1

The Interrelation between Speech Perception and Phonological Acquisition from Infant to Adult*

Cynthia Brown

Introduction

The acquisition of a second language (L2) is clearly *somehow* different from that of a first language (L1): adult second language learners rarely (if ever) achieve the same native competence that children do learning their first language and, conversely, children never experience the degree of difficulty that L2 learners do.[1] This disparity between L2 and L1 acquisition is perhaps most apparent with respect to the acquisition of a second phonological system. Whereas children consistently achieve native competence across the full range of subtle and complex phonological properties of their language,

* Portions of this research were previously presented at the 1994 Boston University Conference on Language Acquisition, the 1994 Second Language Research Forum, and the 1996 Pacific Second Language Research Forum; this chapter is a revised version of chapter four of Brown (1997). Several people have contributed to the completion of this research: I am grateful to Joseph Tomei and John Matthews for lending their voices for experimental stimuli, to Masanobu Ueda for his assistance with the DAT equipment at Hokkaido University, to Dongdong Chen for recruiting Chinese subjects in Montreal, Canada, to Takako Kawasaki for recruiting Japanese subjects in Montreal, Canada and to Michie Namita for recruiting Korean and Chinese subjects in Sapporo, Japan. I would also like to thank Heather Goad and Lydia White for their guidance throughout the course of conducting this research. Finally, this research could not have been completed without the generous contribution of John Matthews to the design and presentation of these studies. The experimental research was supported by SSHRC research grant #410920047 to Lydia White; materials used in stimuli preparation for the picture test reported here were adapted from the Bilingual Aphasia Test (Paradis and Libben, 1987).

second language learners often have extraordinary difficulty mastering the pronunciation and intonation patterns of their L2. This lack of success is often taken as evidence that Universal Grammar (UG) does not operate in second language acquisition; but, perhaps there is another explanation. As White (1989) points out, other factors, in addition to UG, are necessary for successful first and, presumably, second language acquisition (for example, sufficient input and various learning mechanisms). An intriguing line of research suggests that the failure of some L2 learners to attain a native-like competence is attributable to these other factors, rather than to the non-operation of UG.

One such factor that distinguishes second language acquisition from first language acquisition is the fact that the second language learner comes to the task of acquisition already knowing a language. Most current theories of second language acquisition do, in fact, assume that the native language of the learner plays a role in acquisition. Although researchers generally agree that the learner's existing linguistic knowledge exerts some influence on the acquisition process, there is considerable debate as to precisely *what* role the native language plays (e.g., Bley-Vroman's *Fundamental Difference Hypothesis*, 1989, versus White's *Transfer Hypothesis*, 1988; see also papers in Schwartz and Eubank, 1996, on the L2 initial state). Moreover, existing research suggests that the influence of the native grammar is not absolute: some aspects of the L1 seem to prevent successful acquisition of particular L2 structures, whereas other properties of the L2 are acquired with little or no interference from the native grammar (Schwartz, in press). The challenge for second language theory now is to provide a principled explanation for the presence or absence of L1 influence, that is, what determines "partial influence".

Building on the insights of prior phonological research, this chapter develops a model of speech perception, couched within current phonological theory, that accounts for the influence of the native grammar in both infant and adult speech perception. More specifically, by utilizing the theory of Feature Geometry, the proposed model demonstrates how the monotonic acquisition of phonological structure by young children restricts their sensitivity to particular non-native contrasts and how the continued operation of this existing phonological structure in adult speech perception constrains which non-native contrasts adult learners will be sensitive to in the L2 input and, therefore, are capable of acquiring. By forging a link between infant speech perception and phonological acquisition, this research lays the foundation for a unified theoretical account of the interrelation between phonological acquisition and speech perception in children and adults. It also offers an explanation for why learners perceive L2 sounds in terms of their native phonemic categories; by isolating and characterizing those phonological properties of the L1 that impinge upon L2 acquisition, this research identifies why and how this equivalence classification takes place. Finally, by demonstrating how the L1 grammar can both facilitate and hinder acquisition, these findings provide an answer to one of the questions currently central to second language acquisition theory: what determines partial L1 influence? The model outlined in this chapter accounts for the differential success that speakers of different L1s have in acquiring a given non-native contrast; it also accounts for the differential success that speakers with the same L1 have in acquiring various non-native contrasts. Furthermore, the experimental studies reported here demonstrate how the existing phonological system may block accurate perception of the input, thereby preventing the acquisition of novel segmental representations; it also establishes the circumstances in which the native grammar actually facilitates perception of non-native contrasts, demonstrating that when there is sufficient intake to the acquisition device, novel segmental representations can be successfully acquired.

