McGill Working Papers in Linguistics **16 (2): 11 – 64**

AUDITORY DISTANCE IN SECOND LANGUAGE PERCEPTION*

Kathleen Brannen
McGill University

ABSTRACT

This paper examines differential substitution of the L2 English voiceless interdental fricative, [θ]. The L1s investigated in this study — European French, Québec French, and Japanese — reportedly substitute [s], [t] and [s] respectively in production (e.g. Wenk 1979, Gatbonton 1978, Hancin-Bhatt 1994a). Two main hypotheses are explored: 1. Transfer is perceptually based; 2. Substitution involves an assessment of non-contrastive in addition to contrastive features. Results of an AXB task show that advanced learners are unable to perceive certain non-contrastive distinctions; however, unlike Japanese listeners, French listeners *do perceive* Strident and Mellow, features which are non-contrastive in their L1. Results indicate a perceptual basis for the Japanese substitute; however, the difference between Québec and European French does not seem to be perceptually rooted. Another finding is that confusion of [f] and [θ] is greater for French than it is for Japanese listeners. It is proposed that the composition of the L1 phonetic inventory influences which features listeners attend to during perception.

1. INTRODUCTION

Research in second language acquisition has shown that one's second language is strongly influenced by one's native language (see e.g. Robinett and Schachter 1983; Ioup and Weinberger 1987 for articles therein). 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 study is concerned with the phonetic and phonological components.

A predominant characteristic of second language speech is the presence of foreign accent. 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 which is transferred is often subject to variation. This variation has been termed "differential substitution" (Weinberger 19 94). Differential sub-

* I would like to thank Heather Goad for her extensive discussion and encouragement with regard to all aspects of this study. I would also like to thank Lydia White for advice regarding the experimental design and SLA theory, Jeffrey Steele for valuable comments as well as assistance in recruiting participants, and Yvan Rose for his help with the technical component of this research. This paper, in its various stages, has been presented at the Montréal-Ottawa-Toronto (MOT) Workshop (1998 and 2001) Psycholinguistic Shorts (1999 and 2001), and the Canadian Linguistic Association (CLA) (2001). I thank those audiences for their comments. All errors are of course my responsibility alone. This research was supported by an FCAR doctoral fellowship and a McGill FGSR fellowship.

stitution is observed both cross-linguistically and within a language. An example of cross-linguistic differential substitution involves the L2 German target voiceless velar fricative, [x]. In place of this target, native Kazakh speakers will use the voiceless uvular stop, [q], while native Thai speakers will substitute the voiceless velar stop, [k] (Keel 1979). An example of differential substitution within a language (intralanguage variation) comes from German learners of L2 Swedish. The Swedish target, the low, back vowel [ɑː], is either produced as [oː] or as [aː] (Hammarberg 1997)

This paper is concerned with cross-linguistic differential substitution as it applies to the L2 English target voiceless interdental fricative, [θ] (theta), in word-initial onset position.[1] The cross-linguistic perspective is investigated by examining European French (EF), Québec French (QF), and Japanese (JA). These languages have both /t/ and /s/ in their phonemic inventories; yet JA and EF substitute [s] in place of [θ], while QF speakers substitute [t]. These facts suggest that an explanation for differential substitution is not to be found in the makeup of underlying phonemic inventories, which are underspecified for non-contrastive features. While their phonemic inventories are similar, the phonetic representations of the relevant segments differ across these languages. These facts have led me to the hypothesis that the source of differential substitution is based on *phonetic* rather than *phonemic* representations (see also Flege 1995; cf. Brown 1997).

There is some 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 and Strange 1982). This is because word-finally, English /l/ is velarized, making the distinction more salient. This result would not be expected if listeners were basing their evaluation on a phonological representation, underspecified for non-contrastive features.

Phonetic approximation also shows that L2 learners take into consideration non-contrastive phonetic information in the target language. 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].

As already mentioned, EF and JA speakers usually produce theta as [s]; while QF speakers usually produce it as [t] (Weinreich 1966; Gatbonton 1978; Teasdale 1997; Brannen 1998). As concerns perception, studies of JA show that this language group most commonly confuses [θ] with [s] (Hancin-Bhatt 1994a). (To my knowledge, there are no perception studies which examine differential substitution for either EF or QF.) Perceptual confusion is carried over into production for JA speakers. In other words, when a target sound is misheard as being equivalent to a native sound, that native sound will be produced in place of the target. I thus hypothesize that transfer in

1 Henceforth, the term *differential substitution* will refer to differential substitution of theta.

production is due to perceptual factors.

I propose that speech perception involves an evaluation of all features: both contrastive phonemic and non-contrastive phonetic features. The intake form — a fully specified surface representation — is compared with fully specified internal phonetic representations. This is accomplished via a mechanism which assesses the auditory salience of intake features and those of native representations. The salience of one feature may be influenced by another feature with which it co-occurs; that is, one feature may mute or diminish the auditory salience of another feature.

As concerns differential substitution, this mechanism selects the native representation which is closest to the target segment through an algorithmic evaluation of auditory distance. By definition, one or more of a substitute's features do not match that of the target segment. Featural mismatches which are auditorily muted or diminished are preferred over those which are auditorily salient.

The experiment reported in this paper tests whether and which phonemic and phonetic features listeners can perceive. The algorithm outlined in the first part of the paper is applied to the results.

This study fills the void in perception research on EF and QF by examining the degree to which perceptual confusion results in differential production substitutes for these languages. As well, examination of JA and Native English (NE) provides a basis of comparison with the French data and with previous studies.

To situate the current investigation, I begin with a summary of previous research on interdental substitution in Section 2. My theoretical assumptions are in Section 3. In that section, I present a model of speech processing and introduce the features and feature geometry of relevance in the paper. I discuss the concept of processing levels and then present an algorithm for calculating "auditory distance" based on featural mismatches and relative salience. Section 4 gives the phonetic representations for the learner languages investigated in this study. Using the algorithm outlined in Section 3, I show my predictions for the most frequent substitutes for EF, QF, and JA in Section 5. Section 6 elaborates upon the experimental design: participants, stimuli, and task. Section 7 gives the results of the experiment. These are followed by a discussion of the findings.

2. PREVIOUS RESEARCH ON INTERDENTAL SUBSTITUTION

The study of differential substitution is complex. This complexity is partly due to the need to separate out many potentially confounding factors: for example, English L2 proficiency level, orthographic influence, positional effects, perceptual vs. production factors, extralinguistic cues, phonetics vs. phonology, etc.[2] In addition to this, relatively little is known about the phonetics of fricatives and especially about what features or cues are used in their perception and production.

Transfer research involving the English interdental has focussed primarily on pro-

2 These factors must be considered for any type of transfer at the phonological/phonetic level.

duction data. However, there are a few studies which have investigated perception of this sound as well.

For the languages of interest in this paper, this research indicates that for EF, the most common production substitute for target [θ] is [s] (e.g. Berger 1951; Wenk 1979); for QF, it is [t] (Gatbonton 1978; Brannen 1998); and for JA, [s] (Hancin-Bhatt 1994a). Note that [f] has also been reported for EF and QF (Brannen 1998). There are no perception studies which examine the relative degree of confusion between target [θ] and [f s t] for EF and QF; however, for JA, it has been shown that [s] is the most common perceptual substitute (Hancin-Bhatt 1994a). For an extensive list of differential substitutes for [θ], see Appendix A.

In Sections 2.1 to 2.4, I examine studies which investigate the perception of theta from a cross-linguistic perspective — those which are particularly relevant to this paper. To my knowledge, there are only four such studies in relatively recent literature. This points to the penury of perception studies on interdental substitution — a gap that the present study tries to fill.

2.1. NEMSER 1971

Nemser investigates the perception and production of English alveolar stops and interdental fricatives by native speakers of Hungarian. Nemser states that the primary goal of his study is to provide a description of substitution in Hungarian-English contact. While emphasizing the descriptive nature of his study, Nemser nonetheless makes certain assumptions regarding phonetic/phonological structure. In particular, he notes the need to appeal to acoustic relationships, especially with respect to the sibilant and non-sibilant fricatives (p.36). Also, he favours a view where phonetic space is seen as gradient rather than in terms of binary oppositions. As we shall see later in this paper, both of these ideas are relevant to my analysis of interdental substitution, if not in letter, in spirit at least.

To test perception of English [θ], Nemser uses both an oddball and identification task.[3] Results of these tasks showed that, in onset position, the most common perceptual substitute for Hungarian listeners was [f]. As we shall see, this finding is consistent with the results obtained for EF and QF in the present study.

2.2. HANCIN-BHATT 1994

In her dissertation, Hancin-Bhatt investigates the perception and production of English interdental fricatives by native speakers of German, Hindi, Japanese, and Turkish. In addition, she includes a native English control group. Hancin-Bhatt proposes

3 In the oddball task, four sounds are presented: one of the four is different from the other three. The participant must identify the sound that is different. In the identification task, participants were asked to assign a symbol to the sound heard. (There was also an identification task involving Hungarian orthography.)

that differential substitution can be explained if we consider that there are two types of languages: those that maintain the place of articulation specification at the expense of the manner specification (Type A) and those that prefer to maintain the manner specification of the target at the expense of the place (Type B). In order to determine whether a language is Type A or B, Hancin-Bhatt examines the "functional load" of its distinctive features. She relies on Radical Underspecification (RU) theory (e.g. Archangeli 1984) to calculate functional load.

For a given inventory of sounds, RU allows for a variety of representations. Hancin-Bhatt adheres to the principle of Lexical Minimality (e.g. Halle 1959) and considers that the preferred inventory is the one with the smallest number of specified features. For example, if a given inventory has 10 [–continuant] segments and only 8 [+continuant] segments, then the positive value, being less frequent, will be specified in the phonemic inventory. A redundancy rule, [ ] → [–cont], will insert the negative value later in the derivation (presumably in the phonetics).

Hancin-Bhatt considers that Lexical Minimality can determine the functional load of features. The functional load of a given feature — or, more precisely, featural value — is equivalent to the number of times it is specified in the inventory divided by the total number of phonemes. For example, in an inventory containing 15 phonemes and where 8 segments bear the feature [+cont], the functional load for the feature [+cont] will be 8/15 = .53 (Hancin-Bhatt 1994b). In a Type A language, place features bear a higher functional load than manner features, and a redundancy rule will insert [–cont], resulting in the substitute [t]. In a Type B language, manner features bear a higher functional load, and a redundancy rule will insert [+cont], resulting in the substitute [s]. Based on her calculations of functional load, she predicts that, in perception and production, 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 uses an identification task to investigate perception of the interdentals. In this task, participants were asked to listen to nonsense words containing [θ,ð] and to identify the target sound using a number representing the following sounds: [f v θ ð t d s z].[4] With respect to onset position, the results of this study showed that theta was correctly perceived 64% of the time for the English group, 33% for Japanese, 47% for German, 48% for Hindi, and 36% for the Turkish group. The interdentals were most often confused with labiodental fricatives for the German(22%), Hindi(23%), Turkish(24%), and native English groups(26%). After the labiodentals, the next most frequently confused sound by German listeners was [s](5% error rate) vs. [t](0%). Second place for Hindi listeners was shared by [s](5%) and [t](2%); whereas, Turkish listeners heard theta second most often as [t](18% vs. 0% as [s]). The Japanese group most often confused theta with [s](25%) vs. [t](1%). Aside from the errors involving [f], Hancin-Bhatt's predictions were borne out for the

4 Participants were also given the option of responding "don't know", but Hancin-Bhatt does not report these responses; therefore, the numbers she reports do not total 100%.

Turkish listeners and possibly for the German group. However, her predictions were not supported for the Hindi or Japanese groups.

Hancin-Bhatt's model fails to account for the high rates of substitutions involving [f, v]. She claims that perception of [θ,ð] as [f, v] is due to universal acoustic factors. However, it is interesting that among all the language groups, Japanese listeners had the lowest rate of labiodental confusion (17%). This finding will become a factor in the discussion of my results later in this paper. Furthermore, Hancin-Bhatt's model cannot capture the difference between EF and QF: given that they share the same underlying consonant inventory, they would be classed as the same type of language. Hancin-Bhatt's thesis presents several other difficulties, both theoretical and methodological. First, I discuss the theoretical problems, then the methodological shortcomings.

Early in the thesis, Hancin-Bhatt discusses the need for phonetic detail in a feature geometric representation. She introduces a highly articulated feature geometry in which the Coronal node dominates Dental and Alveolar places of articulation. These Coronal dependents in turn branch into [±back] in order to capture Interdental vs. Dental and Alveolar vs. Postalveolar places of articulation. Similarly, the Labial and Dorsal nodes each dominate [±back] in order to capture Bilabial vs. Labiodental and Palatal vs. Velar places of articulation. Unfortunately, she does not develop this analysis in her thesis; rather, she appears to discard it in favour of the functional load analysis based on Radical Underspecification. The present study was partly inspired by Hancin-Bhatt's detailed feature geometry.

With respect to her functional load analysis, it is questionable whether Hancin-Bhatt has correctly represented the phonemic inventories for the languages she examines. Radical Underspecification requires that language-internal phonological processes motivate the configuration of an underlying inventory; in the absence of such processes, it is markedness which dictates the inventory specifications. Hancin-Bhatt provides no such evidence, nor does she rely on markedness; instead, she relies on "minimal specification" as outlined above. This formulation seems to be rather ad hoc, especially since the status of the underlying inventories she adopts is suspect. For example, for Japanese, she includes both /f/ and /h/ as phonemes, when /f/ (or more precisely [ɸ]) is an allophone of /h/, occurring before the high back vowel. Likewise, for German, she includes both /ç/ and /x/ in the phonemic inventory, when these two segments are in allophonic distribution. Thus she confuses phonetic and phonemic information.

