel Intensity in dB for native /s/.
Means and Standard Deviations are provided for each vowel context.

Non-parametric Mann-Whitney U tests indicated that RU has a significantly greater difference in fricative intensity in relation to vowel intensity than does QF in the context of /i a/ ($p < .001$). The difference between RU and JA is also significant before /u/ ($p < .005$). This latter finding can be attributed to the very short realization of JA [ɯ], and the fact that most tokens of it were voiceless.[87] And as can be seen by the numbers in Table 6.6, RU has the greatest discrepancy between fricative and vowel intensity of all the languages. Table 6.7 shows that this effect is largely due to lower amplitude on the fricative in RU.

[86] There is more variation in the EF group than in the other groups. This is largely due to one participant. If this participant is removed from the data, the means for EF are: -14 -13 -12 -16 -17 for /i e a o u/ respectively. This does not result in significance.

[87] In some tokens, there was no [ɯ] pronounced at all. This is because of a rule in Japanese whereby the high vowels [i ɯ] are devoiced or even deleted between voiceless consonants or, less consistently, between a voiceless consonant and a pause (Vance 2008). In the present study, instances where [ɯ] was deleted were excluded from the analysis, but the voiceless productions were included.

	i		e		a		o		u	
	$\bar{X}$IntF	$\bar{X}$IntV	$\bar{X}$IntF	$\bar{X}$IntV	$\bar{X}$IntF	$\bar{X}$IntV	$\bar{X}$IntF	$\bar{X}$IntV	$\bar{X}$IntF	$\bar{X}$IntV
EF	63	71	64	70	64	71	64	71	61	71
QF	64	76	64	77	64	75	64	75	61	77
JA	65	79	64	77	64	77	64	77	64	43
RU	61	77	61	77	61	77	61	77	59	76

Table 6.7. Mean Intensity in dB for native /s/ ($\bar{X}$IntF) and for vowel ($\bar{X}$IntV).

The "quieter" fricative in RU may be another factor that contributes to lessening the auditory distance between target English [θ] and native Russian [s̺]. Thus, measures of intensity indicate that RU [s̺] is not a strongly strident fricative.

6.3.4.3. Length Fricative-Vowel

Results from comparing the relative length of the fricative to that of the following vowel indicated a significant main effect for Vowel ($F(4,180)=26.94, p<.001$), and a significant interaction for Language x Vowel ($F(12,180)=2.12, p<.05$). Table 6.8 shows the means and standard deviations of the difference in length between the fricative and vowel in seconds (sec).

Length F-V (sec)		i	e	a	o	u	Mean Total
MEAN	EF	0.07	0.07	0.04	0.02	0.07	.05
	QF	0.13	0.11	0.09	0.09	0.13	.11
	JA	0.07	0.02	-0.01	0.01	0.08	.03
	RU	0.07	0.02	0.03	0.03	0.06	.04
StDev	EF	0.07	0.07	0.03	0.04	0.06	
	QF	0.02	0.06	0.04	0.05	0.02	
	JA	0.02	0.02	0.03	0.02	0.01	
	RU	0.03	0.04	0.04	0.05	0.04	

Table 6.8. Fricative-Vowel Length in sec for native /s/.
Means and Standard Deviations are provided for each vowel context.

Non-parametric Mann-Whitney U tests showed that QF has a significantly greater difference in vowel length in relation to the fricative than does EF in the context of /o/ ($p=.005$). This same difference between QF and EF nearly reaches significance for /i/ and /u/ ($p=.007$, $p=.009$ respectively).

Moreover, QF displays a significantly larger fricative to vowel length than both JA and RU in all vowel conditions ($p<.005$). And EF shows a significantly greater difference in the context of /a/ ($p=.005$).

	i		e		a		o		u	
	$\bar{X}$LgF	$\bar{X}$LgV	$\bar{X}$LgF	$\bar{X}$LgV	$\bar{X}$LgF	$\bar{X}$LgV	$\bar{X}$LgF	$\bar{X}$LgV	$\bar{X}$LgF	$\bar{X}$LgV
EF	.23	.16	.25	.16	.20	.16	.20	.18	.21	.14
QF	.23	.10	.21	.10	.18	.10	.21	.12	.21	.08
JA	.13	.06	.09	.07	.09	.09	.10	.08	.11	.03
RU	.13	.05	.11	.09	.11	.09	.11	.08	.11	.06

Table 6.9. Mean Length in sec for native /s/ ($\bar{X}$LgF) and for vowel ($\bar{X}$LgV).

Table 6.9 shows the mean length of the fricative for each vowel. From this table, we can see that the difference between EF and QF is not stemming from disparate length of the fricative -- both have rather long fricatives -- but instead from the vowel, with QF having shorter vowels than EF. We also see from Table 6.9 that JA and RU *do* have short fricatives compared to the French groups: their fricatives are about half as long. However, their vowels are very short too. This leads to the fricative-vowel ratio for JA and RU (both short) being analogous to that of EF (both long). Québec French stands apart by having long fricatives and short vowels. And this is more obvious in the context of certain vowels.

Note that these results show that syllable shape, i.e. whether CV or CVC, does not have an influence on vowel length, and hence is not affecting the results. For both EF and QF, the same stimuli were used; for RU, all stimuli were CVC; and for JA, we see in Table 6.9, that the order from longest vowel to shortest vowel was as follows: /a o/ (both CVC) > /e/ (CV) > /i/ (CVC) > /u/ (CV).

The results from this analysis of relative length are compatible with the hypothesis that QF has a more strident fricative than the other languages. The long fricative/short vowel relation in QF will no doubt make the fricative more salient, and hence less likely to be chosen as an interdental substitute. In fact, I would like to suggest that a high fricative to vowel ratio may be a phonetic interpretation of the feature STRIDENT.

6.3.4.4. Comparison to English Measurements

Jongman et al. (2000) carried out a study on the acoustic characteristics of English fricatives. The participants in that study were 20 American English speakers, male and female. The stimuli consisted of CVC syllables in which the fricative was initial and the final consonant was [p]. Table 6.10 presents a summary of their measurements that are pertinent to the current study. These figures conflate voiced and voiceless fricatives as well as vowel context (6 different vowels). Normalized Amplitude corresponds to Intensity F-V (dB) in Table 6.6 above.[88] Jongman et al. measured fricative duration relative to the entire CVC syllable, whereas I measured it relative to CV only; thus only fricative duration is presented in Table 6.10.

English	Spectral Mean (Hz)	Normalized Amplitude (dB)	Fricative Duration (sec)
[f]	6625	-20.8	.17
[θ]	6250	-21.9	.16
[s]	7000	-11	.18
[ʃ]	4875	-9.9	.18

Table 6.10. Summary of measurements of English fricatives in Jongman et al. (2000).

In general then, the measurements reported in this section point to a less strident version of /s/ in JA and RU, and a more strident [s] for QF. Jongman et al. (2000), found that the spectral mean for English [θ,ð] was around 6250 Hz and [s,z] was around 7000 Hz.[89] Therefore, it is possible that a speaker of a language whose spectral mean for /s/ is nearer 6000 Hz might tend to associate target [θ] with their native /s/ (Japanese 6094 Hz, Russian 6366 Hz; see Table 6.5).

In Jongman et al., the noise amplitude for [f] = 55.7, [θ] = 54.7 dB, [s] = 64.9 dB, and [ʃ] = 66.4. The vowel amplitude is 76 dB. So the disparity in amplitude between [θ] and the vowel is greatest. Of the languages tested in the current study, Russian shows the greatest gap in

[88] Intensity and Amplitude are not exactly equivalent. Intensity can be seen as amplitude over time over an area. Thus, intensity equals the square of the amplitude. Given that the dB values given for Intensity in Praat are in the same range as the Normalized Amplitude reported in Jongman et al., I assume these are the same measures.
[89] This is inferred from their Figure 3 (p.1257), middle of the fricative.

intensity between their [s] and the vowel, another factor which may bias these ESL learners to associate [θ] with /s/ or /f/.[90]

As mentioned above, the measures of length in the current study cannot be directly matched to those of Jongman et al., since in the latter, they compared length of fricative relative to length of following VC; whereas, I compared to V alone. Nonetheless, in Jongman et al., all the fricatives are proportionally shorter than the remainder of the syllable, but [θ] is the shortest of the voiceless fricatives. In the present study, in all but one case (JA [a]), the fricative is longer than the vowel; however, for QF, the proportion of fricative to vowel is the greatest compared to EF, JA, and RU. Thus, the relatively long [s] of QF might make it an inappropriate substitute for target [θ].

6.3.5. Summary and Discussion of Spectrographic Analysis

This section has investigated the acoustic properties of the native /s/ of European French, Québec French, Japanese, and Russian. These measurements were compared with analogous ones for Native English obtained from an article by Jongman et al. (2000).

In English, [θ] has a lower spectral mean than [s]. In the current analysis, results revealed a significantly lower spectral mean for Russian and Japanese /s/; particularly when compared to QF. This has been shown to be affected by vowel context. And for Japanese, this is in part attributable to the [ɕ] allophone that occurs before [i]. It was suggested that low spectral mean could lead to JA and RU associating [θ] with their native /s/.

---

[90] Note also that if we remove the speaker who diverges from the other two in the EF group, the values approach those of RU (see footnote 84).

Another measure investigated was the intensity of the fricative in relation to the following vowel. Russian speakers displayed the lowest intensity on /s/ compared to the vowel. Since English [θ] is also of relatively low intensity, this may be another factor driving RU speakers towards /s/ as a substitute.

Finally, the relative length of the fricative to the vowel was also measured. In English, [θ] is quite short compared to [s] and [ʃ]. Out of all the languages examined in this chapter, Québec French stood out in that their native [s] is substantially longer than the following vowel. The difference between EF and QF here is not in the length of the fricative, but in the length of the vowel relative to the fricative, with QF vowels being significantly shorter than EF vowels. I argue that this contributes to the salience of the native [s] in QF, and diverts them away from /s/ as a substitute for [θ], unlike the other languages.

The analysis of the acoustic characteristics of /s/ in this chapter indicates that the feature STRIDENT may have several phonetic correlates. This abstract feature may manifest itself in more ways than simple intensity or amplitude: other correlates such as spectral mean and relative length should also be considered.

At the outset of this section, it was hypothesized that EF and RU would show evidence of having a dental [s̪], and QF, an alveolar [s]. A higher spectral mean would indicate a dental [s̪]. This does not appear to be the case. Both EF and QF had higher spectral means than JA and RU. On the other hand, as discussed in Fuchs & Toda (2010), English [θ] has a relatively low spectral mean because of the wide constriction. So maybe it is not the dental/alveolar dimension which is relevant, but instead width of constriction, or laminality.

It was also predicted that EF, JA, and RU would have a less intense /s/ than QF. Although acoustic measurements revealed this not to be the case for EF and JA, it was confirmed for RU. So again, this is further evidence to support an /s/ substitute in Russian.

Measurements of fricative length in relation to vowel length yielded differences between QF and EF, JA, RU. This measure serves to distinguish QF [s] from the /s/ of the other languages, and provides support for the phonetic manifestation of an enhanced STRIDENT feature in QF.