We will begin by reviewing some of the previous research that has been conducted on the L2 acquisition of segments in order to set the context for the present research program and see why a new analysis is needed. Next, the relevant aspects of phonological theory will be laid out and explained. This will be followed by an examination of the development of the native phonological and perceptual systems, which will then lead us to a theory of phonological interference. After the implications of this theory for second language acquisition are laid out, the results of three experimental studies which test this theory will be reported and discussed. The chapter concludes by considering some of the implications of these experimental data for the theory of phonological interference developed here as well as our theory of second language acquisition.

Historical Context and Theoretical Background

Previous L2 phonological research

Although previous L2 phonological research has addressed the question of *whether* the native language plays a role (e.g., Briere, 1966; Flege, 1981; Wode, 1978, 1992), it has not attempted to answer the question of *why* the native language influences L2 acquisition, nor has it formally articulated the mechanisms by which the native grammar influences this acquisition. Using the tools of current phonological theory, we are now in a position to develop a theory of L2 phonological interference which includes a principled explanation for the existence of L1 influence in some instances and its absence in others, as well as a description of the mechanism(s) by which this influence is exerted.

Conducting research in applied areas such as acquisition requires one to strike a delicate balance between (at least) two continually developing theories: our theory of acquisition and our theory of grammar. In the case of

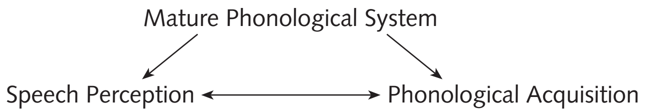


Figure 1.1 Interrelation between speech perception, phonological acquisition and the mature system

L2 phonological acquisition, we must integrate insights from the theory of second language acquisition and current phonological theory. Advances in one of these usually requires us to reinterpret implications of the other in light of these new developments and to recast our theoretical models and experimental hypotheses. Similarly, failure of our acquisition models to correctly account for some aspect of the data force us to consider whether it is the acquisition theory or the linguistic theory underlying our model which is inadequate and needs to be modified. This complex bi-directional relationship often leads to a non-linear flow of progress in acquisition research. We are now once again at a point of reinterpretation, forced by the limitations of current models to reformulate our theory of L2 phoneme acquisition in terms of shifts within both the theory of segmental representation and the theory of second language acquisition.

Successful acquisition of phonological representations requires accurate perception of phonemic contrasts in the input; it is therefore clear that a comprehensive model of L2 phoneme acquisition must integrate not only a theory of second language acquisition and a theory of phonological representation, but also a theory of speech perception. Thus, it is not enough to ask only how the existing phonological system affects acquisition of L2 segments; we must consider all of the relationships illustrated in figure 1.1.

The majority of research on L2 phonological acquisition has investigated the relationship between the mature phonological system and phonological acquisition. But, the interrelation of these factors raises three additional issues that an adequate theory of L2 phoneme acquisition must explain: (1) how does the mature phonological system affect speech perception? (2) how does speech perception affect phonological acquisition? and (conversely) (3) how does phonological acquisition affect speech perception? By isolating the specific research questions addressed by previous L2 phoneme research and highlighting the particular theory of acquisition and/or theory of phonological representation assumed by each approach, we will see why a new analysis is needed.

The earliest systematic approach to the acquisition of L2 segments was undertaken within the contrastive analysis framework, the prevailing theory of second language acquisition of the time (Lado, 1957; Lehn and Slager, 1959; Stockwell and Bowen, 1965). The primary question addressed by this research was how the L1 influenced the acquisition of L2 segments, where acquisition was measured by the learner's ability to produce those segments.