Methodological problems are also evident. In the perception experiment, participants completed an identification task. Because this task involves assigning numbers to sounds heard, it presupposes phonetic virtuosity. Another problem with Hancin-Bhatt's study is that the L2 English proficiency levels of her participants varied from low to high, yet they were all grouped together in the results. This introduces a confounding factor in her study.

In contrast, the present experiment controls for proficiency level by restricting the investigation to advanced learners. Problems with the identification task are avoided by using a more robust experimental paradigm (see Section 6.3.1). Also, the present study attempts to make a clear distinction between phonology and phonetics. Finally, rather than dismissing the common interdental substitute [f], I directly investigate the extent to which the labiodental fricative is perceptually substituted cross-linguistically and explain why it is often confused with theta.

2.3. BROWN 1997

In her dissertation, Brown investigates how Japanese speakers of L2 English perceive the following contrasts: /l/-/r/, /f/-/v/, /p/-/t/, /p/-/f/, and /s/-/θ/. (Note that her study is not a test of differential substitution, since there was only one test contrast involving theta; however, her results serve as an indirect comparison with findings from other researchers.) Brown hypothesizes that second language learners can recombine features that are contrastive in their native language in order to form new representations. On the other hand, she claims they cannot acquire new features; that is, features which do not function contrastively in their native language.

In Japanese, Continuant is contrastive; for example, it serves to distinguish Japanese /t/ from /s/. The Japanese phonemic inventory includes /p/, but not /f/. Since Japanese has the contrastive feature Continuant elsewhere in its inventory, Brown predicts that Japanese listeners will be able to distinguish /p/ from /f/ by recombining features. On the other hand, the Japanese phonemic system does not make use of Strident, the feature which distinguishes /s/ from /θ/.[5] She predicts that Japanese listeners will not perceive the difference between /s/ and /θ/, and, in fact, will never acquire this contrast because Strident is unavailable in the Japanese phonological system, being non-distinctive.

Brown uses a 4IAX task to test whether listeners can perceive features at a phonetic level of processing.[6] The Japanese participants had studied English for an average of eight years in Japan. The results support Brown's hypothesis that it will be easy for L2 learners to hear new contrasts which involve a feature used elsewhere in their inventory. As predicted, Japanese listeners perceived /p/ vs. /f/ as well as did the Native English control group; whereas, their performance on the /s/ vs. /θ/ contrast was significantly worse than Native English controls. Despite the fact that these results support Brown's hypothesis, later we shall see that her hypothesis that listeners fail to perceive features which do not function contrastively in their L1 cannot be upheld in light of my results based on French.

5 Brown actually uses the feature Distributed to distinguish these segments; however, she states that her predictions would be the same if one adopted Strident as the relevant feature: neither Distributed nor Strident are present underlyingly in Japanese (p.222).

6 The 4IAX task involves listening to two pairs of stimuli. One pair consists of the same stimuli, while the members of the other pair consist of different stimuli. The task is to determine which pair has different members.

2.4. LaCharité and Prévost 1999

In their paper, LaCharité and Prévost set out to test Brown's hypothesis. While Brown proposes that features which are non-contrastive in the L1 phonemic inventory cannot be acquired, LaCharité and Prévost hypothesize that not all non-contrastive features are equal. They suggest that terminal features which are absent from the L1 phonemic inventory can be more easily acquired than Major Articulator features (which can have dependents). They use both an AX and ABX task to determine how well advanced Québec French learners of English hear the contrasts [h] vs. Ø, [ŋ] vs. [n], and [θ] vs. [t].[7] According to LaCharité and Prévost, the contrast [h] vs. Ø involves the Major Articulator feature Pharyngeal, and thus should be difficult to acquire, being absent from the feature geometry of QF. The contrast [ŋ] vs. [n] involves a feature (Dorsal) which is used distinctively in QF, although not to make this particular contrast. Being active in the L1 phonology, QF learners should be able to recombine Dorsal with other features to establish this new contrast. LaCharité and Prévost further argue that [θ] vs. [t] should be more perceptible than [h] vs. Ø. While both contrasts involve features which are absent from the QF phonemic inventory, LaCharité and Prévost claim that the former contrast involves a terminal feature, Distributed; therefore, it should be more easily acquired than the latter contrast, which involves a Major Articulator feature. Results of the AX task showed the [h] vs. Ø contrast was more difficult to discriminate than the [θ] vs. [t] contrast. However, LaCharité and Prévost did not find any significant difference between [ŋ] vs. [n] and [θ] vs. [t]. In the ABX task, there was no difference between any of the contrasts, contrary to their hypothesis.

There are some problems in LaCharité and Prévost's hypotheses and interpretations. In particular, it is not clear that Distributed is the feature upon which QF listeners are depending to make the distinction between [θ] and [t]: this contrast also differs on continuancy and place dimensions. Theta is a continuant, while [t] is not; as well, theta is a dental articulation, while [t] is alveolar.[8] The present study attempts to control for these confounding factors by examining contrasts which involve a single feature.

In addition, LaCharité and Prévost did not test [s] vs. [θ], another contrast which, according to Brown, also involves the feature Distributed. As we shall see in Section 7, JA behave differently with respect to [s]/[θ] and [t]/[θ], suggesting that the same feature is not serving to contrast these pairs.

---

7 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. Thus, items vary as to whether they are ABA or ABB. The participant's task is to select whether X=A or whether X=B.

8 Although LaCharité & Prévost do not mention the precise place of articulation of stimulus [t], given that the stimuli were recorded by a Canadian anglophone, we can expect that the stop is alveolar.

3. THEORETICAL FRAMEWORK

In the previous section, we have reviewed four studies that investigate the perception of the interdental fricative by learners of L2 English. I have pointed out the need to investigate phonetic factors rather than adopting a purely phonemic approach to interdental substitution. In this section, I discuss both phonetic and phonological features and propose a model of speech perception which incorporates phonetic factors. I also examine the role of auditory saliency in speech perception.

3.1. SUBSTITUTION HAS A PERCEPTUAL BASIS

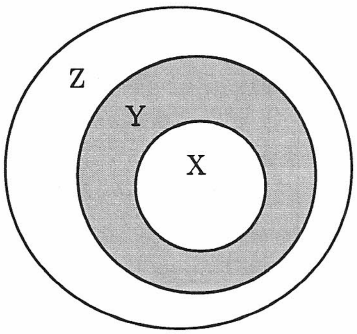
I hypothesize that production is influenced by perception. For example, if a learner substitutes [t] for [θ] in production, it is because target [θ] has been associated with a perceptual representation for [t]. I assume transfer is based on full perception of intake or surface features; in other words, all the features of target [θ] are perceived. Importantly, this assumption does not mean that learners "instantly acquire" [θ]. Although they perceive the individual features of the surface intake, I assume that feature co-occurrence restrictions in their L1 grammar prevent them from combining certain features; that is, from constructing a new output representation. Thus, although the learner perceives all features on [θ], this surface target is merged with their perceptual representation for [t]. In turn, this perceptual representation is (indirectly) mapped to an articulatory representation for [t], resulting in production transfer.

3.2. FEATURES

The present work is conducted within the generative framework. I assume that phonetics as well as phonology falls under the purview of Universal Grammar (UG). By this I mean that phonetics is part of the cognitive module which governs language as opposed to other cognitive systems. Phonology differs from phonetics in two principal ways. First, I consider that phonological representations encode information from all modalities, for example, auditory, articulatory, and visual (see Hardison 1999 for discussion); whereas, phonetic features and representations are either articulatory or auditory, as we shall see shortly. There is nonetheless a close mapping between phonetic and phonological features and representations. Second, phonology and phonetics differ in degree of feature specification.

The phonological component of the grammar is underspecified for non-conrastive features. In this paper, I adopt a version of the theory of Contrastive Underspecification (e.g. Steriade 1987; Calabrese 1988; Mester and Itô 1989; Rice and 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, the feature [–high] is contrastive for mid vowels, but not for low vowels. Empirical evidence has shown that mid vowels are targeted in processes which appeal to [–high]; however, low vowels are non-participants in these processes. Thus, being redundant on low vowels, [–high] is underspecified in the phonology.

In the phonetic component, however, I consider that non-contrastive features are additionally specified. Features found in the phonetic component of a given language form a subset of the universal set of features. The relationship between phonological, phonetic, and universal phonetic features is depicted in Figure 1.[9]



X = Set of Fs for Phonology of Language A
Y = Set of Fs for Phonetics of Language A
Z = Set of Fs for all languages (Universal set).

Figure 1. Phonological and Phonetic Features.

For example, the feature Apical would be specified both in the *phonological and phonetic* components of Australian languages which contrast laminal and apical stops. On the other hand, in French, Apical would be absent from the phonological module (X): since French has only one series of coronal stops, Apical is non-contrastive. However, given that French coronal stops are articulatorily described as being apical (Dart 1991), the feature Apical would be present in the phonetic component (Y).[10] In a language that only has apical segments and no laminal segments, Laminal would be absent from both X and Y in that language. But since Laminal is a feature which is provided by UG, it always resides in Z.

In the latter situation, where a feature is absent from both X and Y in a particular language, I hypothesize that native speakers of this language will nevertheless be able to hear this feature during phonetic processing. In other words, listeners have access to Z in a phonetic mode — listeners have "full perception". This is a strong claim. We will see in Section 7 that I will weaken this hypothesis; however, it is a necessary starting point since we have no a priori knowledge of which features listen-

9 This relationship refers to position in geometry (see Figure 3); it does not imply that phonological and phonetic features are substantively identical.

10 Dart more accurately describes European French stops as Apicolaminal. This does not affect my analysis, since neither Apical nor Laminal is contrastive on stops in French.

ers can perceive.

While the hypothesis states that listeners are able to hear the individual features, as mentioned above, this does not necessarily mean that they are able to introduce new features or recombine existing features into a new geometric 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 representations in their native phonetic inventory. This results in transfer.

3.3. ARTICULATORY AND AUDITORY FEATURES

In the spirit of Flemming (1995), Jun (1995), and Hamilton (1996), amongst others, 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. Very little research has been done on the status of auditory features in terms of monovalency or bivalency. For ease of exposition, I use monovalent features.[11]

Table 1 shows the equivalencies between articulatory and auditory features. The first column indicates on which level in the feature geometry the relevant features reside. (The reader may wish to refer to the feature geometry in Figure 3 in Section 3.5.1 for comparison). The articulatory features are listed in the second column, and their auditory counterparts, in the third column. For the purposes of this paper, I assume a one-to-one correspondence between auditory and articulatory features; however, it is clear that this relationship does not always obtain (see e.g. Flemming 1995 for discussion). I leave this question open to further empirical research.[12]

11 Auditory features seem to behave in a gradient manner, being specified with "more or less" of a given feature. We will see this is the context of muting effects discussed in Section 3.5.3. Whether auditory features are formally represented as gradient requires further research.

12 One situation of relevance to this study where the correspondence is not one-to-one relates to place of articulation. For stops, the relevant auditory features are Grave and Acute (referring to formant transitions); whereas, for strident fricatives, place is cued by peaks of intensity in the noise spectrum (e.g. Jakobson, Fant, & Halle 1967; Tabain 1998).

Table 1. Articulatory Features and their Auditory Equivalents

Level in Feature Geometry	Articulatory Feature	Auditory Feature[13]
Airflow	Continuant, Stop	Aperiodic, Interrupted
Turbulence	Channel, Spread	Strident, Mellow
Major Articulator (Place)[14]	Labial, Coronal	Grave, Acute
Location	Lip, Dental, Alveolar	Lo F2, Mid F2, Hi F2
Minor Articulator	Laminal, Apical	Slow Transition, Fast Transition

This paper is primarily concerned with auditory features. However, for the sake of clarity, I will use the familiar articulatory labels such as Coronal and Stop, rather than the less familiar terms such as Acute or Interrupted. The latter labels may be implied. Exceptions to this convention are the features Strident and Mellow. These auditory features are well known, so I will not refer to their articulatory counterparts.

3.4. Interfaces: Perception-production and Phonology-phonetic

In this section, I elaborate a model of speech perception and production which demonstrates the mechanism behind transfer at both phonemic and phonetic levels. As can be seen in Figure 2 below, the information entering the linguistic module is labeled *intake* (Brown 1993). The features which are present in this intake form depend on the type of processing involved. In phonetic processing, all features are present in the intake. In phonemic processing, only contrastive features are present. The intake form is mapped to the closest L1 representation (see Section 3.5.4). In a phonemic assessment, this is an underlying representation; in a phonetic assessment, it is an auditory representation. During production, the underlying representation feeds into an articulatory representation; the latter determines the output.

---

13 Interrupted, Grave, Acute, Mellow, and Strident are borrowed from Jakobson, Fant, & Halle (1967); I introduce the feature Aperiodic to characterize continuant obstruents. Slow and Fast Transition are taken from Stevens (1998). Location features represent a finer gradation on the place dimension, thought to be cued either by the relation between F2 and F1 (Jakobson, Fant, & Halle 1967:30) or by peaks of spectral energy (Strevens 1960).

14 Although technically Major Articulators, Labial, Coronal, and Dorsal are often referred to as Place features (e.g. Sagey 1986).