These findings also show that vowel context can alter the acoustic characteristics of /s/. Not surprisingly, it was found that JA /s/ before [i] had a lower spectral mean than in other contexts. This is of course attributable to the JA allophonic variant [ɕ] occurring in this environment. The lowering of the spectral mean in this context may explain some of the findings in AXB-2 where JA was shown to have more difficulty on the [s-θ] contrast before high front vowels (see Chapter 4). Some evidence of backing of /s/ before [i] was also found in Russian, and to a lesser degree before back vowels in both JA and RU. In addition, relative length of fricative versus vowel may favour the fricative in cases where the vowel is shorter, making it more salient.

In summary, this section has shown phonetic differences in the articulation of /s/ in each of EF, QF, JA, and RU. It has been argued that the direction of these differences points towards the feature STRIDENT being strongly specified (enhanced) on QF [s]; whereas, for the /s/ of EF, JA, and RU, it is either specified as MELLOW or weakly STRIDENT (muted).

6.4. Summary and Discussion of Production

In this chapter, we have seen results from two production experiments: the Word Production task and the Spectrographic Analysis. The Word Production task demonstrated that production errors largely coincide with perception errors. One major discrepancy to this generalization was the high perception errors involving [f] as compared to the relatively lower rate of /f/ production errors. This was ascribed to the availability of visual information in lexical representations during the production process. Aside from this, the most frequent production substitute corresponded to the most frequent perception substitute for each of the languages tested. These were the substitutes that were predicted by the Auditory Distance Model for phonetic processing. Therefore, the results from the Word Production task are consistent with the hypothesis that perception determines production.

The findings from the Spectrographic Analysis are largely compatible with the hypothesis that a non-strident or less strident /s/ in the native phonetic inventory will be a factor in /s/ substitutions for the target English interdental fricative.

Regarding intensity, it was found that RU [s] had a significantly lower amplitude than that of QF and JA. This is taken an evidence for a less strident [s] in RU, making it a relatively good match to low intensity [θ].

It was predicted that the spectral mean would be lower for QF than for EF and JA, a lower spectral mean being indicative of a less anterior articulation. Results, in fact, were almost the opposite of what was predicted, with QF having a significantly higher spectral mean than JA and RU. However, it was proposed that a lower spectral mean may also be indicative of a wider constriction area, similar to that of English [θ] (Fuchs & Toda 2010). If this were the case, then the /s/ of JA and RU would resemble [θ] in terms of constriction area, being formally defined as LAMINAL; whereas, QF [s] would be more similar to English [s], with a narrower constriction area, defined as APICAL.

Finally, phonetic differences were found between EF and QF, with QF having a significantly longer fricative in relation to the following vowel as compared to EF, JA, and RU. It is argued that this relatively longer fricative in QF could be a manifestation of the feature STRIDENT, and if this is the case, these results would support an enhanced version of the feature STRIDENT in Québec French as compared to the other languages tested.

CHAPTER 7: GENERAL DISCUSSION AND CONCLUSION

The work presented in this thesis constitutes a theoretical and empirical contribution to the fields of second language acquisition and the perception/production and phonetics/phonology interfaces. The theoretical contribution consists of a model of speech processing in which both phonetic and phonological information interacts to yield cross-language differences in perception and production. The model incorporates the Auditory Distance Algorithm which in turn integrates feature enhancement. This model was developed in order to explain second language patterns that are mysterious when viewed from a purely phonological perspective.

In Chapter 1, it was argued that differential patterns of substitution observed in second language production and perception are poorly accounted for within models that rely solely on phonological features, i.e. those which function contrastively in a language. Two L1s or two dialects of a single L1 may fail to display any language-internal evidence to support different contrastive feature specifications: both having the same phonemic inventory of sounds, and relevant phonological processes. For such cases, which include European French and Québec French, the language-internal evidence must be sought elsewhere. It was proposed that it is to be found in the phonetic component of the grammar.

Chapter 2 addressed this by elaborating the Auditory Distance Model (ADM), a model which is applied to second language processing. It was argued that language-specific phonetic features must play a role in cross-language substitution. In the ADM, the acoustic signal is converted into a form which is compatible with the language-specific phonetic component, the intake. Features in the phonetic intake are matched against combinations of features in the L1 phonetic inventory.

In the context of second language acquisition, following Schwartz & Sprouse (1994), I assume that learners theoretically have access to all features available in Universal Grammar. However, in the initial state of L2 acquisition, learners make use of their L1 grammar. In perceptual processing, intake features are matched to stored representations, which include both phonetic and phonological information. Some features may be permanently underspecified through to the phonetic component in the L1, or some combinations of features may not exist in the L1. Thus, in the beginning stages of acquisition, some features or combinations may not be parsed from the L2 target, either because they do not exist in the L1 or because the L2 target consists of a featural combination which is not part of the L1 grammar. These are the cases where transfer is observed, as the L2 learner matches the L2 intake to the phonetically closest L1 representation. As acquisition progresses, it is considered that new features can be acquired and that new combinations of features can be assembled at later stages.

An advantage of the ADM is that it offers an explicit formulation for how target intake sounds are matched to an L1 representation. An algorithm was proposed which calculates the auditory distance between the intake and a combination of L1 phonetic features. A particular feature in this combination may be subject to enhancement; i.e., the perceptual salience of this feature is augmented under these conditions. Enhancement is tied to markedness in the sense that it applies to universally perceptually optimal combinations of features, which in turn accounts for their frequency in languages of the world. For example, when the feature STRIDENT co-occurs with CONTINUANT, this results in the application of the enhancement relation STRIDENT ENHANCES CONTINUANT. This encodes the concept that continuancy is best perceived on a fricative that is sibilant as opposed to a non-sibilant fricative. When a constellation of features does not result in feature enhancement, i.e. does not constitute an optimal combination, this has been referred to as non-enhancement or muting.

By definition, language-specific phonetics admits cross-language variation in the surface specification of a particular phoneme. Consequently, enhancement relations also vary cross-linguistically. A particular feature in the phonetic realization of a phoneme in one language may be subject to enhancement; whereas, in another language, the same feature may be muted. This is similar to the notion of cue weighting in other models, but differs from some in its link to cross-linguistic markedness. Furthermore, the enhancement relations proposed in this thesis are part of the grammar.

The level of detail that has been incorporated into the ADM model is missing from many other models of second language processing such as implementations of the Perceptual Assimilation Model (PAM) of Best (1995). The ADM explicitly details each feature and the perceptual weight accorded to each feature; therefore, it is more amenable to empirical testing.

In this thesis, the ADM model was empirically verified by investigating the perception and production of target English interdental fricatives by L2 learners from European French, Québec French, Japanese, and Russian L1 backgrounds. Application of the Auditory Distance Algorithm in conjunction with the PAM to detailed phonetic specifications for each of these languages predicted perceptual substitutes that conformed to production substitutes.These predictions were tested in Chapters 3 (AXB-1) and 4 (AXB-2). Incorporated into these experimental tasks were two conditions intended to access different processing levels: the phonological level in the Long ISI condition, and the phonetic level in the Short ISI condition. It was predicted that under phonological processing, only features that function contrastively in the L1 would be perceived. For those test languages purported to have the same set of relevant contrastive features, it was expected that behaviour on this task would be the same. However, in the phonetic condition, where non-contrastive phonetic features and enhancement effects are operative, it was anticipated that language-specific perceptual differences would emerge.

The results from the AXB tasks, especially those from the less advanced learners, were largely consistent with the predictions. However, no difference was found between the Short and Long ISI conditions. It was argued that the AXB tasks tapped phonetic processing, but not phonological processing.

In Chapter 5, the hypotheses were tested in a decidedly phonological paradigm, the Picture Identification task. The results from this experiment in some ways corresponded more to the predictions made for phonetic processing than for phonological processing. It was thus concluded that the phonological component of the grammar is richer than has traditionally been assumed, especially for L2 learners. In other words, it embodies some redundant, non-contrastive features in addition to phonemically contrastive features.

The results from all tasks showed that context also affects the perception of obstruents. Perception was influenced by voicing (AXB-1), vowel (AXB-2 and PicID), and syllable position (PicID). This means that listeners are attending to different perceptual cues in voiceless versus voiced contrasts, and are also relying on information from adjacent segments such as formant transitions from neighbouring vowels.

In the two tasks that investigated voicing, AXB-1 (word-initial onset) and PicID (word-final coda), it was found that where differences arose between voiceless and voiced contrasts, the voiceless contrasts were more difficult to perceive than the voiced ones. This is somewhat surprising, given that in word-initial onset and word-final coda positions, voiced obstruents are cross-linguistically less frequent than voiceless ones, and therefore presumably "marked" in some way. Future research should investigate the voicing differences more thoroughly.

Although all three studies showed an influence of vowel context, the results were rather mixed across tasks. Nevertheless, some patterns did emerge. One of these relates to alternations in the L1. Some interactions involving the vowel condition in AXB-2 were argued to be due to L1 allophonic processes before high vowels. Both Japanese and Québec French have these types of processes: JA /si/ → [çi] and /tɯ/ → [tçɯ]; QF /ti/ → [tsi]. Results indicated that in cases where the contrast could be construed as a mismatch between the allophone and the context, these listeners had more difficulty.

However, this did not emerge as significant in the PicID task. The affricate was not tested in the PicID task, so we cannot assess it, but the JA /si/ → [çi] alternation is relevant. Why did /si-θi/ cause particular problems for JA listeners in AXB-2 but not significantly more than other vowels in the PicID task? I speculated that this may be because in the AXB task, which has been deemed to involve phonetic processing, vowel context matters for the application of allophonic processes. However, in the PicID task, which involves phonological processing, neutralization of [s] and [ç] before [i] means that whether [s] or [ç] occurs before [i], it matters not: both will be categorized as /si/.

A finding regarding vowel context which did occur in both the AXB-2 and the PicID studies involved the /f-θ/ contrast. In both tasks, a following high front vowel disfavoured accurate perception. This contrast is already acoustically fragile (Miller & Nicely 1955), but it seems that a HF vowel makes it even more so. It was suggested that the spread lip configuration of the HF vowel may eliminate cues to place of articulation differences between these. In the AXB-2 study, this vowel context affected all language groups, included Native English. However, in the PicID task, the NE group was significantly less influenced by the vowel. This again suggests that the two tasks involve different types of processing. When a phonetic comparison is made, vowel context plays a greater role; however, when the task involves phonological categorization, vowel context is less important, especially for more proficient learners (and of course native speakers).

The PicID task showed that not only did the /f-θ/ contrast interact with the vowel context, but listeners' performance was also affected by syllable position. It was generally more poorly perceived in coda position, but this was mitigated by the vowel. In coda, /f-θ/ was more difficult for all groups after a non-high vowel, but in onset, these sounds were harder to classify before a high front vowel, as mentioned above. There were other interactions with position, but these were less robust.