This approach, however, was unable to account for aspects of the observed acquisition data. In particular, it incorrectly predicted that an L2 learner would have the same degree of difficulty with any and all of the L2 sounds not present in the L1 inventory, when, in fact, learners' performance on different L2 segments in experimental conditions ranges from native-like levels of accuracy to chance performance (see Munro, Flege and MacKay, 1996, for a detailed discussion of this point). This approach also failed to explain why learners with different L1s would substitute different L1 sounds for a given L2 sound (e.g., Japanese speakers substitute [s] for [θ] but Russian speakers substitute [t], despite the fact that these L1s contain both /s/ and /t/ – Hancin-Bhatt, 1994a). These shortcomings, and in fact the most significant limitation of this approach, were due not to its comparison of L1 and L2 inventories, but rather to the level of phonological representation at which the languages where compared: these researchers took the phoneme to be the relevant unit of analysis.

Influenced by developments in generative phonology (and publication of Chomsky and Halle's *The Sound Pattern of English*, 1968), the next wave of research on L2 phoneme acquisition focused their analyses on the differences and similarities in distinctive features between the L1 and L2 (Michaels, 1973, 1974; Ritchie, 1968). According to this line of research, difficulty with particular L2 sounds could be explained in terms of featural differences between the L1 and L2, combined with the learner's perceptual biases. This line of research constituted an advance over the previous contrastive analysis approach in that its focus on the distinctive feature as the relevant unit for comparing the L1 and L2 provided language-internal evidence for differential substitutions. Moreover, it represented the first attempt to address the issue of how the mature phonological system might affect speech perception and how that, in turn, might affect phonological acquisition; however, it did not attempt to formally articulate this L1–L2 perceptual mapping, nor did these researchers provide any experimental evidence for their claims about how the native grammar influenced perception.

In the 1970s and 1980s, several perceptual studies conducted with native speakers and language learners provided the necessary experimental evidence, demonstrating that phonemes are indeed generally perceived in terms of the speaker's native categories (Abramson and Lisker, 1970; Miyawaki, Strange, Verbrugge, Liberman, Jenkins and Fujimura, 1975; Werker and Tees, 1984b; Williams, 1977). Since that time, three models have been proposed to explain how L2 sounds are mapped onto L1 sounds. The first model we will consider restricts itself to the relationship between the mature phonological system and speech perception; it does not address how phonological acquisition relates to these two factors. Best (1993, 1994) has developed the Perceptual Assimilation Model (PAM) to explain the role that a speaker's L1 phonological system plays in the perception of non-native sounds. According to this model, non-native sounds are assimilated to a listener's native categories on the basis of their respective articulatory similarities (more specifically, the spatial proximity of constriction location and active articulators); the degree to which a non-native contrast can be assimilated to native categories determines how well (if at all) a listener will be able to perceive that non-native contrast. While Best's proposal is based on, and supported by, experimental perceptual data, it lacks precise objective criteria for determining how non-native contrasts will be assimilated into native categories. Thus, although the PAM describes the role that a speaker's L1 phonological system plays in the perception of non-native sounds, it does not provide an explanation for why or precisely how this mapping occurs. Moreover, Best's model, concerned primarily with the role of the L1 in the perception of *foreign* sounds, is essentially static; that is, it does not include any means by which the existing L1 phonemic system might be altered by exposure to non-native segmental contrasts and, therefore, does not directly address the acquisition of novel segments.

One model that does address the issue of L2 segment acquisition is the Speech Learning Model (SLM), developed by Flege (1991, 1992, 1995). The SLM attempts to explain how speech perception affects phonological acquisition by distinguishing two kinds of sounds: "new" and "similar". New sounds are those that are not identified with any L1 sound, while similar sounds are those perceived to be the same as certain L1 sounds. Flege suggests that although the phonetic systems used in production and perception remain adaptive over the life span and reorganize in response to sounds in the L2 input, a process of "equivalence classification" hinders or prevents the establishment of new phonetic categories for similar sounds. However, this model does not include a theory-based proposal as to how L2 sounds are equated with L1 sounds; although equivalence classification is stated in probabilistic terms, allowing for the eventual development of L2 categories, there is no concrete proposal for how or when this takes place. Thus, these two models, the PAM and the SLM, attempt to elucidate the interrelation between the mature phonological system, speech perception and L2 phonological acquisition. However, despite their claim that there is an underlying mechanism that maps L2 sounds onto L1 categories, they fail to articulate the nature of that mechanism or adequately formalize the perceptual mapping process.