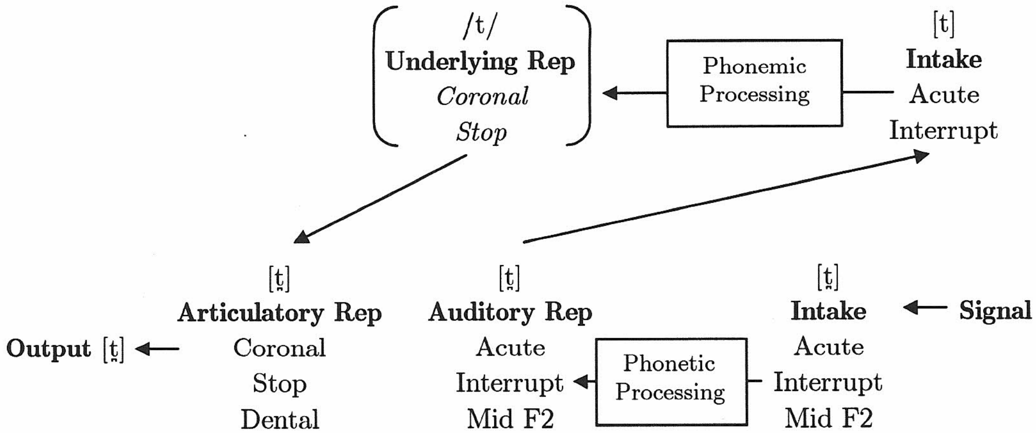


Figure 2. The Perception-production/Phonetics-phonology Link

This view of the grammar is quite speculative at this point; however, it provides a starting point within which to situate issues discussed in this paper.

3.5. AUDITORY DISTANCE

Transfer occurs when the intake form is compared to representations which exist in stored memory.[15] 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. In order to explain exactly how features are compared, it is necessary that I introduce the feature geometry assumed in this paper.

3.5.1. FEATURE GEOMETRY

I assume a hierarchical structure for subsegmental components, borrowing from various models of Feature Geometry (Sagey 1986; Gorecka 1989; Rice and Avery 1995). Figure 3 below shows a highly articulated feature geometry. The geometry looks complex; however, any phonological or phonetic analysis which involves fine distinctions under a major articulator node requires a relatively high degree of detail (e.g. Gnanadesikan 1993; Hancin-Bhatt 1994a). Recall that I have, for the most part, used articulatory labels in this geometry, but that I am actually assuming the auditory counterparts of these features. I presume that corresponding articulatory, auditory, and un-

15 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.

derlying features occupy the same location in the geometry.[16] The features in the geometry reside on levels which correspond to the articulatory parameter which results in particular auditory reflexes (see Table 1). In Figure 3, organizing node labels are in capital letters and level labels are listed down the left side of the geometry.[17] The Airflow level relates to the degree of constriction of obstruents: total constriction defines a Stop (Interrupted), and narrow constriction characterizes a Continuant (Aperiodic). As its name suggests, the Turbulence level relates to degree of noise: Strident (Channel) obstruents are noisy because the airflow is directed through a narrow channel at high velocity; the airstream may also be reflected off an obstacle such as the teeth. Mellow (Spread) sounds, which include stops, are quiet. Major Articulator features reflect the major active articulators involved, for example, the lips (Labial/Grave) or the tongue tip or blade (Coronal/Acute). The Location level relates to the passive articulator, for example, Lip (Lo F2), Alveolar (Hi F2). Location features are proposed in the spirit of Gorecka (1989), who incorporates both Articulator and Site (location) features under a Constriction node in her geometry. Features on the Minor Articulator level involve the precise part of the tongue which is involved in Coronal articulation, for example, either the blade (Laminal/Slow Transition) or the tip (Apical/Fast Transition).

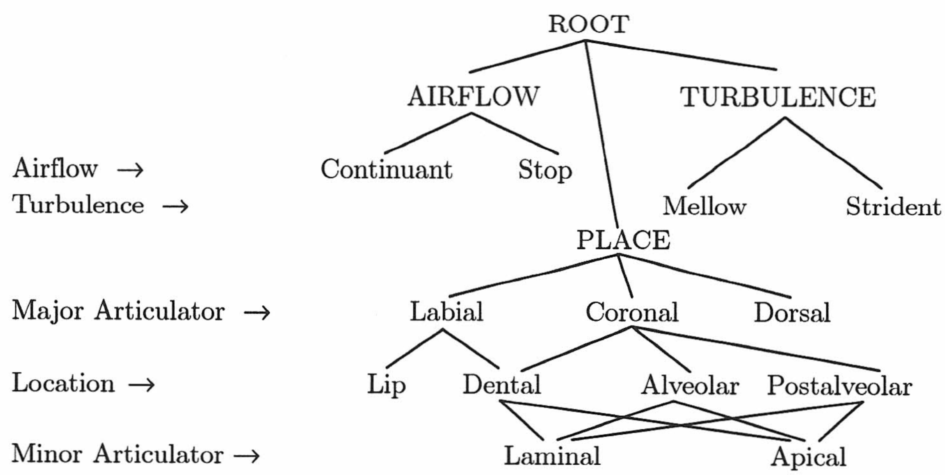


Figure 3. Feature Geometry of Auditory Features

16 This statement is adequate for the purposes of this paper; however, I acknowledge the possibility that some phonetic features may be interpreted differently in the phonology, e.g. a vowel which is phonetically mid functioning as low in phonological processes.

17 Levels, which have no formal status, are to be distinguished from tiers. I adopt the standard view that each feature resides on a separate tier. Levels encode dependency relations in the geometry. Organizing nodes (e.g. PLACE and AIRFLOW), do have formal status, but have no auditory or articulatory content; they serve to group features together.

If used contrastively in a language, these features belong to the phonological component of its grammar. If they are non-contrastive, yet serve to define the precise auditory specifications of a segment, then these features belong to the phonetic component.

The feature geometry is part of UG and thus available to speakers of all languages. As such, speakers have knowledge of dependency relations provided by the geometry. I hypothesize that in phonetic processing, listeners perceive all features in the intake (see Section 5.2). So, when listeners hear a feature in the intake, even if their native language prevents them from combining that feature with another to construct a representation, they nevertheless know to which node that feature reports in the geometry. Importantly, even if the L1 grammar does not contain a particular feature, listeners nonetheless have access to its dependency relation. For example, the JA grammar contains the feature Alveolar, but not Dental. Nonetheless, when JA listeners hear Dental in the ambient language, their knowledge of feature geometry tells them it is a Location feature. Features from native representations are compared with intake features of the same type, for example, Location features are compared with other Location features. When a JA listener hears Dental in the intake, it is compared with Alveolar in the representation from the native inventory.

When native representations are compared with the intake, features which do not match are tagged as mismatches. These mismatches 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. This distance depends on the "weight" of the respective features as will be discussed in the next section.

3.5.2. SALIENCY IN AUDITORY FEATURES

There is a growing literature on the role of auditory salience in determining the shape of phonological systems (e.g. Flemming 1995; Jun 1995; Hamilton 1996; Hume and Johnson 2001). Salient auditory features, such as Strident, are easily picked up from the speech signal because they are prominent. All else being equal, salient features are preferred over less salient or recessive features (see Section 3.5.3).

I suggest that auditory salience may be defined by position in the geometric tree structure. Features which are directly dependent on an organization node are auditorily salient as compared to features which are dependent on other features. Airflow, Turbulence, and Major Articulator features are all dependent on an organizing node; thus, they are salient and, hence, easily recoverable from the auditory intake.[18] In contrast, Location and Minor Articulator features are dependent on other features, and thus they are less salient. Therefore, the auditory reflex elicited by the Location distinctions and Minor Articulators are harder to perceive, for example, Dental/Alveolar

18 The reader may wonder how a feature such as Mellow can be perceptually salient. It is the obvious lack of noise inherent to segments bearing this feature which is salient.

and Laminal/Apical.

To formalize this conception of salience, I introduce the notion of feature weight. I propose that salient auditory features carry more weight than do less salient auditory features. Weight is equivalent to "cost" in featural assessment. Let us see how the idea of featural weighting figures into the interdental substitution phenomenon. As alluded to earlier, the comparison of the target intake with native representations involves assessing the number of featural mismatches. A featural mismatch is assessed as to its cost. It is more costly to substitute a salient feature in place of the input feature than it is to substitute a recessive feature. 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 presented with a choice, it is better to substitute a "wrong" feature which does not readily stand out, than a wrong feature which is blatantly wrong. Salient features bear a cost or weight of 2; recessive auditory features bear a weight of 1.

3.5.3. Relative Saliency: Muting Effects

Despite the fact that, in isolation, two features can be equally salient, features do not occur in isolation. Often, one feature impacts upon the auditory salience of another. Because of this, I consider that auditory salience is relative. The salience of an auditory feature can be affected by the context in which that feature appears.[19] Stevens et al. (1986) discuss how some features tend to enhance the perceptibility of other features; for example, they suggest that Continuant is enhanced by Strident. I assume this to be the case, but will look at it the other way around — one feature can mute the saliency of another feature, for example, Mellow mutes Continuant.[20] In other words, a mellow fricative sounds more like a stop. This suggests an explanation for why many L2 speakers of English substitute the interdental with a stop. In sum, the salient feature Continuant will become less salient or muted in the context of the feature Mellow. Instead of bearing a weight of 2, muting has the effect of reducing its weight to 1. I consider that muting effects happen between auditory features only. An articulatory feature cannot mute an auditory feature or vice-versa. Auditory salience and muting effects are summed up in Table 2. Listed in the first column are features which are inherently recessive as well as those that are muted when they co-occur with another feature; inherently salient features are listed in the second column.

19 Note that this relationship between features is independent of dependency relations in the geometry. Thus muting also occurs in the intake, which consists of an unordered bundle of features.

20 The effects of muting/enhancement were taken into consideration in the design of the experiment. For example, the stimuli which examine the ability to perceive Dental vs. Alveolar were articulated as Laminal and Apical respectively. For example, Laminal needs to be reinforced by Dental in order to be perceptually recoverable (see discussion in Hall 1997: 42).

Table 2. Auditory Salience and Weight

Inherently Recessive and Muted: Weight = 1	*Salient: Weight = 2*
Continuant (with Mellow)	Stop, Continuant
Strident (with Dental)	Mellow, Strident
Lip, Dental, Alveolar, PostAlveolar[21]	Labial, Coronal, Dorsal
Laminal, Apical	

The values given in Table 2 make predictions as to the distribution of segments in inventories. Let us examine what the statements in this table imply. The Major Articulator features, Labial, Coronal, and Dorsal, are all salient. Maddieson's 1984 survey shows that languages show equal preference for labials, coronals, and dorsals, at least as concerns stop articulations. In the first column of Table 2, I state that Mellow mutes Continuant. This means that continuancy is more difficult to hear on a mellow fricative. Because languages strive for maximal distinctiveness in their inventories, all else being equal, languages will prefer a strident fricative over a mellow fricative. This maintains maximal distinction between stops and fricatives. Thus we can predict that languages will prefer $s > \theta, f, \phi$. Maddieson's (1984) survey upholds this prediction. The stridency dimension accounts for why fricatives do not pattern with stops with respect to their distribution on the Major Articulator dimension. In other words, Labial fricatives, being Mellow, are less common than Coronal fricatives in contrast to stops, where labials are as common as coronals.

The other muting effect in the first column of Table 2, Dental mutes Strident, predicts that if a language has both Mellow and Strident coronal fricatives, then these segments will be Dental and Alveolar respectively. Since Strident is muted at the Dental location, [s̪] is too close to [θ] on the stridency dimension to constitute an adequate contrast. This too is confirmed by perusing Maddieson's database. Seven languages have [s] vs. [θ], while only three have [s̪] vs. [θ], and in two of these three cases, [θ] is derived from either /s/ or /t/ (Arabic and Spanish).

Finally, given that Stop and Continuant are equally salient, the prediction is that stops and strident fricatives should be equally represented in languages of the world. This is not entirely true. While all languages have stops, some languages do not have fricatives, for example, Australian languages. This pattern can be attributed to articulatory markedness: fricatives are articulatorily marked as compared to stops.

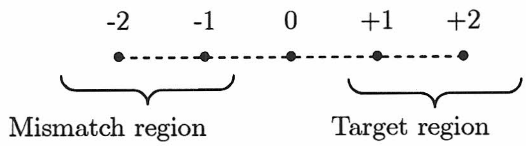
3.5.4. TRANSFER ALGORITHM: CALCULATING AUDITORY DISTANCE

Recall from Table 2 that salient auditory features carry a value of 2, while muted fea-

21 I assume that by itself PostAlveolar would be non-salient; however, this feature is enhanced with Round in [ʃ]. Together these two features become salient.

tures carry a value of 1. When a feature in the stored representation matches one in the intake, it receives a positive value: +1 or +2. If the feature does not match, it receives a negative value: –1 or –2. The evaluation of distance takes place along a scale:

Figure 4. Auditory Distance Scale



Let us look at how this works with a hypothetical example. In Table 3 below, features are represented with letters, and their weight plus their match/mismatch status are indicated. Intake features always have a positive value. In this table, Feature F of the intake is recessive and thus has a weight of 1; feature G is salient and so has a weight of 2. Candidate #1 shares the feature F with the intake; thus, it receives a positive value, +. There is no distance between these features. Feature H of Candidate #1, on the other hand, does not correspond with the intake feature G; it is a mismatch, so it incurs a negative value, –. (Recall from Section 3.5.1 that features on the same level are compared.) The intake feature G has a value of +2, while Candidate #1's feature H is –2; both features are salient and thus the distance between the intake and Candidate #1 is 4.

For Candidate #2, the situation is reversed. Feature G of Candidate #2 is a match for the input feature G. Feature E of Candidate #2 differs from the intake feature F. However, since both features E and F are non-salient, the distance between these two features is only 2. Candidate #2 is closer to the intake form in terms of Auditory Distance; thus, it is chosen as a substitute for the target.