Another observation from these studies is that for [f-θ] in both AXB tasks, and for [s-θ] and [s̺-θ] in the AXB-2 task, the Native English group had a considerably high rate of errors. As well, QF had unexpectedly high errors on [s-θ] and [s̺-θ] in AXB-2. On the other hand, in the PicID task, the patterns were as anticipated: NE showed low rates of errors on all contrasts, significantly lower than the test groups, and QF had significantly fewer errors on /s-θ/. I speculated that this is due again to the nature of the tasks. For the AXB task, especially in the Short ISI condition, it may be that participants sometimes switch into an acoustic processing mode. In other words, they may not be assigning categories to the stimuli, contrary to the PAM. Instead they may be trying to compare auditory traces. The high memory load involved in this task causes these acoustic traces to fade, and perhaps the most vulnerable information is that which distinguishes two fricatives, similar to what is degraded under noise conditions.

Apart from this, the findings from the AXB tasks and the PicID task generally correspond, and support the hypotheses set forth in the thesis. The results for EF, QF, and JA from the perception experiments agree with what has been reported in the production literature: EF and JA are more likely to perceptually confuse [θ-ð] and [s-z]/[s̪-z̪]; whereas QF is more apt to merge [θ-ð] with [t-d]/[t̪-d̪]. This emerged most strongly with less advanced learners, which indicates that native-like perceptual representations of these sounds can eventually be acquired.

For RU, the perception results showed that they pattern with EF in substituting the sibilant fricatives for the interdentals; although a couple of RU participants tended to perceive and produce [f]. The findings for RU contradict reports that have been propagated in the literature stating that RU speakers substitute stops in place of the English interdental fricatives. It should be noted however that for the RU group, there was some variation in both perception and production substitutes, especially in coda position; therefore, this might account for some of the discrepancies in the literature.

Returning again to the [f-θ] pair, as discussed above, it was found that perceptual errors were relatively high for this contrast. Yet, Chapter 6 showed that production of this substitute was generally low. This disparity between perception and production was attributed to visual cues differentiating [f] from [θ] which are encoded in lexical representations. This would explain the low rate of [f] production substitutes. All in all though, the results from the perceptual tests and the Word Production test in Chapter 6 support the hypothesis that production errors mirror perception errors.

The other major hypothesis underpinning this work was that transfer of non-contrastive phonetic features is involved in interdental substitution. The impetus behind this hypothesis was the observation of differential substitution between two dialects of French, QF and EF, which share the same set of consonant phonemes, and for whom it was argued share the same relevant contrastive features. It was suggested that the feature STRIDENT is of particular relevance. In particular, it was proposed that STRIDENT is more salient on an alveolar [s] versus a dental [s̪]. This was formally encoded as ALVEOLAR ENHANCES STRIDENT (or conversely DENTAL

MUTES STRIDENT). Thus, one of the predictions was that QF, whose L1 phonetic inventory includes an enhanced strident alveolar [s], would not associate it with non-strident [θ]; however EF, with a muted strident dental [s̪], would. Results from the perception and Word Production tasks were generally consistent with this analysis.

Another hypothesis behind this work is that languages that select a /t/ substitute for the target interdental fricatives have an anterior coronal fricative for which the feature STRIDENT is salient; whereas, languages that choose /s/ as a substitute do so because their /s/ is less strident. To test this, an acoustic analysis was conducted on the native /s/ of EF, QF, JA, and RU. This analysis revealed cross-language acoustic differences on the dimensions of spectral mean, normalized amplitude, and ratio of fricative to vowel length.

Results suggested that JA and RU /s/ is closer to [θ] than is QF [s] on the spectral mean dimension. In terms of normalized amplitude, RU [s̪] is also more similar to [θ] than for the other language groups. For fricative to vowel length ratio, EF, JA, and RU were all significantly different from QF in having relatively short fricatives compared to the vowel. Data from Jongman et al. (2000) show word-initial [θ] to be the shortest of the four English fricatives. Thus, in terms of length, EF, JA, and RU /s/ is more similar to the interdental compared to QF [s]. These findings, although not exactly as expected, are nevertheless consistent with the hypothesis that QF [s] is phonetically farther from the interdental fricative than is the /s/ of EF, JA, and RU. This suggests that the feature STRIDENT may manifest itself in more ways than simple intensity or amplitude: other correlates such as spectral mean and relative length should also be considered. This reinforces the idea that a particular abstract phonetic or phonological feature can be associated with different acoustic cues.

This thesis has contributed to our knowledge about how learners perceive and produce L2 sounds. To my knowledge, it is the first empirical investigation of the perception of interdental fricatives by English L2 learners from both European French and Québec French backgrounds, and the only experimental study of Russian. As well, with the exception of Hancin-Bhatt (1994a,b), no studies have examined interdental substitution as thoroughly.

In addition, the Auditory Distance Model has offered an explicit framework within which to assess how second (and first) language learners process segmental input. The type of detailed assessment implicit to the ADM is often glossed over in other models of L2 perception and production. Instead of merely stating in some general way that two sounds are similar, it specifies the precise features and weighting of these features through feature enhancement or muting which make them similar, thus generating testable predictions. It is a model that can adapt to increasing proficiency levels as learners improve and acquire new features or combinations of features, and subsequent adjustments in enhancement relations. The predictions made by the ADM with respect to the languages tested in this dissertation were largely borne out.

Results from this work are in general accordance with the hypothesis that perception underlies production. However, there were some discrepancies between perception and production, particularly with respect to the [f-θ] distinction. It was argued that some of these disparities can be explained by factors that influence production, but not perception, such as visual cues, orthography, and ease of articulation.

It was further argued that the findings from the studies in this thesis provide evidence that non-contrastive features are available to L2 learners, as hypothesized, and that in fact, L2 learners may attend to non-contrastive information not only in the featural specifications of a segment, but also from a neighbouring vowel.

Results also showed improvement in the perception of interdental fricatives by more advanced ESL learners, which is an indication that new features can be acquired and that existing features can be recombined to form new representations.

The model introduced in this work needs to be further developed. Future research in this direction should include thorough phonetic descriptions of L1 segments and their diverse contextual variants. In addition, dialectal differences must be taken into account.

Other cases of differential substitution between dialects of the same language should be investigated. For example, German speakers from Germany are known to substitute [s] for the English interdental, while German speakers from Austria have been reported to use [t] (Peust 1996).

To advance our knowledge of what factors influence learners' perception of the interdental fricatives, synthetic versions of the stimuli could be constructed which manipulate acoustic properties such as spectral mean, length, intensity, and the F2 formant.

Results from the studies presented in this thesis indicate that any perceptual testing should aim towards a more natural type of task, i.e. one which involves identification and classification of L2 sounds rather than simple discrimination. Discrimination tasks may be prone to acoustic processing rather than phonetic or phonological processing, inducing listeners to make a physical comparison between sounds without having recourse to stored representations. It seems that identification and classification tasks give us clearer insight into how listeners process language.

Appendix A: Phonetic and Phonemic Assessments -- Full Inventories

Below are tables showing how intake segments are mapped onto L1 representations. Phonetic assessments are presented first, followed by the phonemic assessments.

Phonetic Assessments

At the top of each table is the L1 inventory. Beside each feature is its weight. Recall that the Major Place features (LABIAL, CORONAL, DORSAL) have an inherent weight of 2. Other features have a weight of 1. Enhanced features have a weight of 2.

Feature enhancements are:

STRIDENT ENHANCES CONTINUANT

MELLOW ENHANCES STOP

ALVEOLAR AND POST-ALVEOLAR ENHANCE STRIDENT

ROUND ENHANCES POST-ALVEOLAR

The intake segments and their features are listed down the left side of the table. The distance between an intake feature and the L1 representation feature is shown at the intersection between rows and columns. For each L1 segment, the total distance for all features is shown in the shaded bar for each intake segment. The darkly shaded L1 segment is the selected substitute.

European French Phonetic

EF L1 Inventory →	p, b	f, v	t̪, d̪	s̪, z̪	ʃ, ʒ	k, g
	STOP(2)	CONT(1)	STOP(2)	CONT(2)	CONT(2)	STOP(2)
	LAB(2)	LAB(2)	COR(2)	COR(2)	COR(2)	DOR(2)
	LIP(1)	DENT(1)	DENT(1)	DENT(1)	POST (2)	VEL(1)
			AP(1)	LAM(1)	AP(1)	
	MELL(1)	MELL(1)	MELL(1)	STRID(1)	STRID(2)	MELL(1)
	UNRND(1)	UNRND(1)	UNRND(1)	UNRND(1)	ROUND(1)	UNRND(1)
Intake ↓						
	p, b	f, v	t̪, d̪	s̪, z̪	ʃ, ʒ	k, g
ɸ, β	3	2	9	9	13	5
CONT(1)	3		3	1	1	3
LAB(2)			4	4	4	
LIP(1)		2	2	2	3	2
MELL(1)				2	3	
UNRND(1)					2	
	p, b	f, v	t̪, d̪	s̪, z̪	ʃ, ʒ	k, g
f, v	5	0	7	7	13	5
CONT(1)	3		3	1	1	3
LAB(2)			4	4	4	
DENT(1)	2				3	2
MELL(1)				2	3	
UNRND(1)					2	
	p, b	f, v	t̪, d̪	s̪, z̪	ʃ, ʒ	k, g
θ, ð	9	4	5	3	11	7
CONT(1)	3		3	1	1	3
COR(2)	4	4				2
DENT(1)	2				3	2
LAM(1)			2		2	
MELL(1)				2	3	
UNRND(1)					2	

	p, b	**f, v**	**t̪, d̪**	**s̺, z̺**	**ʃ, ʒ**	**k, g**
s̺, z̺	12	7	8	0	8	12
CONT(2)	4	1	4			4
COR(2)	4	4				4
DENT(1)	2				3	2
LAM(1)			2		2	
STRID(1)	2	2	2		1	2
UNRND(1)					2	
	p, b	**f, v**	**t̪, d̪**	**s̺, z̺**	**ʃ, ʒ**	**k, g**
t̪, d̪	6	7	2	6	14	6
STOP(2)		3		4	4	
COR(2)	4	4				4
DENT(1)	2				3	2
LAM(1)			2		2	
MELL(1)				2	3	
UNRND(1)					2	
	p, b	**f, v**	**t̪, d̪**	**s̺, z̺**	**ʃ, ʒ**	**k, g**
t, d	6	9	2	10	12	6
STOP(2)		3		4	4	
COR(2)	4	4				4
ALV(2)	2	2	2	2	3	2
AP(1)				2		
MELL(1)				2	3	
UNRND(1)					2	
	p, b	**f, v**	**t̪, d̪**	**s̺, z̺**	**ʃ, ʒ**	**k, g**
t̺s̺, d̺z̺	9	8	5	3	13	9
STOP(1)	1	2	1	3	3	1
COR(2)	4	4				4
DENT(1)	2				3	2
LAM(1)			2		2	
STRID(1)	2	2	2		1	2
UNRND(1)					2	

Québec French Phonetic

In cases where the selected substitute is a contextual allophone, i.e. [t̪s̪,d̪z̪], it is possible that these will not be chosen, especially if the triggering context is not met. In these cases, the next closest segment is predicted to be selected.