The most extensive model of speech perception–phonological acquisition interaction to be proposed thus far is Hancin-Bhatt's Feature Competition Model (FCM) (1994a, b). Expanding on the earlier work by Ritchie (1968) and Michaels (1973) described above, this model assumes that the features utilized in a grammar differ with respect to their "prominence": features (and feature patterns) used more frequently in the language's phonology will be more prominent than less frequently used features. Those features that are more prominent in the L1 system will tend to have a greater influence on learners' perception of new L2 sounds; that is, the feature prominences in the L1 will guide how L2 sounds are mapped onto existing L1 categories. Thus, like the PAM and SLM, the FCM assumes that L2 sounds are assimilated to L1 categories, yet this model goes one step further by providing an algorithm for determining feature prominence and, thereby, generating testable predictions for differential perception and substitution of interdentals across learners with different L1s. Furthermore, it is the first comprehensive model to investigate both the relationship between the mature phonological system and speech perception and the relationship between speech perception and the acquisition of L2 phonemic representations. Thus, this model addresses two of the three relationships indicated in (1) above; it also provides a more formal articulation of the L1–L2 perceptual mapping. However, to date its scope has been limited to the study of interdental substitutions; it is not clear whether this model can account for substitutions of other types of segments cross-linguistically or whether it can account for differential difficulty that speakers of a single L1 encounter in the acquisition of various L2 segments. Most importantly, though, the FCM does not address the reciprocal relation between perception and acquisition, namely how (L1) phonological acquisition affects speech perception.

So, while we are moving closer and closer to a formalization of the influence that the mature phonological system has on speech perception (and the consequence of this for L2 acquisition), we still do not understand how the interrelation between speech perception and phonological knowledge originates; therefore, we fail to capture the essential nature of the phonological transfer mechanism. Investigating the development of speech perception and phonological acquisition in young children will enable us to explain why the mature phonological system exerts such a profound influence in adult speech perception; moreover, utilizing the tools of current phonological theory will allow us to articulate the L1–L2 perceptual mapping mechanism more precisely, as well as allow us to explain how the new phonemic categories develop in the L2 learner (i.e., how the relationships in figure 1.1 change over time).

Phonological theory and the representation of phonemes

Whereas previous research on L1 phonological interference primarily considered the phonemic categories of a language, phonological theory within the generative framework assumes that phonemes themselves have an internal structure. Thus, one way current phonological theory provides greater insight into the phenomenon of L1 influence is the distinction made between phonological representations and the components that comprise those representations. L2 phonological researchers now have an additional tool of analysis: the internal sub-components of phonemes constitute a further level of linguistic knowledge which may impinge upon L2 acquisition. However, these components (i.e., distinctive features) are not simply unordered bundles, as was assumed in the SPE framework (and theories of L2 phoneme acquisition couched within this framework). Instead, the distinctive features are

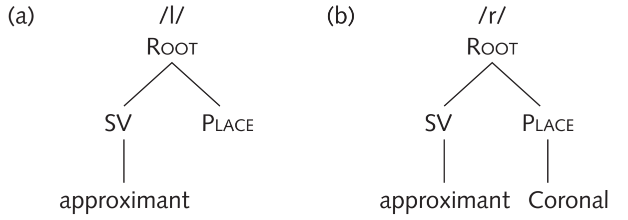


Figure 1.2 Segmental representations for phonemes that are contrastive in a language (English /l/ vs. /r/)

themselves structured – an advancement which has implications for our theory of speech perception and phonological acquisition. Since an understanding of the internal structure of phonemes is necessary for the subsequent discussion of phonological interference, we will begin with a brief review of the relevant aspects of the theory of segmental representation assumed here.[2]

According to the theory of Feature Geometry, phonemes consist of distinctive features which are organized into a systematic hierarchy of constituents (Clements, 1985; Sagey, 1986).[3] Each phoneme has a unique structural representation (i.e., feature geometry) that distinguishes it from other segments in an inventory. Much of the work in phonological theory, as in linguistic theory in general, has been guided by the presumption that redundant information (defined as that information that can be predicted or easily supplied by derivation, e.g., syllable structure) is absent from underlying representations (Chomsky and Halle, 1968). Within Feature Geometry this principle is also extended to segmental representations.