TABLE 3. STORED REPRESENTATIONS FOR SOUNDS: SELECTION CANDIDATES

Intake Target Form: +1; recessive (F) +2; salient (G)	*Match*	*Mismatch*	*Distance*
Candidate #1	+1 (F)	–2 (H)	4 away
→ **Candidate #2**	+2 (G)	–1 (E)	2 away

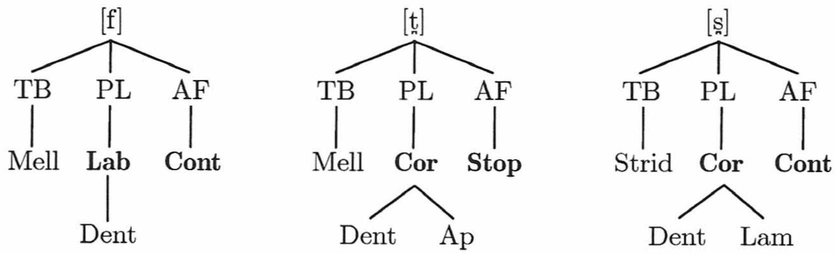
4. AUDITORY REPRESENTATIONS FOR LANGUAGE GROUPS

As discussed in Section 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 each language group in this study. The features which function contrastively in each language are in bold. (To my knowledge, [ʃ] has never been reported as a substitute for [θ] for any of the languages under investigation; this segment is too distant from theta, being Postalveolar, Strident, and Round; therefore, I will not consider it in this study. Refer to Appendix B for algorithmic calculation for this segment.)

4.1. EUROPEAN FRENCH

The phonetic inventory for EF contains a labiodental fricative, an apical dental stop and a laminal dental fricative (Dart 1991; Teasdale 1997).[22, 23] The segments function phonemically and can occur before any vowel. They are represented in Figure 5 below.

Figure 5. European French Phonetic Representations



4.2. QUÉBEC FRENCH

The QF phonetic inventory contains a labiodental fricative, an apical dental stop (Charbonneau and Jacques 1972; Dart 1991), and a laminoalveolar fricative (Teasdale 1997).[24] All of these segments function phonemically, and all can occur before any vowel.

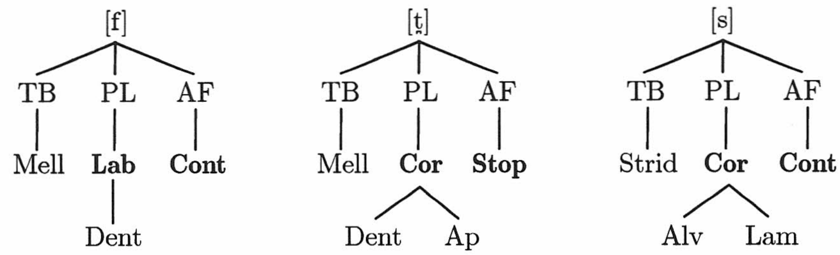
---

22 Dart (1991) shows that [s] varies between speakers: either dental or alveolar. Data from Dart on EF, 'O'dham, and English suggest that there is a universal tendency for [s] to be alveolar. However, the data show that EF speakers produced a dental fricative more often (42%) than did English (22%) and 'O'dham (12%) speakers. This suggests to me a tension between universal and language-specific forces: EF [s] is specified as dental, but universal tendencies pull it toward the alveolar region, hence the observed variation.

23 Dart describes the EF stop as apicolaminal. Specifying EF [t] with both Apical and Laminal would make no difference in my predictions.

24 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.

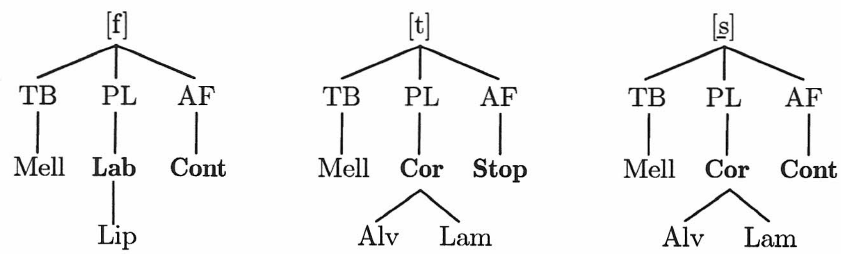
Figure 6. Québec French Phonetic Representations



4.3. JAPANESE

The Japanese phonetic inventory includes a bilabial fricative, a laminal alveolar stop (Someda 1966, cited in Vance 1987:18), and a laminal alveolar (mellow) fricative (Vance 1987; Teasdale 1997).[25] The bilabial fricative [ɸ] occurs as an allophone of /h/ before [ɯ]. Japanese does not have the labiodental fricative either as a phoneme or as an allophone. The allophone [ɸ] is included for symmetry with EF, QF, and NE. The other two segments function phonemically: [s̻] occurs before any vowel except [i], and [t] occurs before non-high vowels.[26]

Figure 7. Japanese Phonetic Representations



4.4. THETA

Below, I give the phonetic representation of the English target interdental fricative. This is the intake that L2 listeners are hypothesized to perceive in phonetic processing. Note that I consider the intake to consist of unordered features; that is, they are not hierarchically organized. Learners match these features to representations from their native inventory. The best fit will be selected as a substitute for the target interdental.

---

25 The description of JA [s] as Mellow comes from Teasdale's spectrographic analysis of this sound.

26 /s/ → [ʃ]/__ [i]; /t/ → [ts]/__ [ɯ]; /t/ → [tʃ]/__ [i].

Intake Target θ
Continuant
Mellow
Coronal
Dental
Laminal

The above representations in Figures 5-7 are fully specified phonetic representations. The main hypothesis of this study is that learners use phonetic, rather than phonemic, information when assessing potential interdental substitutes in naturalistic learning situations. Let us now turn to the predictions I make regarding interdental substitution for the learner groups investigated in this study. First I will show how transfer based on phonemic features fails to account for differential substitution. Next, we will see that if we consider that phonetic features are also assessed in transfer, the correct results are generated.

5. PREDICTIONS

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

5.1 PHONEMIC PREDICTIONS

First I give the predictions based on a phonemic assessment, where only contrastive features are accessible for comparison.

Table 4. Predictions for EF, QF, and JA Based on a Phonemic Assessment

Intake θ	+2 (Cont) +2 (Cor)		
Potential Substitute	*Matches*	*Mismatches*	*Distance*
/s/	+2 (Cont) +2 (Cor)		0 away
/f, ɸ/	+2 (Cont)	–2 (Lab) 4	4 away
/t/	+2 (Cor)	–2 (Stop) 4	4 away

Table 4 shows that the features of /s/, Continuant and Coronal, both match the features on target [θ]. There are no featural mismatches, so the distance of /s/ from [θ] is zero. On the other hand, /f, ɸ/ match the intake on only one feature, Continuant; there is a mismatch on the place feature, Labial. This mismatch results in an auditory distance of 4. For /t/, the match is on Coronal and the mismatch involves Stop, also for a distance of 4. Obviously, an assessment based on contrastive features alone fails to capture differential substitution: it would predict that /s/ would be the preferred substitute for EF, QF, and JA. To my knowledge, /s/ has *never* been reported as a substitute for theta in QF. Let us now see how the algorithm applies to a phonetic level of representation for each of the learner languages investigated in this paper.

5.2 Phonetic Predictions

Recall my hypothesis that all features are perceived in a phonetic assessment. This means that the intake form is fully specified. The intake is compared with phonetic representations from the L1, which are also fully specified for language-specific phonetic properties.

Table 5. Surface Intake

Intake	
θ	+2 (Mell) +1 (Cont)[27] +2 (Cor) +1 (Dent) +1 (Lam)

Table 6. Predictions for EF Based on a Phonetic Assessment

Potential Substitute	*Matches*	*Mismatches*	*Distance*
s̪	+2 (Cont) +2 (Cor) +1 (Lam) +1 (Dent)	-1 (Strid)[28] 3	3 away
f	+2 (Mell) +1 (Cont) +1 (Dent)	-2 (Lab) 4	4 away
t̪	+2 (Mell) +2 (Cor) +1 (Dent)	-2 (Stop) 3 -1 (Ap) 2	5 away

The above table predicts that EF speakers will choose [s̪] as its substitute for the in-

27 Mellow mutes Continuant; hence Cont has a value of 1, not 2 (see Table 2).
28 Dental mutes Strident (see Table 2).

terdental. The next closest substitute is [f]. In the literature, [s] is the most frequently reported substitute for [θ], but [f] has also been reported.

Table 7. Predictions for QF Based on a Phonetic Assessment

Potential Substitute	*Matches*	*Mismatches*	*Distance*
f	+2 (Mell) +1 (Cont) +1 (Dent)	-2 (Lab) 4	4 away
t̪	+2 (Mell) +2 (Cor) +1 (Dent)	-2 (Stop) 3 -1 (Ap) 2	5 away
s	+2 (Cont) +2 (Cor) +1 (Lam)	-2 (Strid) 4 -1 (Alv) 2	6 away

For QF, [f] is selected as the first candidate. The next closest substitute is [t̪] — this is what is most commonly reported in the literature as a production substitute for QF. I do not consider the prediction that [f] wins to be a problem for my analysis because of the high perceptual confusion between [θ] and [f] cross-linguistically.[29] In fact, studies tend to gloss over the fact that [f] is used as a substitute (see Section 2.2). The algorithm proposed here incorporates these perceptual difficulties.

Table 8. Predictions for JA Based on a Phonetic Assessment

Potential Substitute	*Matches*	*Mismatches*	*Distance*
s̻	+2 (Mell) +1 (Cont) +2 (Cor) +1 (Lam)	-1 (Alv) 2	2 away
t	+2 (Mell) +2 (Cor) +1 (Lam)	-2 (Stop) 3 -1 (Alv) 2	5 away
ɸ	+2 (Mell) +1 (Cont)	-2 (Lab) 4 -1 (Lip) 2	6 away

The above table predicts that JA speakers will choose their Laminoalveolar [s̻] as a

29 See Section 7.6 for discussion of a hypothesis that learners use visual cues on [f] and [θ] as an aid in distinguishing these two segments.

substitute. This is what is reported in the literature. The stop comes in second.[30]

In summary, we have seen that in EF, the dental fricative [s̪] is the best substitute for theta 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 labiodental fricative which is predicted to be the closest substitute, while in JA, it is alveolar fricative [s̺] by virtue of it being a mellow continuant.

5.3. EXCURSUS ON HINDI

It is interesting to see that the auditory distance algorithm can also make other predictions. For target apicoalveolar [t], Hindi speakers of L2 English substitute the retroflex [ʈ], even though they have a dental [t̪]. This is shown in Table 9.

Table 9. Predictions for Hindi Substitute for English Target [t]

Intake			
t	+2 (Mell) +2 (Stop) +2 (Cor) +1 (Alv) +1 (Ap)		
Potential Substitute	*Matches*	*Mismatches*	*Distance*
ʈ	+2 (Mell) +2 (Stop) +2 (Cor) +1 (Ap)	-1 (PostAlv) 2	2 away
t	+2 (Mell) +2 (Stop) +2 (Cor)	-1 (Dent) 2 -1 (Lam) 2	4 away

To summarize, the auditory distance algorithm predicts that the differential substitution facts can be captured if we assume that listeners make a selection from their L1 phonetic inventory, rather than their phonemic inventory. Crucially, however, the algorithm requires that listeners hear all features in the intake form — it assumes full perception. The experimental part of this study sets out to test this full perception hypothesis by examining to what extent various features are perceived cross-linguistically in a phonetic vs. phonemic processing mode.

30 Note the discrepancy in distance between target theta and [t] versus [s̺] (5 and 2 respectively). For this reason, one would expect the stop to be a poor substitute for theta in Japanese.

6. EXPERIMENTAL DESIGN

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

6.1 PHONETIC VS. PHONOLOGICAL PROCESSING

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

Acoustic: Processing of fine, non-linguistic distinctions; i.e. the 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 paper are the phonetic and phonemic (i.e. phonological) levels. Werker and 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.[31] 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.

In accordance with Werker and Logan (1985), 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 acoustic signal of the first stimulus has faded. It is hypothesized that what fades are the non-contrastive, phonetic features. By contrast, in the Short ISI condition, the interval between stimuli is short enough for both contrastive and non-contrastive, phonetic features to be preserved. In a within-subjects design, Werker and Logan found phonetic processing at a 250msec ISI and evidence of phonemic processing at

31 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.

1500msec ISI.[32] These are the intervals I use to evoke phonetic and phonological processing respectively.

Importantly, I assume that in discrimination tasks such as the AXB task, what is being compared are features. In other words, an intake form is compared with another intake form. This contrasts with what occurs in transfer. In the latter case, an intake form is compared with L1 segmental representations.

6.2. PARTICIPANTS

The 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.[33]

All learners in this study began learning English after the age of seven, the purported critical period for phonological acquisition (Scovel 1988); thus they are *not* considered to be native bilinguals 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.[34] All participants had normal hearing according to self-report.

6.3. TEST DESIGN

6.3.1 TASKS

The experimental paradigm used was an AXB forced choice oddity 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, for example, $\text{thigh}_1\text{thigh}_2\text{tie}$

A X B

32 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.

33 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.

34 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.

The participant's task is to determine whether A=X or whether B=X. If A=X, the subjects were to press button A. If B=X, the subjects were to press button B.