QF L1 Inventory →	p, b	f, v	t̪, d̪	s, z	t̪s̪, d̪z̪	ʃ, ʒ	k, g
	STOP(2)	CONT(1)	STOP(2)	CONT(2)	STOP(1)	CONT(2)	STOP(2)
	LAB(2)	LAB(2)	COR(2)	COR(2)	COR(2)	COR(2)	DOR(2)
	LIP(1)	DENT(1)	DENT(1)	ALV(1)	DENT(1)	POST(2)	VEL(1)
			LAM(1)	LAM(1)	LAM(1)	AP(1)	
	MELL(1)	MELL(1)	MELL(1)	STRID(2)	STRID(1)	STRID(2)	MELL(1)
	UNRND(1)	UNRND(1)	UNRND(1)	UNRND(1)	UNRND(1)	ROUND(1)	UNRND(1)
Intake ↓							
	p, b	**f, v**	**t̪, d̪**	**s, z**	**t̪s̪, d̪z̪**	**ʃ, ʒ**	**k, g**
ɸ, β	3	2	9	10	10	13	5
CONT(1)	3		3	1	2	1	3
LAB(2)			4	4	4	4	
LIP(1)		2	2	2	2	3	2
MELL(1)				3	2	3	
UNRND(1)						2	
	p, b	**f, v**	**t̪, d̪**	**s, z**	**t̪s̪, d̪z̪**	**ʃ, ʒ**	**k, g**
f, v	5	0	7	10	8	13	5
CONT(1)	3		3	1	2	1	3
LAB(2)			4	4	4	4	
DENT(1)	2			2		3	2
MELL(1)				3	2	3	
UNRND(1)						2	

	p, b	f, v	t̪, d̪	s, z	t̪s̪, d̪z̪	ʃ, ʒ	k, g
θ, ð	9	4	3	6	4	11	9
CONT(1)	3		3	1	2	1	3
COR(2)	4	4					4
DENT(1)	2			2		3	2
LAM(1)						2	
MELL(1)				3	2	3	
UNRND(1)						2	
	p, b	**f, v**	**t̪, d̪**	**s, z**	**t̪s̪, d̪z̪**	**ʃ, ʒ**	**k, g**
s̪, z̪	12	7	6	3	3	8	10
CONT(2)	4	1	4		3		4
COR(2)	4	4					4
DENT(1)	2			2		3	2
LAM(1)						2	
STRID(1)	2	2	2	1		1	
UNRND(1)						2	
	p, b	**f, v**	**t̪, d̪**	**s, z**	**t̪s̪, d̪z̪**	**ʃ, ʒ**	**k, g**
t̪, d̪	6	7	0	9	3	14	6
STOP(2)		3		4	1	4	
COR(2)	4	4					4
DENT(1)	2			2		3	2
LAM(1)						2	
MELL(1)				3	2	3	
UNRND(1)						2	
	p, b	**f, v**	**t̪, d̪**	**s, z**	**t̪s̪, d̪z̪**	**ʃ, ʒ**	**k, g**
t, d	6	9	4	9	7	12	6
STOP(2)		3		4	1	4	
COR(2)	4	4					4
ALV(2)	2	2	2		2	3	2
AP(1)			2	2	2		
MELL(1)				3	2	3	
UNRND(1)						2	

	p, b	f, v	t̪, d̪	s, z	t̪s̪, d̪z̪	ʃ, ʒ	k, g
t̪s̪, d̪z̪	9	9	3	6	0	11	9
STOP(1)	1	3	1	3		3	1
COR(2)	4	4					4
DENT(1)	2			2		3	2
LAM(1)	2	2	2	1		1	2
STRID(1)						2	
UNRND(1)						2	

Japanese Phonetic

For Japanese, it has been argued that DENTAL and ROUND are absent from the phonetic inventory; therefore these features are not parsed from the intake.[91] This is shown by the crossed out features. In cases where the selected substitute is a contextual allophone, i.e. [ɸ ts,dz ɕ,ʑ tɕ,dʑ], it is possible that these will not be chosen, especially if the triggering context is not met. In these cases, the next closest segment is predicted to be selected.

JA L1 Inventory →	**p, b**	**ɸ**	**t, d**	**s̪, z̪**	**ts, dz**	**ɕ, ʑ**	**tɕ, dʑ**	**k, g**	**h**
	STOP(2)	CONT(1)	STOP(2)	CONT(1)	STOP(1)	CONT(1)	STOP(1)	STOP(2)	CONT(1)
	LAB(2)	LAB(2)	COR(2)	COR(2)	COR(2)	COR(2)	COR(2)	DOR(2)	LAR(2)
	LIP(1)	LIP(1)	ALV(1)	ALV(1)	ALV(1)	POST(1)	POST(1)	VEL(1)	
			AP(1)	LAM(1)	LAM(1)	LAM(1)	LAM(1)		
	MELL(1)	MELL(1)	MELL(1)	MELL(1)	STRID(2)	MELL(1)	STRID(2)	MELL(1)	MELL(1)
	UNRD(1)	UNRD (1)	UNRD(1)	UNRD(1)	UNRD(1)	UNRD(1)	UNRD(1)	UNRD(1)	

[91] In these tables, the affricates of JA are specified as STRIDENT; however, given that these are derived from mellow segments, it is likely that these affricates are MELLOW as well. This would mean that the feature STRIDENT is also absent from the phonetic inventory of JA.

Intake ↓									
	p, b	**ɸ**	**t, d**	**s̪, z̪**	**ts, dz**	**ɕ, ʑ**	**tɕ, dʑ**	**k, g**	**h**
ɸ, β	3	0	9	6	11	6	11	5	4
CONT(1)	3		3		2		2	3	
LAB(2)			4	4	4	4	4		4
LIP(1)			2	2	2	2	2	2	
MELL(1)					3		3		
~~UNRND(1)~~									
	p, b	**ɸ**	**t, d**	**s̪, z̪**	**ts, dz**	**ɕ, ʑ**	**tɕ, dʑ**	**k, g**	**h**
f, v	3	0	7	4	9	4	9	3	4
CONT(1)	3		3		2		2	3	
LAB(2)			4	4	4	4	4		4
~~DENT(1)~~									
MELL(1)					3		3		
~~UNRND(1)~~									
	p, b	**ɸ**	**t, d**	**s̪, z̪**	**ts, dz**	**ɕ, ʑ**	**tɕ, dʑ**	**k, g**	**h**
θ, ð	7	4	5	0	5	0	5	7	4
CONT(1)	3		3		2		2	3	
COR(2)	4	4						4	4
~~DENT(1)~~									
LAM(1)			2						
MELL(1)					3		3		
~~UNRND(1)~~									
	p, b	**ɸ**	**t, d**	**s̪, z̪**	**ts, dz**	**ɕ, ʑ**	**tɕ, dʑ**	**k, g**	**h**
s̪, z̪	10	7	8	3	4	3	4	10	6
CONT(2)	4	1	4	1	3	1	3	4	
COR(2)	4	4						4	4
~~DENT(1)~~									
LAM(1)			2						
STRID(1)	2	2	2	2	1	2	1	2	2
~~UNRND(1)~~									

	p, b	**ɸ**	**t, d**	**s̪, z̪**	**ts, dz**	**ɕ, ʑ**	**tɕ, dʑ**	**k, g**	**h**
t̪, d̪	5	7	2	3	4	3	4	4	7
STOP(2)	1	3		3	1	3	1		3
COR(2)	4	4						4	4
~~DENT(1)~~									
LAM(1)			2						
MELL(1)					3		3		
~~UNRND(1)~~									

	p, b	**ɸ**	**t, d**	**s̪, z̪**	**ts, dz**	**ɕ, ʑ**	**tɕ, dʑ**	**k, g**	**h**
t, d	6	9	0	5	6	7	8	6	7
STOP(2)		3		3	1	3	1		3
COR(2)	4	4						4	4
ALV(2)	2	2				2	2	2	
AP(1)				2	2	2	2		
MELL(1)					3		3		
~~UNRND(1)~~									

	p, b	**ɸ**	**t, d**	**s̪, z̪**	**ts, dz**	**ɕ, ʑ**	**tɕ, dʑ**	**k, g**	**h**
ts̪, dz̪	7	8	5	4	0	4	1	7	8
STOP(1)	1	2	1	2		3		1	2
COR(2)	4	4						4	4
~~DENT(1)~~									
LAM(1)			2						
STRID(1)	2	2	2	2		1	1	2	2
~~UNRND(1)~~									

Phonemic Assessments

Below are the tables for assessments based on phonemic features. It is assumed that phonemic assessments are categorical; hence no distance is measured, nor does feature enhancement apply. Thus, each mismatch between intake feature and L1 feature simply incurs a mark (x). The phonetic features DENTAL and ALVEOLAR are assimilated into the phonological feature ANTERIOR.

EF Phonemic

EF L1 Inventory →	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g
	STOP	CONT	STOP	CONT	CONT	STOP
	LAB	LAB	COR	COR	COR	DOR
			ANT	ANT	POST	
Intake ↓						
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g
ɸ, β	x	0	xx	x	x	xx
CONT	x		x			x
LAB			x	x	x	x
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g
f, v	x	0	xx	x	x	xx
CONT	x		x			x
LAB			x	x	x	x
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g
θ, ð	xx	x	x	0	x	xx
CONT	x		x			x
COR	x	x				x
ANT					x	
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g
s̪, z̪	xx	x	x	0	x	xx
CONT	x		x			x
COR	x	x				x
ANT					x	

	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g
t̪, d̪	x	xx	0	x	xx	x
STOP		x		x	x	
COR	x	x				x
ANT					x	
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**
t, d	x	xx	0	x	xx	x
STOP		x		x	x	
COR	x	x				x
ANT					x	
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**
t̪s̪, d̪z̪	x	xx	0	x	xx	x
STOP		x		x	x	
COR	x	x				x
ANT					x	

QF Phonemic

QF L1 Inventory →	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g
	STOP	CONT	STOP	CONT	CONT	STOP
	LAB	LAB	COR	COR	COR	DOR
			ANT	ANT	POST	
Intake ↓						
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**
ɸ, β	x	0	xx	x	x	xx
CONT	x		x			x
LAB			x	x	x	x

	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g
f, v	x	0	xx	x	x	xx
CONT	x		x			x
LAB			x	x	x	x
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**
θ, ð	xx	x	x	0	x	xx
CONT	x		x			x
COR	x	x				x
ANT					x	
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**
s̪, z̪	xx	x	x	0	x	xx
CONT	x		x			x
COR	x	x				x
ANT					x	
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**
t̪, d̪	x	xx	0	x	xx	x
STOP		x		x	x	
COR	x	x				x
ANT					x	
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**
t, d	x	xx	0	x	xx	x
STOP		x		x	x	
COR	x	x				x
ANT					x	
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**
t̪s̪, d̪z̪	x	xx	0	x	xx	x
STOP		x		x	x	
COR	x	x				x
ANT					x	