One such theory of underspecification is Minimally Contrastive Underspecification. According to this position, a segmental representation contains only the information needed to contrast it from all other segments in the system; any further specification will be provided by a system of phonetic implementation (Avery and Rice, 1989; cf. Archangeli, 1984, 1988, on Radical Underspecification). Thus, the precise representation of a segment will depend entirely upon which segments it contrasts with in the particular inventory. For example, in English, where the lateral approximant /l/ and central approximant /r/ are contrastive, /l/ may be represented as (a) in figure 1.2 while /r/ may be represented as (b), omitting irrelevant structure (Piggott, 1993; Brown, 1993b, 1995).[4]

The fact that these segments have different representations reflects the fact that they are contrastive phonemes in the language; the presence of [coronal] in the representation of only one of them is sufficient to distinguish these segments in the grammar.

Conversely, when these two segments are not contrastive in a language, they will not have distinct representations. For example, in Japanese, both [l] and [ɾ] are freely varying allophones of a single phoneme, so there will be

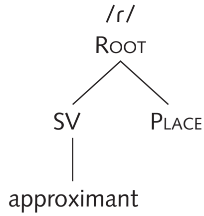
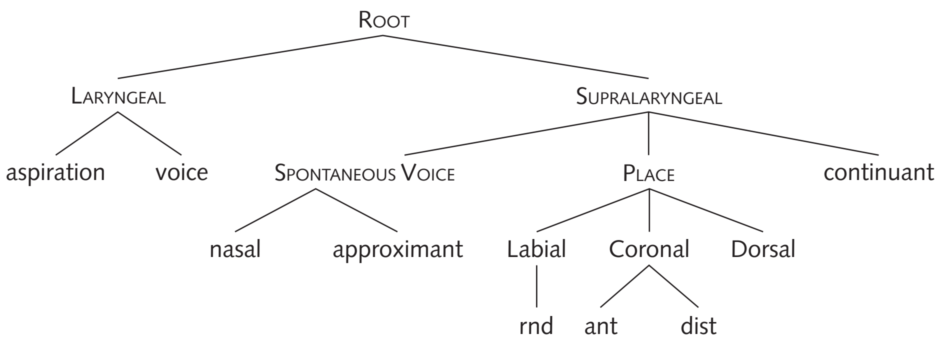


Figure 1.3 Segmental representation of Japanese /ɾ/



Note: rnd = round, ant = anterior, dist = distributed

Figure 1.4 A model of Feature Geometry

only one underlying representation for these two surface segments. This is given in figure 1.3.

Despite the fact that this Japanese segment is realized as a coronal, in accordance with our theory of underspecification, it is not specified for the feature [coronal] because it does not contrast with any other coronal approximants.[5] The phonetic realization of this segment as an [l] or an [ɾ] (which, unlike the English /r/, is a flap) varies freely (International Phonetic Association, 1979; Vance, 1987). Thus, in this way a speaker's knowledge of which sounds in his or her language are contrastive is represented by distinctive segmental representations.

The full set of features manipulated in the world's languages and their dependency relations can be represented in terms of a single, universal Feature Geometry. This universal geometry is given in figure 1.4 for illustration.[6]

This Feature Geometry is contained in the phonological component of Universal Grammar, the innate language faculty ascribed to the child by generative theorists. Like a syntactic principle or parameter, this geometry constrains the acquisition process and provides the learner with information about what phonemic oppositions are possible in natural languages.

Thus, while no one language manipulates all components of this universal Feature Geometry, every phoneme in the world's languages can be represented in terms of the features and structural relations present in this geometry. Two structural relations between features – constituency and dependency – are particularly important and must be captured in *any* theory of segments. Dependency is a structural relationship between features, such that the presence of a dependent node (or feature) in the representation of a segment entails the presence of its superordinate node in that representation. For example, in the model in figure 1.4, the feature [anterior] is a dependent of the Coronal node and all representations specified for [anterior] will also contain the Coronal node. Constituency refers to the structural relation that holds among features that are dominated by a common node in the geometry. Since phonological processes manipulate constituents of segmental structure, the features of a constituent all pattern together in the phonological operations of a grammar. Again, considering figure 1.4, a phonological process that manipulates the Coronal node will also affect the [anterior] and [distributed] features.