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 and Gottfried 1995). In the AXB task, this does not occur because the target is in the middle. Beddor and 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 because of a difference in volume, not a difference in features. With an AXB task, this pitfall is largely avoided. Finally, the AXB task is shorter than the 4IAX task — the former requires three stimuli per item, while the latter requires four, while retaining the advantages of the 4IAX (e.g. low response bias). Considering the large amount of contrasts examined in the present study, this was a necessary consideration.

Importantly, I assume that in discrimination tasks such as the AXB task, transfer does not occur. In other words, the features of an intake form are compared with those of another intake form. For example, in QF, [t̪] is substituted for [θ] in L2 speech; however, when the participant is presented with an item such as [sɑt] [sɑt] [θɑt] in an AXB task, the features of [s] are compared with those of [θ], not with those of [t̪] — the comparison is between [s] and [θ], not between [s] and [t̪].

Participants heard 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.

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 was presented in its entirety before the Short ISI. The reason for this was that I wanted to ensure phonological processing in the Long ISI condition. Werker and 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 and 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 and 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. 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 6800 using Grado SR60 headphones. Each day, they began with four practice items. Then participants were presented blocks of 135 items. Items were randomly ordered within each block. 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 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. 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 will be reported in this study.[35] At the end of the five-day testing period, L2 learners participated in an oral interview. This interview was recorded. The interview elicited a subject profile and provided a debriefing. Towards the end of the interview, the goal of the experiment was explicitly discussed, in the hopes of eliciting a higher number words containing [θ]. Production results have not been completely analyzed yet.

6.3.2 STIMULI

Table 10 gives a list of test contrasts presented in the experiment. The first column gives the feature or features which serve to contrast two segments. Both voiceless and voiced contrasts were presented in the experiment, but I am only reporting the results for the voiceless contrasts. This is because there is reason to believe that voiceless and voiced segments are not treated equally by learners (in both L2 and L1). For example, it seems that continuancy is more salient on voiceless fricatives than it is on voiced fricatives. This may lead to differential substitution depending on voicing, for

35 This is because of a fault in the methodology. Participants had the option of responding as soon as they heard X or of 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.

example, stops being substituted for voiced fricatives, but Continuant being maintained on voiceless fricatives. Perception of voiced contrasts will be left for further research. Filler items were also included in the experiment at a ratio of approximately 1:2 with the test items. This was done to provide distracters in order that the participants would not become cognizant of the purpose of the experiment. I will not report on filler items in this paper. All stimuli appeared in simple Onset position of monosyllables. The vowels in these monosyllables were divided into three groups: A = low vowels; I = front (non-low) vowels; U = back (non-low) vowels.

Table 10 shows that the majority of contrasts which were tested differed by a single feature. By manipulating one feature, we avoid introducing confounding factors. This enables us to have a clearer picture of the precise features which learners can perceive.

TABLE 10. TEST CONTRASTS[36]

TEST ITEMS					
Distinguishing Feature(s)	**Segments**	**Number of tokens per ISI by vowel**			
		A	**I**	**U**	**Total tokens**
Lip vs. Dental	[f] vs. [ɸ]	9	9	6	24
Labial vs. Coronal	[f] vs. [θ]	9	6	15	30
Stop vs. Continuant	[t̪] vs. [θ]	18	12	6	36
Strident vs. Mellow	[s̪] vs. [θ]	21	18	9	48
Stop vs. Continuant and Alveolar vs. Dental and Apical vs. Laminal	[t] vs. [θ]	27	30	6	63
Alveolar vs. Dental and Apical vs. Laminal	[t] vs. [t̪]	30	27	21	78

6.3.2.1. STIMULI QUALITY

All stimuli respect English syllable structure and phonotactics.[37] Both words and non-words were tested in the experiment. The real words incorporated in the study

36 The contrast [t̪] vs. [t̪s̪] was also tested. However, I do not report on this contrast partly because of indeterminacy in the correct representation of affricates and partly in the interest of space. Note, however, that performance on this contrast was at ceiling.

37 Note however that many of the test contrasts involve non-English segments. This was necessary in order to test my hypotheses.

are all high frequency as rated by seven native English speakers on a 5-point scale.[38] Non-words were checked by three linguists. They were rejected if they were homonyms, or close homonyms, to a real English word or if they were similar to a socially unacceptable real word.[39]

Because of the limited number of participants in this study, several tokens of each type were constructed in order to increase the reliability of the results. Four factors were manipulated in the construction of the stimuli: Onset, Vowel, Rhyme, and Wordhood. Only the first two factors were analyzed for this paper.[40]

1. Onset — All onsets were simple. The segment of interest in the experiment, [ɸ f θ t̪ s̪ t] or one of the fillers, was always located in onset position.

2. Vowel — Either low (A), front (I), or back (U). Front/Unround (I): [ij ɪ ej ɛ]; Back/Round (U): [uw ow]; Low (A): [æ ʌ ɑ aj aw].[41] 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 vowels are round. The diphthongs [aj aw] were classified as Low according the category of the head. In general, true diphthongs were avoided. However, they were used in real words to increase the number of stimuli. Diphthongs introduce a potential confounding factor, given that they involve several features, for example, Low and Front and High in [aj]. Vowel quality was tested to see if it has an influence on the perception of word-initial contrasts. Previous research, for example, 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.

3. Rhyme — The rhyme was always heavy, respecting word minimality. Thus the syllable was either:

a. Open (V) with a diphthong (vowel + glide); or
b. Closed (C) by a consonant or consonant cluster.[42]

The codas used were [p k m n]. The consonants [m n] were selected because nasals are favoured over other segments in coda position due their relatively high sonority. 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

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

39 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.

40 Preliminary examination indicates no difference between words and non-words; thus these conditions were grouped together for the present analysis.

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

42 The effect of open vs. closed syllables was not analyzed for this paper.

complexity and frequency.[43] In the non-word condition (see (4) below), only simple codas were used; however, in word contrasts, it was necessary to allow for branching codas. Words beginning with [θ] 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 indeed.

4. Wordhood — Word (W) or non-word (NW). Wordhood was consistent within each item triad. For example, in the contrast [t] vs. [θ], both items were either words, for example, *taught* vs. *thought*, or non-words, for example, *teep* vs. *theep*.

6.3.2.2. STIMULI QUANTITY

Two stimuli are considered to be of the same type if they agree on all four factors named above. For example, *teak, team, teen, tip, tick, Tim, tin, tape* are all of the same type: Onset [t], Front vowel (I), Closed syllable (C), Word (W). Each type had 6–8 tokens. These tokens were either identical or non-identical. Note that identical tokens were not physically identical (see Section 6.3.2.3). The number of tokens constructed depended on the contrast in which they appeared.

<u>Non-identical Tokens of the same Type</u>. The set [*teak, team, teen, tip, tick, Tim, tin, tape*] consists of eight non-identical tokens.

<u>Identical Tokens of the same Type</u>. In the real word condition, the construction of non-identical 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. [s̪] involving types [θ] A V W vs. [s̪] A V W, two tokens could be constructed, *thigh* vs. *sigh* and *thaw* vs. *saw*. However, for the contrast [θ] vs. [t], involving types [θ] A V W vs. [t] A V W, only one token could be constructed for each type, *thigh* vs. *tie*. The word *thaw* has a counterpart, *taw*, but it is of low frequency and probably unfamiliar to participants.

An identical token was not repeated more than six times. This was done in order to avoid repetition effects. Thus, for types in which only one token could be constructed for a given contrast, this token was repeated to a maximum of six times. For example, the contrast *thigh* vs. *tie* was repeated six times. However, where two non-identical tokens could be constructed, these were repeated to a maximum of eight tokens per type. For example, the contrasts *thigh* vs. *sigh* and *thaw* vs. *saw* were each repeated four times, giving a total of eight tokens for the types [θ] A V W vs. [s̪] A V W.

6.3.2.3. RECORDING OF STIMULI

Stimuli were recorded by three trained phonologists (talkers) in a professional sound

43 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.

studio (DNA Studio) using a Tascan DA 30 DAT recorder and AKG 414 EB microphone with cardiod pattern setting and pop filter. 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; 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 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 Section 6.1 that the acoustic mode involves processing non-linguistic distinctions; the listeners assess physical identity, for example, fundamental frequency, amplitude. This is the type of processing we use in discriminating two bell tones, for example. If listeners were making decisions using an acoustic factor, we could say nothing about speech perception; thus, it was primordial that we tap a linguistic mode of processing.

The data were transferred into SoundEdit 16.01 (Macromedia Sound Team 1994) at a sampling rate of 22kHz. The experimental paradigm was constructed using PsyScope 1.2 (Cohen et al. 1993). The order of presentation of stimuli within the AXB triads was counterbalanced: AAB, ABB, BBA, and BAA. For example, in the AAB order, participants would hear *thigh, thigh, tie*, while in the ABB order, they would hear *thigh, tie, tie*. This was done because it has been shown that the perception of one stimulus can affect the perception of the following stimulus (Polka and Werker 1994). However, order of presentation has not yet been analyzed.

6.4. 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 features listeners can perceive in both a phonological and phonetic processing mode. In order to determine this, we need a working definition of what it means to "perceive features". In her dissertation, Brown (1997:161) uses an 83% perceptibility criterion. Following Brown, I also introduce a criterion, set at a slighter higher rate of 85% accuracy (15% error rate), in order to consider that a featural contrast is perceived.

1. [f] vs. [ɸ]
   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. Since I hypothesize that listeners perceive all features in a phonetic processing mode, in the Short ISI (phonetic) condition, all groups should perceive this contrast.

2. [f] vs. [θ]
   The features involved in this contrast — Labial vs. Coronal — are distinctive for all languages.[44] Thus it is predicted that participants will reach criterion in both conditions.

3. [t̪] vs. [θ]
   The features contrasting these segments — Stop vs. Continuant — are contrastive in all languages tested.[45] Again, participants should perceive this contrast in both conditions.

4. [s̪] vs. [θ]
   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; however, according to my hypothesis, they should be able to hear the difference in the phonetic condition.

5. [t] vs. [θ]
   This contrast is related to the contrast in (3) above, [t] vs. [θ]. However, in addition to testing Stop vs. Continuant, it also examines Alveolar vs. Dental and Apical vs. Laminal. In the Long ISI, listeners should perceive this contrast based on Stop vs. Continuant. In the Short ISI, discriminability should be further enhanced by Alveolar/Dental and Apical/Laminal.

6. [t] vs. [t̪]
   The features involved in this contrast — Alveolar vs. Dental and Apical vs. Laminal — are not distinctive in any of the languages investigated.[46] It is predicted that this contrast will not be perceived at the phonological level, but that it will be heard at the phonetic level.

Recall that I am assuming the AXB task involves an intake-to-intake comparison (see Section 6.3.1). In other words, an AXB comparison is made in one step as:

$A \leftrightarrow B \leftrightarrow B$, not in two steps as:

$$\begin{array}{ccc} A & B & B \\ \updownarrow & \updownarrow & \updownarrow \\ C \leftrightarrow & D \leftrightarrow & D \end{array}$$

If the latter situation were the case, one would expect absolutely no difference be-

44 Brown (1997) 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.

45 [t̪] was produced by the talkers as Laminal; thus both the [t̪] and [θ] stimuli share this feature.

46 Kenstowicz (1994:30) presents arguments that English makes the [θ] vs. [s] distinction based on Laminal vs. Apical. If this is the case, then NE should hear the Laminodental vs. Apicoalveolar contrast in the Long ISI condition. As we shall see later, they fail to do so. Note that in EF and QF, both /s/ and /ʃ/ are laminal. This suggests that Laminal and Apical do not serve to make phonemic contrasts in French.

tween [θ s̪ s̪] and [s̪ s̪ s̪]. In effect, since it is hypothesized that participants will perceive non-contrastive features in a phonetic processing mode, and thus we would expect differences in error rates between [θ s̪ s̪] and [s̪ s̪ s̪].

Thus, in terms of auditory distance, it is predicted that the test contrasts will be perceived as follows, from worst to best. Auditory distance is indicated in parentheses. *Phonological condition*: f-ɸ, s̪-θ, t-t̪ (0) > f-θ, t̪-θ, t-θ (4). *Phonetic condition*: f-ɸ (2) > t̪-θ, s̪-θ (3) > f-θ, t-t̪ (4) > t-θ (8). Let us now turn to the results of this investigation.

7. Results

Each participant's error rate was calculated for each contrast. A mixed Analysis of Variance (ANOVA) was run on the data.

7.1 Main Effects

This section reports the results for each factor, collapsing the effects of other factors.

7.1.1. Language

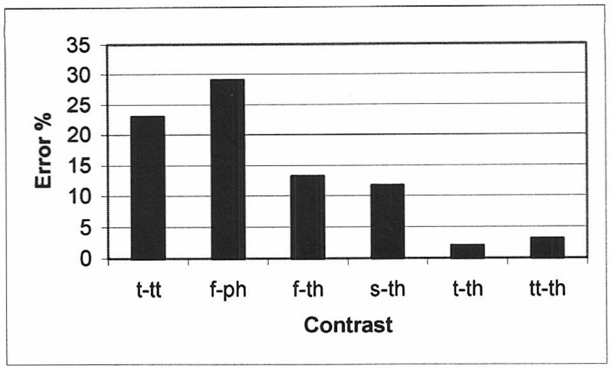
There was no main effect for Language. This means that, overall, one language group did not perform significantly worse or better than any other group.

7.1.2. ISI

There was no main effect for interstimulus interval. This means that the ability to perceive contrasts was the same in both the Long and Short ISI conditions across contrasts and across vowels.

7.1.3. Contrast

There was a significant main effect for Contrast, [$F(6,90) = 147.43$, $p < .00001$]. This means that certain contrasts are easier to perceive than others if we collapse language groups, ISI, and vowel conditions.