JA Phonemic

JA L1 Inventory →	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g	h
	STOP	CONT	STOP	CONT	CONT	STOP	CONT
	LAB	LAB	COR	COR	COR	DOR	LAR
			ANT	ANT	POST		
Intake ↓							
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g	h
ɸ, β	x	0	xx	x	x	xx	x
CONT	x		x			x	
LAB			x	x	x	x	x
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g	h
f, v	x	0	xx	x	x	xx	x
CONT	x		x			x	
LAB			x	x	x	x	x
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g	h
θ, ð	xx	x	x	0	x	xx	x
CONT	x		x			x	
COR	x	x				x	x
ANT					x		
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g	h
s̪, z̪	xx	x	x	0	x	xx	x
CONT	x		x			x	
COR	x	x				x	x
ANT					x		
	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g	h
t̪, d̪	x	xx	0	x	xx	x	xx
STOP		x		x	x		x
COR	x	x				x	x
ANT					x		

	p, b	f, v	t, d	s, z	ʃ, ʒ	k, g	h
t, d	x	xx	0	x	xx	x	xx
STOP		x		x	x		x
COR	x	x				x	x
ANT					x		
	p, b	**f, v**	**t, d**	**s, z**	**ʃ, ʒ**	**k, g**	**h**
t̪s̪, d̪z̪	x	xx	0	x	xx	x	xx
STOP		x		x	x		x
COR	x	x				x	x
ANT					x		

Appendix B: Perceptual Assimilation Model -- Phonetic Assimilations

In the tables below, the intake form for each member of a contrast is associated with the closest L1 category. The auditory distance between intake and L1 category is indicated under "Fit". If each intake member is funneled into the same L1 category, and the fit for each is identical, then this is considered to be a *Single Category Assimilation* (SC). If each intake member is funneled into the same L1 category, but one of the assimilations is greater than the other, then this is considered to be a *Category Goodness Assimilation* (CG). If each intake member is funneled into a separate L1 category, then this is considered to be a *Two Category Assimilation* (TC). Predictions on degree of difficulty for each contrast are given underneath.

EF PAM Assimilations

EF

Intake		**L1 Cat**	**Fit**		**Intake**		**L1 Cat**	**Fit**	
f,v	→	f,v	0	vs.	ɸ,β	→	f,v	2	CG
f,v	→	f,v	0	vs.	θ,ð	→	s̺,z̺	3	TC
t̺,d̺	→	t̪,d̪	2	vs.	θ,ð	→	s̺,z̺	3	TC
s̺,z̺	→	s̺,z̺	0	vs.	θ,ð	→	s̺,z̺	3	CG
t,d	→	t̪,d̪	2	vs.	θ,ð	→	s̺,z̺	3	TC
t,d	→	t̪,d̪	2	vs.	t̺,d̺	→	t̪,d̪	2	SC
t̺,d̺	→	t̪,d̪	2	vs.	t̺s̺,d̺z̺	→	s̺,z̺	3	TC

EF

Predictions:	**Hardest**	t,d	vs.	t̪,d̪	SC
	Medium	f,v	vs.	ɸ,β	CG
		s̪,z̪	vs.	θ,ð	CG
	Easiest	f,v	vs.	θ,ð	TC
		t̪,d̪	vs.	θ,ð	TC
		t,d	vs.	θ,ð	TC
		t̪,d̪	vs.	t̪s̪,d̪z̪	TC

QF PAM Assimilations

QF

Intake		L1 Cat	Fit		Intake		L1 Cat	Fit	
f,v	→	f,v	0	vs.	ɸ,β	→	f,v	2	CG
f,v	→	f,v	0	vs.	θ,ð	→	t̪,d̪	3	TC
t̪,d̪	→	t̪,d̪	0	vs.	θ,ð	→	t̪,d̪	3	CG
s̪,z̪	→	t̪s̪,d̪z̪	3	vs.	θ,ð	→	t̪,d̪	3	TC
t,d	→	t̪,d̪	4	vs.	θ,ð	→	t̪,d̪	3	CG
t,d	→	t̪,d̪	4	vs.	t̪,d̪	→	t̪,d̪	0	CG
t̪,d̪	→	t̪,d̪	0	vs.	t̪s̪,d̪z̪	→	t̪s̪,d̪z̪	0	TC

QF

Predictions:	**Medium**	f,v	vs.	ɸ,β	CG
		t̪,d̪	vs.	θ,ð	CG
		t,d	vs.	θ,ð	CG
		t,d	vs.	t̪,d̪	CG
	Easy	f,v	vs.	θ,ð	TC
		s̪,z̪	vs.	θ,ð	TC
		t̪,d̪	vs.	t̪s̪,d̪z̪	TC

JA PAM Assimilations

JA

Intake		**L1 Cat**	**Fit**		**Intake**		**L1 Cat**	**Fit**	
f,v	→	ɸ,β	0	vs.	ɸ,β	→	ɸ,β	0	SC
f,v	→	ɸ,β	0	vs.	θ,ð	→	**s̠,z̠**	0	TC
t̪,d̪	→	t,d	2	vs.	θ,ð	→	**s̠,z̠**	0	TC
s̪,z̪	→	**s̠,z̠**	3	vs.	θ,ð	→	**s̠,z̠**	0	CG
t,d	→	t,d	0	vs.	θ,ð	→	**s̠,z̠**	0	TC
t,d	→	t,d	0	vs.	t̪,d̪	→	t,d	2	CG
t̪,d̪	→	t,d	2	vs.	t̪s̪,d̪z̪	→	ts,dz	0	TC

JA

Predictions:	**Hardest**	f,v	vs.	ɸ,β	SC
	Medium	t,d	vs.	t̪,d̪	CG
		s̪,z̪	vs.	θ,ð	CG
	Easiest	t̪,d̪	vs.	θ,ð	TC
		t̪,d̪	vs.	t̪s̪,d̪z̪	TC
		f,v	vs.	θ,ð	TC
		t,d	vs.	θ,ð	TC

Appendix C: Russian Phonetic and Phonemic Assessments

Below are tables showing how intake segments are mapped onto Russian L1 representations. Phonetic assessments are presented first, followed by the phonemic assessments. Palatalized and velarized segments are not included due to space restrictions; however, omission of these does not affect the choice of substitute.

Russian Phonetic Assessment

At the top of the table is the RU L1 inventory. Beside each feature is its weight. Recall that the Major Articulator features (LABIAL, CORONAL, DORSAL) have an inherent weight of 2. Other features have a weight of 1. Enhanced features have a weight of 2.

Feature enhancements are:

STRIDENT ENHANCES CONTINUANT

MELLOW ENHANCES STOP

ALVEOLAR AND POST-ALVEOLAR ENHANCE STRIDENT

ROUND ENHANCES POST-ALVEOLAR

The intake segments and their features are listed down the left side of the table. The distance between an intake feature and the L1 representation feature are shown at the intersection between rows and columns. For each L1 segment, the total distance for all features is shown in the shaded bar for each intake segment. The darkly shaded L1 segment is the selected substitute. Russian [s̺̪] and [t̺̪] are underspecified for APICAL/LAMINAL.

Russian Inventory→	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	Stop(2)	Cont(1)	Stop(2)	Cont(2)	Stop(1)	Cont(2)	Stop(1)	Stop(2)	Cont(1)
	Lab(2)	Lab(2)	Cor(2)	Cor(2)	Cor(2)	Cor(2)	Cor(2)	Dor(2)	Dor(2)
	Lip(1)	Dent(1)	Dent(1)	Dent(1)	Alv(1)	Post(2)	Post(2)	Vel(1)	Vel(1)
					Ap(1)	Ap(1)	Lam(1)		
	Mell(1)	Mell(1)	Mell(1)	Strid(1)	Strid(2)	Strid(2)	Strid(2)	Mell(1)	Mell(1)
	UnRd(1)	UnRd(1)	UnRd(1)	UnRd(1)	UnRd(1)	Rd(1)	Rd(1)	UnRd(1)	UnRd(1)
Intake ↓	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
ɸ, β	3	2	9	8	11	13	14	9	6
Cont(1)	3		3		2	1	2	3	
Lab(2)			4	4	4	4	4	4	4
Lip(1)		2	2	2	2	3	3	2	2
Mell(1)				2	3	3	3		
UnRd(1)						2	2		
	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
f, v	5	0	7	6	11	13	14	11	6
Cont(1)	3		3		2	1	2	3	
Lab(2)			4	4	4	4	4	4	2
Dent(1)	2				2	3	3	2	2
Mell(1)				2	3	3	3	2	2
UnRd(1)						2	2		

	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
θ,ð	9	4	3	3	9	11	10	9	6
Cont(1)	3		3	1	2	1	2	3	
Cor(2)	4	4						4	4
Dent(1)	2				2	3	3	2	2
Lam(1)					2	2			
Mell(1)				2	3	3	3		
UnRd(1)						2	2		
	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
s̪,z̪	12	7	6	0	6	6	9	12	9
Cont(2)	4	1	4		3		3	4	1
Cor(2)	4	4						4	4
Dent(1)	2				2	3	3	2	2
Lam(1)									
Strid(1)	2	2	2		1	1	1	2	2
UnRd(1)						2	2		
	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
t̪,d̪	7	6	0	6	8	14	9	6	8
Stop(2)	1	3		3	1	4	1		3
Cor(2)	4	3		1				4	3
Dent(1)	2				2	3	3	2	2
Lam(1)					2	2			
Mell(1)				2	3	3	3		
UnRd(1)						2	2		

	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
t̪,d̪	7	6	0	6	6	12	11	6	8
Stop(2)	1	3		3	1	4	1		3
Cor(2)	4	3		1				4	3
Dent(1)	2				2	3	3	2	2
Ap(1)							2		
Mell(1)				2	3	3	3		
UnRd(1)						2	2		

	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
t,d	7	8	2	8	4	12	11	6	8
Stop(2)	1	3		3	1	4	1		3
Cor(2)	4	3		1				4	3
Alv(1)	2	2	2	2		3	3	2	2
Ap(1)							2		
Mell(1)				2	3	3	3		
UnRd(1)						2	2		

	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
ts̪,dz̪	9	8	3	3	5	11	6	9	10
Stop(1)	1	2	1	3		3		1	2
Cor(2)	4	4						4	4
Dent(1)	2				2	3	3	2	2
Lam(1)					2	2			
Strid(1)	2	2	2		1	1	1	2	2
UnRd(1)						2	2		

	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
ʃ, ʒ	13	13	9	9	5	0	2	13	13
Cont(2)	1	2	1	2				1	2
Cor(2)	4	3		1				4	3
Post(2)	3	3	3	3	3			3	3
Ap(1)							2		
Strid(2)	3	3	3	1				3	3
Round(1)	2	2	2	2	2			2	2
	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
tʃ, dʒ	13	13	9	9	5	4	2	13	13
Stop(1)	1	2	1	2		3		1	2
Cor(2)	4	3		1		1		4	3
Post(2)	3	3	3	3	3			3	3
Ap(1)							2		
Strid(2)	3	3	3	1				3	3
Round(1)	2	2	2	2	2			2	2
	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
k, g	6	8	6	10	10	15	13	0	4
Stop(2)		3		3	1	3	1		3
Dor(2)	4	3	4	3	4	4	4		1
Vel(1)	2	2	2	2	2	3	3		
Mell(1)				2	3	3	3		
UnRd(1)						2	2		

	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
x, ɣ	12	8	12	6	8	8	11	8	4
Cont(2)	4	1	4	1	3		3	4	1
Dor(2)	3	2	3	2	3	3	3	1	
Vel(1)	2	2	2	2	2	3	3		
Strid(2)	3	3	3	1				3	3
UnRd(1)						2	2		

Russian Phonemic Assessment

Below is the table for an assessment based on phonemic features. It is assumed that phonemic assessments are categorical; hence no distance is measured, nor does feature enhancement apply. Thus, each mismatch between intake feature and L1 feature simply incurs a mark (x). The phonetic features DENTAL and ALVEOLAR are assimilated into the phonological feature ANTERIOR.