Languages will differ with respect to their phoneme inventories and, hence, with respect to the set of phonological features they manipulate. However, the organization of those features, as given by the universal Feature Geometry, will be the same in every language. The learner's task is to determine which of the phonological features contained in this universal geometry are used to contrast phonemes in the language he or she is learning and to construct the appropriate representations. In the remainder of this chapter, we will consider whether L2 learners can acquire non-native segmental representations as well as how, and to what extent, the native phonological system influences this process.

A Theory of Phonological Interference

In developing a theory of L1 influence, one of the issues we must address is *why* the L1 grammar exerts this influence. With respect to phonological interference, the relevant question is why foreign sounds are perceived in terms of the learner's native sound categories. In order to fully understand why the phonological system affects perception in this way, we must examine the genetic development of these systems, as well as any interdependence between them. Studying the development of the L1 system will offer us insight into its operation in mature speakers; in addition, an understanding of how phonological knowledge is acquired in L1 acquisition will enable us to determine what conditions are necessary for successful L2 acquisition.

L1 phoneme acquisition

Since segments are distinguished in a grammar by their internal feature geometries, acquisition of a phonemic contrast involves the acquisition of the relevant structure (i.e., distinctive features) that differentiates those two phonemes (Rice and Avery, 1995, based on Jakobson, 1941). The child is able to contrast the two phonemes in his or her grammar once that phonological structure has been acquired and a representation has been constructed. Brown and Matthews (1993, 1997) demonstrate experimentally that children's ability to differentiate phonemes phonologically develops gradually over time and in a systematic order that is consistent across children (see also Barton, 1980; Edwards, 1974; Garnica, 1973; Shvachkin, 1948, for related studies). Based on these results, they argue that UG provides the child's emerging grammar with a minimal amount of segmental structure (in fact, only those portions of the feature geometry that are universal) which is subsequently expanded over the course of acquisition until the adult feature geometry for the particular language is attained.

The systematic order of acquisition results from the hypothesized acquisition process and the nature of Feature Geometry itself; the child will only elaborate the feature geometry in his or her grammar in ways that are consistent with the hierarchical organization of features in UG. Specifically, the particular dependency and constituency relations that are encoded in the feature geometry in UG will be respected in the geometry posited by the child. For example, the presence of a dependent feature in a representation entails the presence of that feature's superordinate node. By extension, superordinate structure must be posited in the child's feature geometry before dependent structure can be elaborated. As a result, children will phonologically distinguish those segments that require less structure to differentiate before distinguishing those segments that require highly articulated structure. Thus, phonological structure is added to the child's grammar in a uniform, step-by-step fashion.

This step-by-step elaboration of the child's feature geometry in his or her grammar is driven by the child's detection of contrastive use of segments in the input (Jakobson, 1941; Rice and Avery, 1995). Once a child notices that two segments are used contrastively (i.e., are distinct phonemes), the phonological structure that differentiates the two segments is added to his or her grammar. If the child never perceives contrastive use of two segments (because, for example, they are allophones of a single phoneme in that language), the structure that differentiates them will never be posited. Therefore, the mere *presence* of two contrastive segments in the input (while necessary) is not sufficient to trigger acquisition; the learner must *detect* the contrast in the input.

Infant speech perception

In order for a learner to detect that two sounds are used contrastively, the learner must be able to discriminate the two sounds perceptually. Hence, proper development of the phonological system is dependent on properties of the speech perception mechanism. Given the fact that a child may be born into any language environment, it is imperative that he or she be equipped with adequate cognitive machinery to perceive (or, at the very least, be predisposed to perceive) the whole range of possible phonetic contrasts (cf. Burnham, 1986). Researchers have, in fact, demonstrated that infants as young as one month old are able to acoustically discriminate not only the sounds of the ambient language but many non-native contrasts as well (Eilers, Gavin and Oller, 1982; Eimas, Siqueland, Jusczyk and Vigorito, 1971; Streeter, 1976; Trehub, 1976; see Mehler, 1985, for a review).