Key: t-tt = [t] vs. [t̪]; f-ph = [f] vs. [ɸ]; f-th = [f] vs. [θ]; s-th = [s̪] vs. [θ]; t-th = [t] vs. [θ]; tt-th = [t̪] vs. [θ].

Figure 8. Overall Error Percentage by Contrast.

Figure 8 reveals that, overall, performance was poorest on the [f] vs. [ɸ] contrast and best on the [t] vs. [θ] contrast.

7.1.4. VOWEL

There was a significant main effect for Vowel, [$F(2,30) = 16.89$, $p < .0001$]. This means that contrasts are more difficult to perceive in the context of certain vowels if we collapse language groups, contrasts, and ISI. 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.

7.2. INTERACTIONS

This section reports the interactions between factors.

7.2.1. LANGUAGE X ISI

There was no interaction between Language and ISI.[47] This means that behaviour in each of the ISI conditions was not dependent on language group. Recall that the ISI condition was intended to isolate phonemic vs. phonetic processing. Given that there was no main effect and no interactions for ISI, this leaves us with the question: What level of processing was accessed during this experiment? In order to answer this question, the following procedure was conducted.

47 Nor was there any interaction between Contrast and ISI.

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 listeners were *not* using an acoustic mode (see Section 7.2.2.).) Two contrasts will be focused on to address this issue: [ɸ] vs. [f] and [t] vs. [t̪]. The features involved in these contrasts are non-distinctive in all the languages examined. Also, Figure 8 has shown that these two contrasts had the highest overall error rates.

If listeners are using a phonetic mode, they should be able to discriminate the above contrasts at a rate which is significantly better than chance. To verify this, I conducted t-tests comparing the means for each language group against a hypothesized chance score of 50%.

For the contrast [ɸ] vs. [θ], a one-tailed *t*-test comparing the average of each language group with a hypothesized chance score of 50% was significant, $t = 12.48$, $p < .001$. This means that the performance of all language groups on this (non-distinctive) contrast was significantly better than chance. For the contrast [t] vs. [t̪], again the *t*-test was significant, $t = 18.97$, $p < .001$, indicating that participants were discerning this contrast at a significantly better than chance rate. Thus, on these non-phonemic contrasts, participants had accuracy rates which were significantly better than 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.[48]

We must now ask the question: Why was there no difference between the ISI conditions? It seems that this study has run into the same problems as were encountered by Werker and Logan (1985) in a comparable within-subjects design. Since Werker and 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 phonemic strategy.[49] This turned out to be wrong. It seems that the AXB paradigm is not overly conducive to phonemic processing.[50]

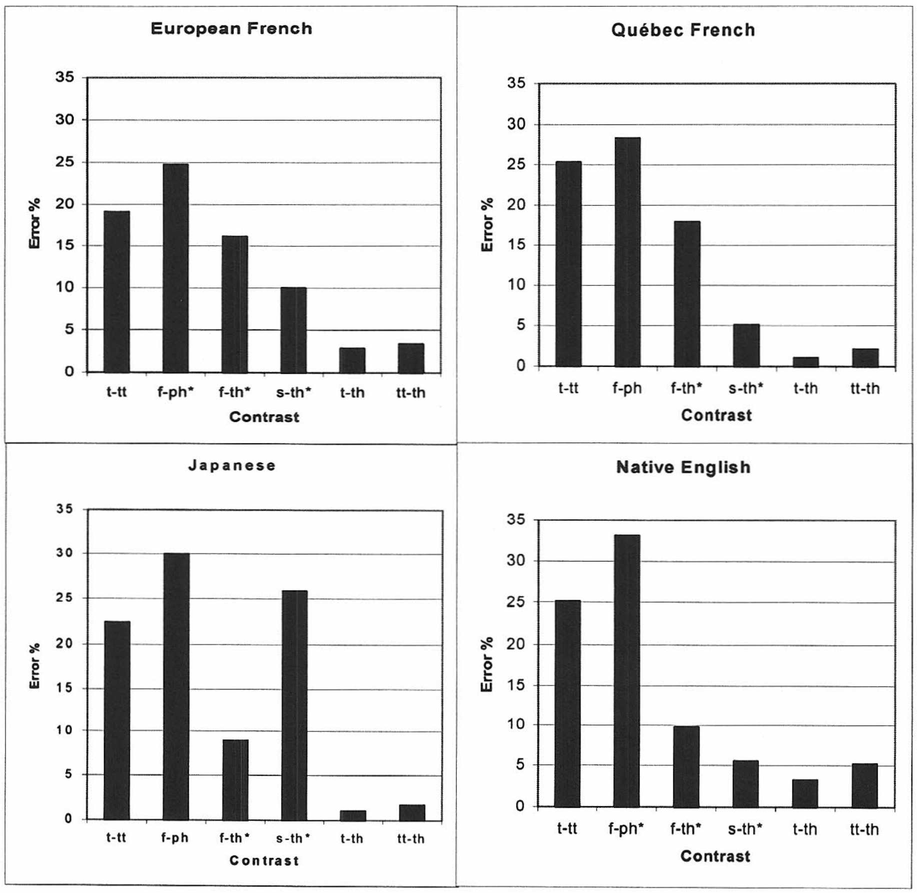
48 Even though I use a 50% criterion to determine level of processing, I use the 85% perceived criterion discussed in Section 6.4 in order to interpret the data in a categorical rather than gradient manner.

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.

50 Curtin, Goad, & Pater (1998) encounter the same difficulties with the ABX paradigm.

7.2.2. LANGUAGE X CONTRAST [51]

There was a significant interaction between Language and Contrast, $[F(18,90) = 7.9, p < .00001]$. Figure 9 below gives the error percentages on the test contrasts for each language group.



Key: * = differs significantly from the same contrast in at least one other language. t-tt = [t] vs. [t̪]; f-ph = [f] vs. [ɸ]; f-th = [f] vs. [θ]; s-th = [s̪] vs. [θ]; t-th = [t] vs. [θ]; tt-th = [t̪] vs. [θ].

Figure 9. Error percentages by contrast for each language group

51 There was also a significant interaction between Contrast and Vowel, $[F(12,180) = 10.4, p < .00001]$. These interactions indicate that a following vowel influences the ability to perceive differences between fricatives. In the interest of space, I will not be discussing this interaction.

As can be seen in Figure 9, there are differences in the ability to perceive certain contrasts depending on the native language. Most strikingly, posthoc tests showed that the Strident vs. Mellow contrast ([s̪] vs. [θ]) is more difficult for the JA listeners than it is for the other language groups. Let us examine each contrast individually. I will begin with the [t] vs. [t̪] contrast.

[t] vs. [t̪] — Alveolar vs. Dental and Apical vs. Laminal: This contrast was relatively difficult to discriminate for all groups — accuracy rates were below criterion; that is, below 85% accuracy or above 15% error rate. This is contra the full perception hypothesis.

Table 11.
Results According to 15% Criterion for Apicoalveolar vs. Laminodental

Language	Perceived? (Error rate)
EF	No (19%)
QF	No (25%)
JA	No (22%)
NE	No (25%)

[f] vs. [ɸ] — Lip vs. Dental: This contrast was the most difficult to discriminate for all groups. No group reached criterion, contra the full perception hypothesis.

Table 12. Results According to 15% Criterion for Lip vs. Dental

Language	Perceived? (Error rate)
EF	No (25%)
QF	No (28%)
JA	No (30%)
NE	No (33%)

[f] vs. [θ] — Labial vs. Coronal: Only JA and NE reached criterion on this contrast; these listeners also performed significantly better than EF and QF on this item. This will be discussed in Section 7.5.1.

Table 13. Results According to 15% Criterion for Labial vs. Coronal

Language	Perceived? (Error rate)
EF	No (16%)
QF	No (18%)
JA	Yes (9%)
NE	Yes (10%)

[s̪] vs. [θ] — Strident vs. Mellow: This contrast reveals the largest difference between language groups in this study. Japanese is the only group that did not reach criterion on this contrast.

Table 14. Results According to 15% Criterion for Strident vs. Mellow

Language	Perceived? (Error rate)
EF	Yes (10%)
QF	Yes (5%)
JA	No (26%)
NE	Yes (6%)

[t̪] vs. [θ] — Stop vs. Continuant: All groups did well on this contrast.

Table 15. Results According to 15% Criterion for Stop vs. Continuant

Language	Perceived? (Error rate)
EF	Yes (3%)
QF	Yes (2%)
JA	Yes (2%)
NE	Yes (5%)

[t] vs. [θ] — Stop vs. Continuant, Apicoalveolar vs. Laminodental: All groups easily reached criterion on this contrast. Given that no group perceived Laminodental vs. Apicoalveolar in the [t] vs. [t̪] contrast, we must attribute the perceptual discrimination of [t] vs. [θ] to Stop vs. Continuant.[52]

Table 16. Results According to 15% Criterion for Stop vs. Continuant; Apicoalveolar vs. Laminodental

Language	Perceived? (Error rate)
EF	Yes (3%)
QF	Yes (1%)
JA	Yes (1%)
NE	Yes (3%)

7.2.3. SUMMARY OF INTERACTIONS.

Results suggest that, even with a long interstimulus interval, the AXB experimental paradigm is most successful in conducing phonetic processing. Despite indications

52 Note that the fact that JA listeners perceive this contrast while they fail to perceive the [s̪] vs. [θ] contrast strongly suggests that these pairs do not involve the same features, contrary to LaCharité & Prévost's claim (see Section 2.4).

that the task tapped a phonetic level of representation, outcomes from the test contrasts show that the full perception hypothesis is not entirely supported. Importantly, however, some listeners are able to perceive features which are *not* contrastive in their L1. In particular, results on [s̪] vs. [θ] have shown that both EF and QF groups were able to perceive the difference between Strident vs. Mellow, features which do not function phonemically in the native language.

New Phonetic Representations Based on Results

As discussed in the preceding section, not all features were perceived in a phonetic processing mode. Section 4 of this paper showed phonetic representations for the learner languages based on full perception. In this section, these phonetic representations are revised to reflect the results of the experiment.

European French Results: The results from Tables 12 and 13 above suggest that EF listeners fail to hear the Apicoalveolar vs. Laminodental and Lip vs. Dental contrasts. This scenario means that EF has the following representations.

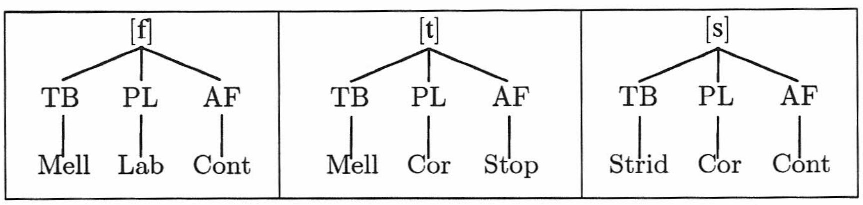


Figure 10. EF Representations for Potential Substitutes Based on Results

Québec French Results: The results in Tables 12 and 13 show that QF listeners also fail to distinguish between the features Apicoalveolar vs. Laminodental and Lip vs. Dental. The representations based on QF results are given below.

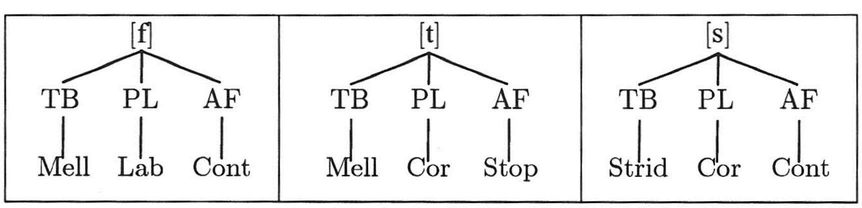


Figure 11. QF Representations for Potential Substitutes Based on Results

Japanese Results: Again, Tables 12 and 13 indicate that JA listeners fail to perceive Apicoalveolar vs. Laminodental and Lip vs. Dental. In addition, these listeners were unable to distinguish between Strident and Mellow. The new representations for JA segments are given below.

sentations for JA segments are given below.

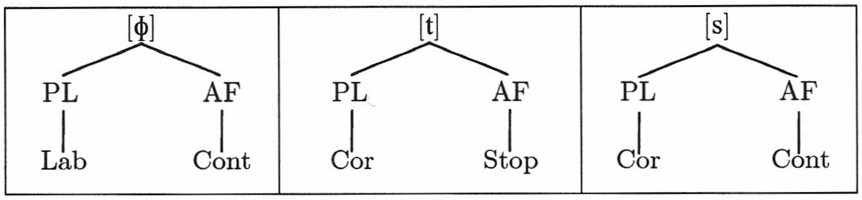


Figure 12. JA Representations for Potential Substitutes based on Results

7.4. REAPPLYING THE ALGORITHM

In Section 5.2, the metric of auditory distance was applied to fully specified representations. Given that the results of this experiment suggest that learners were not able to perceive all features, this section reapplies the auditory distance algorithm using the revised representations presented in the preceding section.

Before we proceed, let us keep in mind two assumptions: First, if a feature does not reach the perceptibility criterion, I assume it is absent from both the intake and representation. Second, when a feature is absent from the intake and representation, it cannot participate in muting.[53] Thus, features present in the intake as well as those in L1 representations will vary across language groups. Let us now reapply the transfer algorithm based on just those features that are perceived by the participants (see Section 3.5.5).

Both European and Québec French listeners displayed similar behaviour: both fail to perceive all Location and Minor Articulator features. This goes against the hypothesis that phonetic differences in the representation of coronal obstruents for these languages would lead to different perceptual acuities.