RU L1 Inventory →	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	Stop	Cont	Stop	Cont	Stop	Cont	Stop	Stop	Cont
	Lab	Lab	Cor	Cor	Cor	Cor	Cor	Dor	Dor
			Ant	Ant	Ant	Post	Post		
Intake ↓	Mell	Mell	Mell	Strid	Strid	Strid	Strid	Mell	Mell

ɸ, β	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	x		xx	xx	xxx	xx	xxx	xx	x
Cont	x		x		x		x	x	
Lab			x	x	x	x	x	x	x
Mell				x	x	x	x		
f, v	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	x		xx	xx	xxx	xx	xxx	xx	x
Cont	x		x		x		x	x	
Lab			x	x	x	x	x	x	x
Mell				x	x	x	x		
θ, ð	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	xx	x	x	x	xx	xx	xxx	xx	x
Cont	x		x		x		x	x	
Cor	x	x						x	x
Ant						x	x		
Mell				x	x	x	x		
s̪, z̪	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	xxx	xx	xx		x	x	xx	xxx	xx
Cont	x		x		x		x	x	
Cor	x	x						x	x
Ant						x	x		
Strid	x	x	x					x	x
t̪, d̪	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	x	xx		xx	x	xxx	xx	x	xx
Stop		x		x		x			x
Cor	x	x						x	x
Ant						x	x		
Mell				x	x	x	x		

t̪,d̪	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
	x	xx		xx	x	xxx	xx	x	xx
Stop		x		x		x			x
Cor	x	x						x	x
Ant						x	x		
Mell				x	x	x	x		
t,d	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
	x	xx		xx	x	xxx	xx	x	xx
Stop		x		x		x			x
Cor	x	x						x	x
Ant						x	x		
Mell				x	x	x	x		
s,z	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
	xxx	xx	xx		x	x	xx	xxx	xx
Cont	x		x		x		x	x	
Cor	x	x						x	x
Ant						x	x		
Strid	x	x	x					x	x
t̪s̪,d̪z̪	p,b	f,v	t̪,d̪	s̪,z̪	ts	ʃ,ʒ	tʃ	k,g	x
	xx	xxx	x	x		xx	x	xx	xxx
Stop		x		x		x			x
Cor	x	x						x	x
Ant						x	x		
Strid	x	x	x					x	x

tʃ, dʒ	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	xx	xxx	xx	xx	x	x		xx	xxx
Stop		x		x		x			x
Cor	x	x						x	x
Post			x	x	x				
Strid	x	x	x					x	x
k, g	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	x	xx	x	xxx	xx	xxx	xx		x
Stop		x		x		x			x
Dor	x	x	x	x	x	x	x		
Mell				x	x	x	x		
x, ɣ	p, b	f, v	t̪, d̪	s̪, z̪	ts	ʃ, ʒ	tʃ	k, g	x
	xxx	xx	xxx	x	xx	x	xx	xx	x
Cont	x		x		x		x	x	
Dor	x	x	x	x	x	x	x		
Strid	x	x	x					x	x

Appendix D: Perceptual Assimilation Model -- Russian Phonetic Assimilations

In the tables below, the intake form for each member of a contrast is associated with the closest L1 category. The auditory distance between intake and L1 category is indicated under "Fit". If each intake member is funnelled into the same L1 category, and the fit for each is identical, then this is considered to be a *Single Category Assimilation* (SC). If each intake member is funnelled into the same L1 category, but one of the assimilations is greater than the other, then this is considered to be a *Category Goodness Assimilation* (CG). If each intake member is funnelled into a separate L1 category, then this is considered to be a *Two Category Assimilation* (TC).

Intake			**Fit**		**Intake**		**L1 Cat**	**Fit**	
ɸ	→	f	2	vs.	θ	→	t̪ , s̪	3	TC
f	→	f	0	vs.	θ	→	t̪ , s̪	3	TC
t̪	→	t̪	0	vs.	θ	→	t̪ , s̪	3	CG/TC
t	→	t̪	2	vs.	θ	→	t̪ , s̪	3	CG/TC
s̪	→	s̪	0	vs.	θ	→	t̪ , s̪	3	CG/TC
s	→	s̪	5	vs.	θ	→	t̪ , s̪	3	CG/TC
t̪s̪	→	s̪	2	vs.	θ	→	t̪ , s̪	3	SC/TC

Appendix E: Stimuli for Picture Identification Task

Here is an example block of the list of the picture triads presented during the Picture Identification experiment. The position of the pictures and the aural target word presented were counterbalanced across test blocks. Under the Position column in this table below, O = Onset and C = Coda.

Picture Triads			Consonant/Vowel	Position
thin	fin	boot	f/HF	O
theses	bath	feces	f/HF	O
bath	Thor	four	f/NH	O
laid	fought	thought	f/NH	O
think	sink	death	s/HF	O
theorem	bat	serum	s/HF	O
thought	lathe	sought	s/NH	O
thaw	saw	bash	s/NH	O
sigh	thigh	bathe	s/NH	O
sore	fin	Thor	s/NH	O
mouth	sum	thumb	s/NH	O
booth	thin	tin	t/HF	O
taught	thought	lather	t/NH	O
bays	thigh	tie	t/NH	O
Thor	tore	fought	t/NH	O
torn	thorn	four	t/NH	O
tug	thug	mouse	t/NH	O
shin	deaf	thin	ʃ/HF	O
bass	thaw	Shah	ʃ/NH	O
Fred	Thor	shore	ʃ/NH	O
shorn	thorn	ladder	ʃ/NH	O
mass	shot	thought	ʃ/NH	O
shy	beige	thigh	ʃ/NH	O
wreath	reef	sought	HF/f	C
death	deaf	shorn	NH/f	C

bath	bass	serum	NH/s	C
math	mass	shot	NH/s	C
moth	moss	shy	NH/s	C
mouth	mouse	sigh	NH/s	C
booth	boot	shore	HB/t	C
bath	bat	Shah	NH/t	C
math	mat	shred	NH/t	C
wrath	rat	sink	NH/t	C
bath	bash	shin	NH/ʃ	C
wrath	rash	sore	NH/ʃ	C
teethe	tease	thaw	HF/z	C
bathe	bays	sum	NH/z	C
lathe	laid	tease	NH/d	C
bathe	beige	taught	NH/ʒ	C

APPENDIX F: STIMULI FOR WORD PRODUCTION TASK

PAIR		Status: Test (T) Filler (F)	Position: Onset (O) Coda (C)
thank	thing	T	O
thanked	things	T	O
thanks	thirsty	T	O
think	thin	T	O
thinks	thick	T	O
thirst	thank	T	O
death	mouth	T	C
math	bath	T	C
south	north	T	C
teeth	faith	T	C
tooth	with	T	C
truth	path	T	C
arm	art	F	
baby	bottle	F	
bed	dream	F	
cap	cat	F	
chair	cheese	F	
computer	internet	F	
dress	dry	F	
ear	eat	F	
elephant	tiger	F	
first	fat	F	
hair	hat	F	
key	call	F	
leg	light	F	
me	my	F	

meet	mouse	F	
neat	knee	F	
new	no	F	
nose	not	F	
pie	pick	F	
pill	sick	F	
pot	knife	F	
share	shoe	F	
sing	see	F	
snow	rain	F	
soap	water	F	
table	kitchen	F	
tap	tail	F	
tea	train	F	
television	radio	F	
top	tap	F	
way	where	F	
window	glass	F	
year	yes	F	

REFERENCES

Akamatsu, T. (1997). *Japanese Phonetics: Theory and Practice.* Munich: Lincom Europa.

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology , 5*, 183-208.

Archangeli, D., & Pulleyblank, D. (1989). Yoruba vowel harmony. *Linguistic Inquiry , 20*, 173-217.

Avery, P. (1996). The Representation of Voicing Contrasts. PhD Dissertation. University of Toronto.

Avery, P., & Rice, K. (1989). Segmental structure and coronal underspecification. *Phonology , 6*, 179-200.

Babel, M., & McGuire, G. (2010). *A cross-modal account for synchronic and diachronic patterns of /f/ and /θ/.* Laboratory Report, University of Southern California at Santa Cruz, Linguistics Research Centre.

Beddor, P., & Gottfried, T. (1995). Methodological issues in cross-language speech perception research with adults. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-Language Speech Research* (pp. 207-232). Timonium, MD: York Press.

Behren, S., & Blumstein, S. (1988). On the role of amplitude of the fricative noise in the perception of place of articulation in voiceless fricative consonants. *Journal of the Acoustical Society of America , 84,* 295-298.

Benki, J. (1998). Evidence for Phonological Categories from Speech Perception. Doctoral Dissertation, University of Massachuetts Amherst.

Berger, M. (1951). The American English Pronunciation of Russian Immigrants. PhD Dissertation, Columbia University.

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. (pp. 171-204). Timonium, MD: York Press.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics , 20*, 305-331.

Best, C. T., McRoberts, G., & Sithole, N. (1988). "Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance , 14*, 345–360.

Bialystok, E., & Cummins, J. (1991). Language, cognition, and education of bilingual children. In E. Bialysok (Ed.). London: Cambridge University Press.

Boersma, P. (1998). *Functional Phonology*. The Hague: Holland Academic Graphics.

Boersma, P., & Hamann, S. (2007). The evolution of auditory contrast. (909). Rutgers Optimality Archive. Retrieved from Rutgers Optimality Archive.

Brannen, K. (1996). Erroneous Articulatory Routines: A Performance-based Model of Speech Production. Master's Thesis, University of Ottawa.

Brannen, K. (1998). *The acquisition of interdentals in L2.* University of Ottawa, MOT Phonology Workshop.

Brannen, K. (2002). The role of perception in differential substitution. *Canadian Journal of Linguistics , 47* (1), 1-46.

Brown, C. (1993). The role of the L1 grammar in the L2 acquisition of segmental structure. *McGill Working Papers in Linguistics , 9* (1&2), 180-210.

Brown, C. (1997). Acquisition of Segmental Structure: Consequences for Speech Perception and Second Language Acquisition. PhD Dissertation. McGill University.

Brown, C. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research , 14*, 136-193.