If we reapply the algorithm based on this information, we arrive at the following transfer order for both dialects of French: [t] > [ʃ] ~ [s]. Since [t] comes out as the most favoured substitute, this fits well with reported facts for QF; but not for those of EF, where [s] is most commonly reported. The phonetic predictions, based on full perception, gave the order: [s̪] > [ʃ] > [t̪] for EF (Table 6), while those for QF (Table 7) gave the order: [ʃ] > [t̪] > [s].

For JA, because they only perceive Stop, Continuant, and Major Articulator features, the algorithm gives the following order: [s] > [ɸ] ~ [t̪]. The predictions for JA in Table 8 were [s̪] > [t] > [ɸ]. The most frequent substitute reported in the literature for JA is [s].

Before discussing these calculations further, I would like to point out an asymmetry in the data which requires that we consider an additional muting effect.

53 These assumptions were implicit in the calculation of phonemic predictions in Table 4.

7.5. ANOTHER LOOK AT MUTING EFFECTS

7.5.1. AN ADDITION

The most robust finding in this study is that JA listeners fail to perceive the Strident vs. Mellow contrast. This starkly contrasts with the other languages groups for whom this contrast was relatively easily perceived. Another finding of interest is the behaviour of JA listeners with respect to the [f] vs. [θ] (Labial vs. Coronal) contrast in comparison with EF and QF. While JA listeners perform poorly on [s̪] vs. [θ] as compared to EF and QF, their performance on [f] vs. [θ] is significantly better than that of EF and QF (see Table 13). Interestingly, my results for JA partially replicate those of Hancin-Bhatt (1994). Recall that she examined Japanese, German, Hindi, Turkish, and Native English L1s. In her perceptual identification task, JA listeners had the highest error rates on [s] vs. [θ], but they had the lowest rate of confusion for [f] and [θ] as compared to the other groups.

I propose that this asymmetry is due to the absence of Mellow in the JA grammar. In other words, I would like to suggest a link between the ability to perceive Strident/Mellow vs. Labial/Coronal on fricatives. Strevens (1960), from his examination of the acoustics of fricatives, considers that [ɸ], [f], and [θ] form an acoustic natural class in being equivalently low in intensity. Noting the difficulty in determining which factors identify members within the group, he proposes that place distinctions are based on "centre of gravity"; in other words, the location of spectral peaks. The location of peaks in [f] and [θ] are quite similar. The fricatives [s] and [ç] also form an acoustic natural class: according to Strevens, they are high intensity (strident) fricatives. In contrast to [f] and [θ], place cues on [s] and [ç] are quite distinct. It seems, therefore, that place of articulation is less salient on mellow fricatives. I suggest that this relationship has a formal expression: Mellow + Continuant mute Major Articulator (e.g. Labial and Coronal). This means that Labial and Coronal will be less salient in conjunction with Mellow and Continuant. Thus, the features Coronal on [θ] and Labial on [f] would both receive a value of 1.

Japanese speakers do not hear the Strident vs. Mellow contrast; thus, neither of these features is specified in their surface intake or in their L1 grammar. Since there is no Mellow to mute Labial and Coronal, these place features are more salient for JA listeners, and the contrast is more readily discerned. On the other hand, both EF and QF hear Mellow, so the presence of Mellow in their grammars reduces the salience of place cues that differentiate [f] and [θ] for these listeners. This means that the grammar dictates which cues the listener focuses on: for JA listeners, it is place (Major Articulator) and for EF and QF listeners, it is intensity (Mellow).

As Table 17 shows, the auditory distance between [f] and [θ] is reduced for EF and QF. This results in [f] being chosen as the best substitute for [θ] in both EF and QF. This will be further discussed in Section 7.6.

Table 17. Algorithmic Calculation for [f] in EF and QF Based on Mell + Cont Mutes MajArtic.

Potential Substitute	*Matches*	*Mismatches*	*Distance*
f	+2 (Mell) +1 (Cont)	−1 (Lab) 2	2 away

In closing this discussion on muting of place features, notice that NE listeners, although they have Mellow in their grammar, pattern with JA in having relatively low error rates on this contrast. Perhaps this is due to the fact that the [f] vs. [θ] segmental contrast exists in the L1 grammar. There may be a constraint which ensures that the features upon which a native contrast relies remain salient.

The question remains, however, as to why EF and QF perceive the Strident/Mellow contrast, while JA listeners do not. As noted in Section 4.3, I consider that Japanese /s/ is a low-intensity (mellow) fricative. This would mean, therefore, that grammars with low-intensity fricatives only are permanently underspecified for Mellow and Strident.[54]

7.5.2. EUROPEAN AND QUÉBEC FRENCH: [s] OR [t]?

Let us now examine how the French listeners perceived the test contrasts [s̪]-[θ] and [t̪]-[θ]. During the recording of stimuli, all these segments were produced as Dental and Laminal. As such, [s̪]-[θ] differs on the Strident/Mellow dimension; while, [t̪]-[θ] differs on the Stop/Continuant dimension. Both contrasts are subject to muting effects: Dental mutes Strident being relevant to the former, and Mellow mutes Continuant relevant to the latter. If it can be shown that a language is specified for Strident, Mellow, and Dental, then we would expect similar error rates on both [s̪]-[θ] and [t̪]-[θ]. Recall that while Strident and Mellow are specified in EF and QF, Dental is not. Because Dental is absent, Dental mutes Strident is inactive, and as a result, French listeners should perceive [s̪]-[θ] better than [t̪]-[θ]. However, Figure 9 shows this is not the case. There is no significant difference in error rates for [s̪]-[θ] vs. [t̪]-[θ] for either EF or QF.[55] These results suggest that Mellow mutes Continuant is not active in these grammars either. This may be because the learners in this study are at an advanced level of L2 English. It is possible that Mellow mutes Continuant is active at earlier stages of acquisition, but is overridden at more advanced stages. In other

54 This would predict free variation on the stridency dimension (see Keating 1988).

55 Note however that, although statistically insignificant, EF listeners have higher error rates on the [s̪]-[θ] contrast than do QF listeners. This may be related to the fact that the EF group showed comparatively better (though again statistically insignificant) performance on the Dental/Alveolar contrast ([t]-[t̪]). Because there is a tendency for EF listeners to hear Dental better than other groups, this feature may be exerting some influence, muting Strident and thus making [s̪] perceptually closer to [θ] for this group.

words, the auditory distance between [t̪]-[θ] may be less in earlier as compared to later stages of acquisition. This will be tested in future research with beginning learners.

In conclusion, results show that the substitute [f] has a perceptual basis for EF and QF; however, differential substitution in these two languages does not have a strong perceptual foundation, at least for learners at an advanced level of L2 acquisition. Revised calculations give the following order of transfer; auditory distance is given in parentheses (see Appendix D for details): EF and QF: [f] (2) > [t] ~ [s] (4); JA: [s] (0) > [ɸ] ~ [t] (4).

7.6. Discussion and Conclusion

The main hypothesis which instigated this work is that the choice of substitute for the target interdental is based on an auditory phonetic comparison of a fully specified intake (surface form) with native language internal representations. This hypothesis is only partially supported. Not all non-contrastive features were perceived: no group perceived the features Lip, Dental, Alveolar, Apical, or Laminal. However, the non-contrastive features Mellow and Strident were perceived by EF, QF, and NE listeners. It is proposed that Mellow is involved in muting effects which result in [f] being auditorily close to [θ], and thus the best choice for interdental substitute for EF and QF. The fact that EF and QF perceive the non-contrastive features Strident and Mellow accounts for why [s] is not perceptually merged with [θ] to the extent that it is for JA listeners. However, as we have seen, there is no significant auditory difference between [s] and [t] for either EF or QF; thus, phonetic features do not play a determining role in the choice between coronal stop or fricative as substitute. Recall that a phonemic assessment (based on contrastive features alone) cannot capture the QF facts either. The failure to find significant perceptual differences between [s] and [t] for EF and QF listeners may be due to the fact that these participants are advanced learners of L2 learners. Further research with beginning learners may reveal an auditory phonetic basis for differential substitution for these language groups.

The results of the algorithmic calculations presented at the end of Section 7.5.2. are consistent with the perception and production facts for JA. However, in both EF and QF, [f] is selected as the best perceptual substitute, yet [f] is not the most commonly reported production substitute in either EF or QF. I suggest that the reason for the relatively few reports of [f] in the literature (other than the paucity of research), may be that visual cues to the labiality of [f] and the coronality of [θ] are available for encoding in the representations for these segments. This predicts that French learners who acquire English using auditory materials alone would have a propensity to substitute [f] more frequently in perception and production.

In this paper, it has been proposed that the failure to perceive the difference between [s̪] and [θ] is due to a lack of Strident and Mellow in the grammar. It has further been suggested that this arises when the coronal fricative in a language can be phonetically analyzed as being low-intensity. In such a language, one would expect place cues to be more salient for the listener than intensity cues. As a consequence, listeners from such a language would be predicted to perceive the difference between [f] and [θ] relatively well. In Brown's (1997) study, Mandarin and Korean listeners performed on a par with JA listeners in failing to perceive the [s] vs. [θ] distinction. These languages should be investigated regarding the phonetic properties of their [s] and the performance of listeners on the [f] vs. [θ] contrast.

In conclusion, the results of this study coupled with the auditory distance model indicate that when target [θ] 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, muting effects involving these features play a role in choice of substitute. Importantly, EF and QF do perceive Strident and Mellow, despite the fact that these features do not function contrastively in either language. This finding goes against Brown's 1997 claim that features which are non-contrastive in a language will not be perceived either phonetically or phonologically.

REFERENCES

Archangeli, D.1984. Underspecification in Yawelmani phonology and morphology. Doctoral Dissertation. MIT.

Beddor, Patrice S. and Terry L. Gottfried.1995. Methodological issues in cross-language speech perception research with adults. In *Speech perception and linguistic experience: Issues in cross-language research,* ed. Winifred Strange, Baltimore: York Press.

Berger, Marshall D.1951. The American English pronunciation of Russian immigrants. Doctoral Dissertation. Columbia University.

Best, Catherine T. and Winifred Strange.1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20:305-330.

Brannen, Kathleen.1998. The acquisition of interdentals in L2. Paper read at Montréal-Ottawa-Toronto Phonology Workshop, Ottawa.

Brown, Cynthia A.1993. The role of the L1 grammar in the L2 acquisition of segmental structure. *McGill Working Papers in Linguistics* 9: 180-210.

Brown, Cynthia A.1997. Acquisition of segmental structure: Consequences for speech perception and second language acquisition. Doctoral Dissertation. McGill University.

Calabrese, A.1988. Towards a theory of phonological alphabets. Doctoral Dissertation. MIT.

Charbonneau, René and Benoît Jacques.1972. [ts] et [dz] en français canadien. In *Papers in linguistics and phonetics to the memory of Pierre Delattre*, ed. A. Valdmen. The Hague: Mouton.

Chomsky, Noam & Morris, Halle.1968. *The sound pattern of English.* Cambridge, MA: MIT Press.

Cohen, J.D., B. MacWhinney, M. Flatt, and J. Provost.1993. PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, & Computers* 25:257-271.

Curtin, Suzanne, Heather Goad and Joe Pater.1998. Phonological transfer and levels of representation: The perceptual acquisition of Thai voice and aspiration by English and French speakers. *Second Language Research* 14: 389-405.

Dart, Sarah N.1991. Articulatory and acoustic properties of apical and laminal articulations. Doctoral Dissertation. *UCLA Working Papers in Phonetics* 79:1-155.

Dickerson, W.1976. The psycholinguistic unity of language learning and language change. *Language Learning* 26:215-231.

Flege, James Emil.1980. Phonetic approximation in second language acquisition. *Language Learning* 30:117-134.

Flege, James Emil.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, James Emil.1995. Second language speech learning: Theory, findings, and problems. In *Speech perception and linguistic experience: Issues in cross-language research*, ed. Winifred Strange. Baltimore: York Press.

Flemming, Edward S.1995. Auditory representations in phonology. Doctoral Dissertation. UCLA.

Gatbonton, Elizabeth.1983. Patterned phonetic variability in second language speech: A gradual diffusion model. In *Second language learning: Contrastive analysis, error analysis, and related aspects*, ed. Betty Wallace Robinett and Jacquelyn Schachter, 240-255. Ann Arbor: University of Michigan Press. [1978.]

Gnanadesikan, Amalia E.1993. The geometry of coronal articulations. In *North Eastern Linguistic Society 24 Proceedings*, 125-139.

Gorecka, Alicja.1989. *Phonology of articulation.* Doctoral Dissertation. MIT.

Hall, T. Alan.1997. *The phonology of coronals.* Amsterdam/Philadelphia: John Benjamins.

Halle, Morris.1959. *The sound pattern of Russian.* The Hague: Mouton.

Hamilton, Phillip J.1996. Phonetic constraints and markedness in the phonotactics of Australian aboriginal languages. Doctoral Dissertation. University of Toronto.

Hammarberg, Björn.1997. Conditions on transfer in phonology. In *Second-language speech: Structure and process*, ed. Allan James and Jonathan Leather, 162-180. Berlin/New York: Mouton de Gruyter.

Hancin-Bhatt, Barbara Jean.1994a. Phonological transfer in second language perception and production. Doctoral Dissertation. University of Illinois.

Hancin-Bhatt, Barbara Jean.1994b. Segment transfer: A consequence of a dynamic system. *Second Language Research* 10:241-269.

Hardison, Debra M.1999. Bimodal speech perception by native and nonnative speakers of English: Factors influencing the McGurk Effect. In *Phonological issues in language learning*, ed. Jonathan Leather. Malden: Blackwell.