Calabrese, A. (1988). Towards a Theory of Phonological Alphabets. PhD Dissertation, Massachusetts Institute of Technology.

Caramazza, A., & Yeni-Komshian, G. (1974). Voice onset time in two French dialects. *Journal of Phonetics , 2*, 239-245.

Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytical Studies.* New York: Cambridge University Press.

Charbonneau, R., & Jacques, B. (1972). [ts] et [dz] en français canadien. In A. Valdmen (Ed.), *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre.* The Hague: Mouton.

Cho, H., & Giavazzi, M. (2009). Perception of voicing in fricatives. *Proceedings of the 18th International Congress of Linguists.* John Benjamins.

Clements, G. N. (1992). Phonological primes: Features or gestures? *Phonetica , 49*, 181-193.

Clements, N. (2005). The Role of Features in Phonological Inventories. In E. Raimy, & C. Cairns (Eds.), *Contemporary Views on Architecture and Representations in Phonological Theory.* Cambridge, MA: MIT Press.

Clements, N. (2006). The feature [strident]. Manuscript. Laboratoire de phonétique et phonologie, CNRS/U. Paris.

Cohen, J.D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers , 25*, 257-271.

Côté, M.-H. (2005). Phonologie française. Manuscript, University of Ottawa.

Crothers, J., Lorentz, J., Sherman, D., & Vihman, M. (1979). *Handbook of Phonological Data from a Sample of the World's Languages.* Stanford.

Curtin, S., Goad, H., & Pater, J. V. (1998). Phonological transfer and levels of representation: The perceptual acquisition of Thai voice and aspiration by English and French speakers. *Second Language Research , 14* (4), 389-405.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *Journal of the Acoustical Society of America , 116* (6), 3668-3678.

Dart, S. (1991). Articulatory and Acoustic Properties of Apical and Laminal Articulations. *UCLA Working Papers in Phonetics , 79*, 1-155.

Dauer, R. (1993). *Accurate English: A Complete Course in Pronunciation.* Englewood Cliffs, NJ: Prentice Hall Regents.

Delattre, P. (1951). *Principes de phonétique française.* Middlebury, VA.

Dresher, B. E. (2009). *The Contrastive Hierarchy in Phonology.* New York: Cambridge University Press.

Eckman, F. (1977). Markedness and the contrastive analysis hypothesis. *Language Learning , 27*, 315-330.

Escudero, P. (2005). Linguistic Perception and Second Language Acquisition: Explaining the Attainment of Optimal Phonological Categorization. PhD Dissertation. University of Utrecht.

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.

Flege, J. E. (1989). Chinese subjects' perception of the word-final English /t/-/d/contrast: Performance before and after training. *Journal of the Acoustical Society of America , 86* (5), 1684-1697.

Flege, J. E. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Heubner, & C. Ferguson (Eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theory* (pp. 249-289). Philadelphia: John Benjamins.

Flege, J.E. (1995) Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and Linguistic Experience* (pp.233-277). Timonium, MD: York Press.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer, & N. Schiller (Eds.), *Phonetic and Phonology in Language Comprehension and Production* (pp. 319-355). Berlin: Mouton de Gruyter.

Flege, J. E., & Eefting, W. (1987). Cross-Language Switching in Stop Consonant Perception and Production by Dutch Speakers of English. *Speech Communication, 6*, 185-202.

Flege, J. E., MacKay, I. R., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America , 106* (5), 2973-2987.

Flemming, E. (1995/2002). *Auditory Representations in Phonology*. Routledge.

Fowler, C. (1994). Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation. *Perception & Psychophysics , 55*, 597-610.

Fuchs, S., & Toda, M. (2010). Do differences in male versus female /s/ reflect biological or sociophonetic factors? In S. Fuchs, & M. Toda (Eds.), *Turbulent Sounds: An Interdisciplinary Guide* (pp. 293-314). Berlin: Walter de Gruyter.

Gardner, R., & Lambert, W. E. (1972). *Attitudes and Motivation in Second-Language Learning*. Rowley, MA: Newbury House.

Gatbonton, E. (1978). Patterned phonetic variability in second language speech: A gradual diffusion model. *Canadian Modern Language Revier , 34*, 335-347.

Gerrits, E. (2001). The Categorisation of Speech Sounds by Adults and Children. Doctoral Dissertation, University of Utrecht.

Gnanadesikan, A. E. (1993). The feature geometry of coronal subplaces. *University of Massachusetts Occasional Papers in Linguistics , 16*, 27-67.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "l" and "r". *Neuropsychologia , 9*, 317-323.

Grenon, I. (2010). The Bi-Level Input Processing Model of First and Second Language Perception. PhD Dissertation. University of Victoria.

Guion, S. G., Flege, J. E., Akahane, R.-Y., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustic Society of America , 107* (5), 2711-2724.

Hall, T. A., & Hamann, S. (2003). Towards a typology of stop assibilation. *ZAS Papers in Linguistics , 32*, 111-136.

Hall, T. A., & Zygis, M. (2010). An overview of the phonology of obstruents. In S. Fuchs, M. Toda, & M. Zygis (Eds.), *Turbulent Sounds: An Interdisciplinary Guide* (pp. 1-36). Berlin: De Gruyter.

Hamilton, P. J. (1996). Phonetic Constraints and Markedness in the Phonotactics of Australian Aboriginal Languages. *Toronto Working Papers in Linguistics* , PhD Dissertation.

Hamilton, W. S. (1980). *Introduction to Russian Phonology and Word Structure.* Columbus, OH: Slavica.

Hancin-Bhatt, B. (1994a). *Phonological Transfer in Second Language Perception and Production.* PhD Dissertation. University of Illinois at Urbana-Champaign.

Hancin-Bhatt, B. (1994b). Segment transfer: A consequence of a dynamic system. *Second Language Research , 10* (3), 241-269.

Hannahs, S. (2007). French phonology and L2 acquisition. In D. Ayoun (Ed.), *French Applied Linguistics* (pp. 50-74). Amsterdam/Philadelphia: John Benjamins.

Hanulíková, A., & Weber, A. (2010). Production of English interdental fricatives by Dutch, German, and English speakers. *New Sounds 2010: Sixth International Symposium on the Acquisition of Second Language Speech* (pp. 173-178). Poznan: Adam Mickiewicz University.

Hardcastle, W. (1976). *The Physiology of Speech Production.* New York: Academic Press.

Hardison, D. M. (1999). Bimodal Speech Perception by Native and Nonnative Speakers of English: Factors Influencing the McGurk Effect. In J. Leather (Ed.), *Phonological Issues in Language Learning* (pp. 213-283). Oxford: Blackwell.

Hazan, V., Sennema, A., Faulkner, A., Ortega-Llebaria, M., Iba, M., & Chung, H. (2006) The use of visual cues in the perception of non-native consonant contrasts. *Journal of the Acoustical Society of America , 119,* 1740-1751.

Hedrick, M., & Ohde, R. (1993). Effect of relative amplitude and formant transitions on perception of place of articulation by adult listeners with cochlear implants. *Journal of the Acoustical Society of America , 94*, 2005-2027.

Heinz, J., & Stevens, K. N. (1961). On the properties of voiceless fricative consonants. *Journal of the Acoustical Society of America , 33,* 589-596.

Hughes, G., & Halle, M. (1956). Spectral properties of voiceless fricative consonants. *Journal of the Acoustical Society of America , 28,* 303-310.

Hume, E., & Johnson, K. (2001). *The Role of Speech Perception in Phonology.* San Diego: Academic Press.

Hurch, B. (1986). Elementi di una fonologia contrastiva naturale. *Lingua e Stile , 21*, 3-26.

Ingram, J. C. L. & Park, S.-G. (1998) Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by

Japanese and Korean listeners. *Journal of the Acoustical Society of America , 103(2),* 1161-1174.

Ioup, G., & Weinberger, S. (1987). *Interlanguage Phonology.* Cambridge, MA: Newbury House.

Iverson, G. A. (1995). Aspiration and Laryngeal Representation in. *Phonology* (12), 369-396.

Iverson, P., Ekanayake, D., Hamann, S., Sennema, A., & Evans, B. G. (2008). Category and Perceptual Interference in Second-Language Phoneme Learning: An Examination of English /w/-/v/ Learning by Sinhala, German, and Dutch Speakers. *Journal of Experimental Psychology*, *34* (5).

Jakobson, R. (1941/1968). *Kindersprache, Aphasie, und Allgemeine Lautgesetze.* (A. Keiler, Trans.) Uppsala.

Jakobson, R., Fant, G., & Halle, M. (1963/1969). *Preliminaries to speech analysis.* Cambridge: MIT Press.

James, A. R. (1994). Second language acquisition: Phonology. In R. E. Asher, J. M. Y. Simpson (Eds.), *The Encyclopedia of Language and Linguistics* (pp. 3728-3732). Oxford: Pergamon Press.

James, A. R. (1996). Second language phonology. In P. Jordens, & J. Lalleman (Eds.), *Investigating Second Language Acquisition* (pp. 293-319). Berlin: Mouton de Gruyter.

Jamieson, D. G., & Morosan, D. E. (1986). Training non-native speech contrasts in adults: Acquisition of the English /th/-/dh/ contrast by francophones. *Perception & Psychophysics* , 205-215.

Johnson, K. (2003). *Acoustic and Auditory Phonetics* (2nd ed.). Oxford: Blackwell.

Johnson, K., & Babel, M. (2007). *Perception of fricatives by Dutch and English speakers.* Berkeley: University of Calgary.

Jones, D., & Ward, D. (1969). *The Phonetics of Russian.* Cambridge: Cambridge University Press.

Jongman, A. (1989). Duration of frication noise required for identification of English fricatives. *Journal of the Acoustical Society of America , 85* (4), 1718-1725.

Jongman, A., Wayland, R., & Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America , 108* (3), 1252-1263.

Keating, P. A. (1990). Phonetic representations in a generative grammar. *Journal of Phonetics , 18,* 321-334.

Kenstowicz, M. (1994). *Phonology in generative grammar.* Oxford: Blackwell.

Kenstowicz, M. (2003). Salience and Similarity in Loanword Adaptation: A Case Study from Fijian. *Language Sciences* .

Keyser, S. J., & Stevens, K. N. (2001). Enhancement revisited. In M. Kenstowicz (Ed.), *Ken Hale: A Life in Language* (pp. 271-291). Cambridge, MA: MIT Press.

Keyser, S. J., & Stevens, K. N. (2006). Enhancement and Overlap in the Speech Chain. *Project MUSE* , 33-63.

Kim, H. (2001). A phonetically based account of phonological stop assibilation. *Phonology , 18*, 81-108.

Kiparsky, P. (1985). Some consequences of lexical phonology. *Phonology , 2,* 85-138.

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications.* London: Longman.

Kučera, H., & Francis, N. (1967). *Computational analysis of present-day American English.* Providence: Brown University Press.

LaCharité, D., & Prévost, P. (1999). Le rôle de la langue maternelle et de l'enseignement dans l'acquisition des segments de l'anglais langue second par des apprenants francophones. *Langues et linguistique , 25*, 81-109.