Hume, Elizabeth and Keith Johnson.2001. A model of the interplay of speech perception and phonology. In *The role of speech perception in phonology*, ed. Elizabeth Hume and Keith Johnson. New York: Academic.

Ioup, Georgette and Steven H. Weinberger.1987. *Interlanguage phonology: The acquisi-*

tion of a second language sound system. Cambridge, MA: Newbury House.

Jakobson, Roman, Fant, C. Gunnar M., and Morris Halle.1967. *Preliminaries to speech analysis: The distinctive features and their correlates.* Cambridge, MA: MIT Press.

James, A.R.1984. Phonic transfer: The structural bases of interlingual assessments. *Proceedings of the Tenth International Congress of Phonetic Sciences.* Dordrecht: Foris.

Jamieson, Donald G. and David E. Morosan.1986. Training non-native speech contrasts in adults: Acquisition of the English /θ/-/ð/ contrast by francophones. *Perception and Psychophysics* 40:205-215.

Jun, Jongho.1995. Perceptual and articulatory factors in place assimilation: An optimality theoretic approach. Doctoral Dissertation, University of California, Los Angeles.

Keating, Patricia.1988. Underspecification in phonetics. *Phonology* 5:275-292.

Keel, W.1979. The acquisition of German as a second language. In *Studies in first and second language acquisition*, ed. F. Eckman and A. Hastings. Rowley, MA: Newbury House.

Kenstowicz, Michael.1994. *Phonology in generative grammar.* Cambridge, MA: Blackwell.

Kenstowicz, Michael and Charles Kisseberth.1979. *Generative phonology: Description and theory.* New York: Academic.

Kruatrachue, Foongfuang.1955. Thai and English: A comparative study of phonology for pedagogical applications. Doctoral Dissertation. Indiana University.

Kučera, H. and W.N. Francis.1967. *Computational analysis of present-day American English.* Providence: Brown University Press.

LaCharité, Darlene and Philippe Prévost.1999. Le rôle de la langue maternelle et de l'enseignement dans l'acquisition des segments de l'anglais langue seconde par des apprenants francophones. *Langues et linguistique* 25:81-109.

Lado, Robert.1957. *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor: University of Michigan Press.

Lombardi, Linda.2000. Second language data and constraints on manner: Explaining substitutions for the English interdentals. *Rutgers Optimality Archive.* September.

Maddieson, Ian.1984. *Patterns of sounds.* Cambridge, MA: Cambridge University Press.

Mester, R. Armin and Junko Itô.1989. Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65:258-293.

Michaels, D.1974. Sound replacements and phonological systems. *Linguistics* 176:69-81.

Miller, G.A. and P.A. Nicely.1955. An analysis of perceptual confusion among some English consonants. *Journal of the Acoustical Society of America* 27:338-346.

Mohanan, K.P.1993. Fields of attraction in phonology. In *The last phonological rule: Reflections on constraints and derivations*, ed. John Goldsmith, 61-116. Chicago: University of Chicago Press.

Morosan, David E. and Donald G. Jamieson.1989. Evaluation of a technique for training new speech contrasts: Generalization across voices, but not word-position or task. *Journal of Speech and Hearing Research* 32:501-511.

Nemser, William.1971. *An experimental study of phonological interference in the English of Hungarians.* Bloomington: Indiana University.

Peust, Carsten.1996. Sum: th-substitution. *The Linguist List* 7.1108.

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

Rice, Keren and Peter Avery.1995. Variability in a deterministic model of language acquisition: A theory of segment acquisition. In *Phonological acquisition and phonological theory*, ed. J. Archibald. Hillsdale: Lawrence Erlbaum.

Ritchie, William C.1983. On the explanation of phonic interference. In *Second language learning: Contrastive analysis, error analysis, and related aspects*, ed. Betty Wallace Robinett and Jacquelyn Schachter. Ann Arbor: University of Michigan Press.

Robinett, Betty Wallace and Jacquelyn Schachter, eds.1983. *Second language learning: Contrastive analysis, error analysis, and related aspects.* Ann Arbor: University of Michigan Press.

Sagey, Elizabeth.1986. The representation of features and relations in non-linear phonology. Doctoral Dissertation. MIT.

Sangster, Catherine M.2000. Lenition of alveolar stops in Liverpool English. *Oxford University Working Papers in Linguistics: Philology and Phonetics* 5:51-63.

Scovel, Thomas.1988. *A Time to Speak: A psycholinguistic inquiry into the critical period for human speech.* New York: Newbury House.

Shadle, C.H.1985. The acoustics of fricative consonants. Doctoral Dissertation, MIT.

Shadle, C.H., S.J. Mair, and J.N. Carter.1996. Acoustic characteristics of the front fricatives [f, v, θ, ð]. *Proceedings of the First ESCA Workshop on Speech Production Modeling.* ESCA: Autrans.

Sheldon, A. and W. Strange.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.

Someda, T.1966. Ei, futsugo to no hikaku ni okeru nihongo no chōon no ippanteki haikei ni tsuite. *Onsei no Kenkyū* 12:327-346.

Steriade, Donca.1987. Redundant values. In *Chicago Linguistic Society 23: Parasession on Autosegmental and Metrical Phonology,* ed. A. Bosch, B. Need, and E. Schiller. Chicago.

Steriade, Donca.1995. Underspecification and markedness. In *The handbook of phonological theory*, ed. John A. Goldsmith, 114-174. Cambridge, MA: Blackwell.

Stevens, Kenneth N.1998. *Acoustic phonetics.* Cambridge, MA: MIT Press.

Stevens, Kenneth N., S.J. Keyser, and H. Kawasaki.1986. Toward a phonetic and phonological theory of redundant features. In *Invariance and variability in speech processes*, ed. J. Perkell and D. Klatt. Hillsdale: Lawrence Erlbaum.

Strevens, Peter.1960. Spectra of fricative noise in human speech. *Language and Speech* 3:32-49.

Tabain, Marija.1998. Non-sibilant fricatives in English: Spectral information above 10 kHz. *Phonetica* 55:107-130.

Teasdale, Allison.1997. On the differential substitution of English [θ]: A phonetic approach. *Calgary Working Papers in Linguistics.* February.

Vance, Timothy J.1987. *An introduction to Japanese phonology.* Albany: SUNY Press.

Weinberger, Steven Howard.1994. Theoretical foundations of second language phonology. Doctoral Dissertation, University of Washington.

Weinreich, U.1966. *Languages in contact.* New York: Linguistic Circle of New York.

Wenk, B.J. 1979. Articulatory setting and de-fossilization. *Interlanguage Studies Bulletin* 4:202-220.

Werker, Janet F. and Richard Tees.1984. Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75:1866-1878.

Werker, Janet F. and John Logan.1985. Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37:35-44.

APPENDIX A

Table I. Transfer Variants Reported for Production of the L2 English Target *θ*.[56]

Substitute	Native Language	Source
[f]	Afrikaans, Chinese (Hong Kong), English (Black American, Cockney)	Peust (1996)
	FRENCH (QUÉBEC)	Brannen (1998)*
[f] beginners	**FRENCH (FRANCE)**	Wenk (1979)*
[t]	Chinese (Malaysia/Singapore), English (New York), Spanish, Yiddish, Dutch	Peust (1996)
		James (1986)*
	FRENCH (QUÉBEC)	Brannen (1998)*
	German (Austria)	Peust (1996)
	Hindi	Hancin-Bhatt (1994a)*
	Hungarian	Nemser (1971)*
	Norwegian, Polish	Peust (1996)
	Russian	Lombardi (2000), Ritchie (1968/1983), Weinberger (1990)
		Weinreich (1966)
	Serbo-Croatian	Kenstowicz & Kisseberth (1979)
	Tagalog, Thai	Lado (1957)
	Turkish	Hancin-Bhatt (1994a)*
[t]unschooled	**FRENCH (FRANCE)**	Berger (1951)*
[t] or [tʰ]	Thai	Kruatrachue (1955)*
[t] or [tθ]	**FRENCH (QUÉBEC)**	Gatbonton (1978)*
[t] or [f]	Czech	Peust (1996)
[s]	English (Liverpool)	Sangster (2000)*
	FRENCH (FRANCE)	Kenstowicz & Kisseberth (1979), Teasdale (1997), Weinreich (1966)
	German	Hancin-Bhatt (1994a)*
	Hebrew, Portuguese, Swedish	Peust (1996)
	JAPANESE	Hancin-Bhatt (1994a)*
	Sinhalese	Michaels (1973)*
[s] schooled	**FRENCH (FRANCE)**	Berger (1951)*
[s] intermediate	**FRENCH (FRANCE)**	Wenk (1979)*

56 Original research is indicated by an asterisk. The other reports are either anecdotal or secondary sources.

[θs]	**JAPANESE**	Dickerson (1976)*
[s] or [t] advanced[57]	**FRENCH (FRANCE)**	Wenk (1979)*
[s] or [t]	Turkish	Angus (1937)*
[s] or [f] high intermediate	**FRENCH (FRANCE)**	Brannen (1998)*

Table II. Transfer Variants Reported for Perception of the L2 English target *θ*.

Substitute	Native Language	Source
[f]	Hungarian	Nemser (1971)*
	English	Hancin-Bhatt (1994a)*
		Miller & Nicely (1955)*
		Tabain (1998)*
	German	Hancin-Bhatt (1994a)*
	Hindi	Hancin-Bhatt (1994a)*
[s]	**JAPANESE**	Brown (1997)*, Hancin-Bhatt (1994a)*
[t]	**FRENCH (QUÉBEC)**	LaCharité & Prévost (1999)*
	Turkish	Hancin-Bhatt (1994a)*

APPENDIX B

Algorithmic Calculation for ʃ

PREDICTIONS FOR PHONETIC PROCESSING

EF, QF, AND JA

Potential Substitute	*Matches*	*Mismatches*	*Distance*
/ʃ/	+2 (Cont) +2 (Cor) +1 (Lam)	-2 (Strid) 4 -1 (PostAlv) 2 -1 (Round) 2	8 away

57 Studies reporting [t] as a substitute in EF merge voiceless and voiced variants. As noted earlier, I have reason to believe that voicing affects whether target theta is substituted with a stop or fricative.

APPENDIX C

PERCEPTION OF TEST CONTRASTS: RELATIVE COMPARISON (NEWMAN-KEULS POSTHOC TESTS)

CONTRAST	ERROR RATE COMPARISON (> MEANS "SIGNIFICANTLY MORE ERRORS")	SIGNIFICANCE LEVELS
t̪ ~ θ STOP/CONT	EF = QF = JA = NE	N/A
t ~ θ STOP/CONT; APALV/LAMDENT	EF = QF = JA = NE	N/A
s̪ ~ θ STRID/MELL	JA > EF,QF,NE	$p < .001$
f ~ θ LAB/COR	EF,QF > JA,NE	EF $p < .05$; QF $p < .01$
f ~ ϕ LIP/DENT	NE > EF; QF= EF,JA,NE	$p < .05$
t ~ t̪ APALV/LAMDENT	EF = QF = JA = NE	N/A

APPENDIX D

Revised Algorithmic Calculations Based on Results

EF Algorithm Calculation for Results

Intake
θ +2 (Mell) +2 (Cont) +1 (Cor)

Potential Substitute	*Matches*	*Mismatches*	*Distance*
f	+2 (Mell) +2 (Cont)	-1 (Lab) 2	2 away
s̪	+2 (Cont) +2 (Cor)	-2 (Strid) 4	4 away
t̪	+2 (Mell) +2 (Cor)	-2 (Stop) 4	4 away

QF Algorithm Calculation for Results

Intake
θ +2 (Mell) +2 (Cont) +1 (Cor)

Potential Substitute	*Matches*	*Mismatches*	*Distance*
f	+2 (Mell) +1 (Cont)	-1 (Lab) 2	2 away
t̪	+2 (Mell) +2 (Cor)	-2 (Stop) 4	4 away
s	+2 (Cont) +2 (Cor)	-2 (Strid) 4	4 away

JA Algorithm Calculations for Results

Intake
θ +2 (Cont) +2 (Cor)

Potential Substitute	*Matches*	*Mismatches*	*Total*
s̪	+2 (Cont) +2 (Cor)		0 away
ɸ	+2 (Cont)	-2 (Lab) 4	4 away
t	+2 (Cor)	-2 (Stop) 4	4 away

Résumé

Ce papier examine la substitution différentielle de la fricative interdentale non-voisée en anglais langue seconde (L2), [θ]. Les langues maternelles (L1) examinées dans cette étude, le français européen, le français québécois, et le japonais, au dire de tout le monde, substituent en production [s], [t̪] et [s] respectivement (e.g. Wenk 1979, Gatbonton 1978, Hancin-Bhatt 1994a). Deux hypothèses principales sont explorées: 1. le transfert est basé sur la perception; 2. la substitution implique une évaluation de traits non-contrastifs en plus de traits contrastifs. Les résultats d'un test AXB montrent que les apprenants avancés sont incapables de percevoir certaines distinctions non-contrastives; cependant, contrairement aux auditeurs japonais, les auditeurs français perçoivent les traits stident et moelleux, traits qui ne sont pas contrastifs dans leur L1. les résultats indiquent qu'il y a un fondement perceptuel pour le substitut japonais; cependant, les différences entre le français québécois et européen ne semblent pas basées sur la perception. Un autre résultat est que la confusion de [f] et [θ] est plus grande pour les auditeurs français que pour les auditeurs japonais. La notion que la composition de l'inventaire phonétique de la L1 a une influence sur les traits auxquels les auditeurs portent leur attention au cours de la perception est proposée.