Ladefoged, P. (2001). *Vowels and Consonants: An Introduction to the Sounds of Languages.* Maldon, Mass: Blackwell.

Lahiri, A., & Evers, V. (1991). Palatalization and coronality. In C. Paradis, & J.-F. Prunet (Eds.), *The Special Status of Coronals* (pp. 79-100). San Diego: Academic Press.

Lamendella, J. (1977). General principles of neurofunctional organization and their manifestation in primary and non-primary acquisition. *Language Learning , 27*, 155-196.

Lenneberg, E. (1967). *Biological Foundations of Language.* New York: Wiley.

Li, F. (2008). The Phonetic Development of Voiceless Sibilant Fricatives in English, Japanese, and Mandarin Chinese. Doctoral Dissertation, Ohio State University.

Lieberman, M. (2010). The acquisition of phonemic liquid contrasts in L2 English. Evaluation Paper. McGill University.

Lindman, H. (1974). *Analysis of Variance in Complex Experimental Designs.* San Francisco: W.H. Freeman & Co.

Lombardi, L. (2003). Second language data and constraints on Manner: Explaining substitutions for the English interdentals. *Second Language Research , 19* (3), 225-250.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition. , 12*, 251-285.

Maddieson, I. (1984). *Patterns of Sounds.* Cambridge: Cambridge University Press.

Mah, J. (2011). 'Segmental Representations in Interlanguage Grammars: The Case of Francophones and English /h/. PhD Dissertation. McGill University.

Major, F. (1987). A model for interlanguage phonology. In G. Ioup, & S. Weinberger (Eds.), *Interlanguage Phonology: The Acquisition of a Second*

Language Sound System (pp. 101-125). New York: Newbury House/Harper & Row.

Major, F. (1995). Native and nonnative phonological representations. *International Review of Applied Linguistics in Language Learning , 33*, 143-161.

Major, R. (2001). *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Mahwah, NJ: Lawrence Erlbaum.

Mann, V., & Repp, B. (1980). Influence of vocalic context on the perception of the [ʃ]-[s] distinction. *Perception & Psychophysics, 28* (3), 213-228.

McCarthy, J. (1988). Feature geometry and dependency: A review. *Phonetica, 43,* 84-108.

McGurk, H. & MacDonald, J. (1976) Hearing lips and seeing voices. *Nature*, 264, 746-748.

Mester, R. A., & Itô, J. (1989). Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language , 65*, 258-293.

Michaels, D. (1974). Sound replacements and phonological systems. *Linguistics , 126*, 69-81.

Miller, G. A., & Nicely, P. (1955). An analysis of perceptual confusion among some English consonants. *Journal of the Acoustical Society of America, 27*, 338-346.

Miller-Ockhuizen, A., & Zec, D. (2003). Phonetics and Phonology of Contrastive Palatal Affricates. *Working Papers of the Cornell Phonetics Laboratory, 15*, 130-193.

Muñoz-Sánchez, A. (2003). The Effect of Phonological Status on the Acquisition of New Contrasts: Evidence from Spanish and Japanese L2 Learners of English. PhD Dissertation. University of California, San Diego.

Nartey, J. (1982). On fricative phones and phonemes. *UCLA Working Papers in Linguistics, 55*.

Nathan, G. (1989). Natural Phonology Predicts L2 Phonology. Manuscript, Southern Illinois University at Carbondale and Haskins Laboratories.

Nemser, W. (1971). An Experimental Study of Phonological Interference in the English of Hungarians. Bloomington, IN: Indiana University.

Neufeld, G. G. (1988). Phonological asymmetry in second language learning and performance. *Language Learning , 38* (4), 531-560.

Ohala, J., & Solé, M. J. (2010). Turbulence and Phonology. In S. Fuchs, M. Toda, & M. Zygis (Eds.), *Turbulent Sounds: An Interdisciplanary Guide* (pp. 37-102). Berlin: De Gruyter.

Paradis, C., & LaCharité, D. (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics , 33*, 379-430.

Paradis, C., & Prunet, J.-F. (1991). *The Special Status of Coronals: Internal and External Evidence.* San Diego: Academic Press.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism.* Amsterdam: John Benjamins.

Peperkamp, S., & Dupoux, E. (2003). Reinterpreing loanword adaptations: The role of perception. *15th ICPhS*, (pp. 367-370). Barcelona.

Peust, C. (1996). Sum: th-substitution. *The Linguist List , 7.1108* .

Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle linguistique de Prague , 4*, 79-96.

Prince, A., & Smolensky, P. (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar.* New Brunswick, NJ: Rutgers University Center for Cognitive Science.

Rice, K., & Avery, P. (1989). On the interaction between sonority and voicing. *Toronto Working Papers in Linguistics , 10*, 65-92.

Rice, K., & Avery, P. (1995). Variability in a deterministic model of language acquisition: A theory of segmental elaboration. In J. Archibald

(Ed.), *Phonological Acquisition and Phonological Theory* (pp. 23-42). Hillsdale, NJ: Lawrence Erlbaum.

Robinett, B., & Schachter, J. (Eds.). (1983). *Second Language Learning.* University of Michigan.

Rochet, B. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech Perception and Linguistic Experience.* Timonium, MD: York Press.

Roy, M.-J. (1992). Le rôle des contraintes phonologiques dans l'adaptation d'emprunts anglais en français québécois. Master's Thesis, Laval University.

Rubach, J. (1994). Affricates as Strident Stops in Polish. *Linguistic Inquiry , 25* (1), 119-144.

Sagey, E. (1986/1991). *The Representation of Features and Relations in Nonlinear Phonology.* New York: Garland Press.

Schmidt, A. M. (1996). Cross-language identification of consonants Part I: Korean perception of English. *Journal of the Acoustical Society of America, 99*, 3201-3211.

Schumann, H., Holroyd, J., Campbell, R. N., & Ward, F. A. (1978). Improvement of foreign language pronunciation under hypnosis: A preliminary study. *Language Learning , 28*, 143-148.

Schwartz, B. & Sprouse, R. (1994). Word order and nominative case in nonnative language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra & B. Schwartz (Eds.), *Language Acquisition Studies in Generative Grammar*, (pp.317-368). Amsterdam: John Benjamins.

Scovel, T. (1988). *A Time to Speak.* Cambridge, MA: Newbury House.

Shadle, C. H. (1985). The Acoustics of Fricative Consonants. Doctoral Dissertation, Massachusetts Institute of Technology.

Shadle, C. H., Mair, S., & Carter, J. (1996). Acoustic characteristics of the front fricatives [f,v,θ,ð]. *Proceedings of the First ESCA Workshop on Speech Production Modeling* (pp. 193-196). Autrans: European Speech Communication Association.

Sharwood-Smith, M. (2004). In two minds about grammar: On the interaction of linguistic and metalinguistic knowledge in performance. *Transactions of the Philological Society , 102* (3), 255-280.

Sheldon, A. & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics , 3*, 243-261.

Silverman, D. (1995). Phasing and Recoverability. PhD Dissertation. University of California Los Angeles.

Solé, M.-J. (2003). Aerodynamic characteristics of onset and coda fricatives. *15th ICPhS* , (pp. 1-4). Barcelona.

Soli, S. D. (1981). Second formants in fricatives: Acoustic consequences of fricative-vowel coarticulation. *Journal of the Acoustical Society of America, 70* (4), 976-984.

Steriade, D. (1987). On class nodes. Manuscript, Massachusetts Institute of Technology.

Steriade, D. (1995). Underspecification and Markedness. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 114-174). Cambridge/Oxford: Blackwell.

Steriade, D. (2001). Directional asymmetries in place assimilation. In E. Hume, & K. Johnson (Eds.), *The Role of Speech Perception in Phonology* (pp. 219-250). San Diego: Academic Press.

Stevens, K. N. (1998). *Acoustic Phonetics.* Cambridge, MA: MIT Press.

Stevens, K. N. (1985) Evidence for the role of acoustic boundaries in the perception of speech sounds. In V. Fromkin (Ed.), *Phonetic Linguistics* (pp. 243-255). New York: Academic.

Stevens, K. N., & Keyser, S. J. (1989). Primary features and their enhancement in consonants. *Language , 65*, 81-106.

Stevens, K. N., Keyser, S. J., & Kawasaki, H. (1986). Toward a Phonetic and Phonological Theory of Redundant Features. In J. S. Perkell, & D. H. Klatt (Eds.), *Invariance and Variability in Speech Processes* (pp. 426-463). Hillsdale, New Jersey: Lawrence Erlbaum.

Strevens, P. (1960). Spectra of fricative noise in human speech. *Language and Speech , 3*, 32-49.

Sussman, H., & Shore, J. (1996). Locus equations as phonetic descriptors of consonantal place of articulation. *Perception & Psychophysics , 58*, 936-946.

Tabain, M. (1998). Non-sibilant fricatives in English: Spectral information above 10kHz. *Phonetica , 55*, 107-130.

Teasdale, A. M. (1997). On the differential substitution of English [θ]: A phonetic approach. *Calgary Working Papers in Linguistics , 19*, 71-85.

Toda, M., & Honda, K. (2003). *An mri-based cross-linguistic study of sibilant fricatives.* 6th International Seminar on Speech Production, Manly, Australia.

Trubetzkoy, N. (1939/1969). *Principles of Phonology* (2nd ed.). (C. A. Baltaxe, Trans.) Berkeley/Los Angeles: University of California Press.

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

Vance, T. (2008). *The Sounds of Japanese.* Cambridge: Cambridge University Press.

van Hessen, A. J. & Schouten, M. E. H. (1999) Categorical perception as a function of stimulus quality. *Phonetica, 56*, 56–72.

Venneman, T. (1988). *Preference laws for syllable structure and the explanation of sound change.* Berlin: Mouton de Gruyter.

Vihman, M. (1996). *Phonological Development.* Oxford: Blackwell.

Wagner, A., Emestus, M., & Cutler, A. (2006). Formant transitions in fricative identification: The role of native fricative inventory. *Journal of the Acoustical Society of America , 120* (4), 2267-2277.

Weinberger, S. H. (1988). Theoretical Foundations of Second Language Phonology. *PhD Dissertation* . University of Washington.

Weinreich, U. (1953). *Languages in Contact: Findings and Problems.* New York: Linguistic Circle of New York.

Wenk, B. J. (1979). Articulatory setting and de-fossilization. *Interlanguage Studies Bulletin , 4*, 202-220.

Werker, J., & Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics , 37* (1), 35-44.

Werker, J. F., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development , 7* (1), 49-63.

Wilde, L. F. (1985). Analysis and Synthesis of Fricative Consonants. Doctoral Dissertation, Massachusetts Institute of Technology.

Wode, H. (1983). *Papers on Language Acquisition, Language Learning, and Language Teaching.* Heidelberg: J. Groos.

Wright, R. (2004). A review of perceptual cues and cue robustness. In B. Hayes, R. Kirchner, & D. Steriade (Eds.), *Phonetically-Based Phonology* (pp. 34-57). Cambridge University Press.