Since the detection of contrasts in the input is crucial for the acquisition of phonemic representations, we need to consider how this capacity changes (if at all) as the child develops. In particular, whether or not an L2 learner has the capacity to perceive a non-native contrast will be a factor in determining if he or she will be able to construct the phonological representations necessary to distinguish the two segments phonologically. It is now well established that the ability to acoustically discriminate non-native contrasts decreases rapidly in infancy with exposure to a specific language, until the child is able to discriminate only those contrasts present in the language being acquired (see Werker and Polka, 1993, for a review of studies that establish this observation). In a series of studies, Janet Werker and her colleagues demonstrate that the decline in the ability to acoustically discriminate non-native contrasts occurs within the first year of life (Werker, Gilbert, Humphrey and Tees, 1981; Werker and LaLonde, 1988; Werker and Tees, 1984a, b; and also Best and McRoberts, 1989; Best, 1994).[7] What is particularly fascinating is that this decline in perceptual capacity does not appear to be temporally uniform for all non-native contrasts. Experimental results indicate that perceptual sensitivity to certain non-native contrasts is lost before sensitivity to others, suggesting that loss of perceptual sensitivity to non-native contrasts is gradual and proceeds in a systematic order. An explanation for this decline in speech perception abilities, in particular one that integrates the role of linguistic experience, is still needed.[8] Both Werker and Best have tentatively suggested that the decline in the ability to discriminate some non-native contrasts may reflect the first stage of phonological development in the child, though neither is specific as to which aspect of the developing phonology might be responsible for this change. In the following section, I examine some findings from infant speech perception research and suggest a causal link between the development of a learner's feature geometry and the subsequent decline in perceptual capabilities. Establishing such a link will have important consequences for the acquisition of a second phonological system.

The role of the L1 phonological system in speech perception

If we consider the findings from the infant speech perception research together with the research on phonological acquisition, an interesting parallelism emerges. We see that infants' perceptual capacities gradually "*degrade*" from all potential contrasts to only native contrasts (with some interesting exceptions), while their ability to discriminate segments phonologically gradually *improves* from no contrasts to only native contrasts. An exhaustive comparison of the stages of phonological and perceptual development is not feasible at this point, due to the limited number of non-native contrasts that have been investigated thus far. However, an examination of the data that we do have available suggests an intriguing possibility. According to Brown and Matthews, children first phonologically differentiate labials from velars, followed by labials from coronals. They do not distinguish segments that require a coronal node, such as /l/ and /r/, until relatively late. So, the node that distinguishes velar segments from other places of articulation is posited by the child before the node to distinguish among coronal segments is posited. Measuring auditory perception, Werker found the mirror order, with the perception of contrasts involving velars declining before contrasts involving coronals. Based on this convergence of the learner's perceptual and phonological capacities on the set of native sounds, Brown (1993a, 1998) proposes that there is a causal link between the learner's phonological development and the concomitant decline in his or her ability to acoustically discriminate non-native sounds.[9]

According to Brown's proposal, the acquisition of phonological structure (more specifically, the elaboration of feature geometry) in the child's grammar imposes upon his or her perceptual system the specific boundaries within which phonemic categories are perceived. In other words, the degradation of the perceptual capacities and the increase in the ability to distinguish sounds phonologically are the result of the same internal mechanism, namely the construction of phonological representations. This layer of phonological structure subsequently mediates between the acoustic signal and the linguistic processing system.

If we are correct in postulating that the acquisition of a phonological system determines the course of speech perception development, then it is reasonable to assume that the phonological system continues to constrain speech perception in adults. Mediating between the acoustic signal and the linguistic system, the phonological structure of the native grammar can be viewed as a filter that funnels acoustically distinct stimuli into a single phonemic category. This results in the well-documented phenomenon of categorical speech perception, whereby speakers of a language are able to easily distinguish members of different native phonemic categories and relatively unable to distinguish members of the same native phonemic category (Abramson and Lisker, 1970; Mattingly, Liberman, Syrdal and Halwes, 1971; Miller and Eimas, 1977; Pisoni, 1973; Repp, 1984). In other words, the mature speaker perceives the sounds of his or her native language, filtered through the existing phonological system, as distinct segments.

The hypothesis that the acoustic signal and phonological categories are mediated by a level of feature organization is quite intuitive given the fact

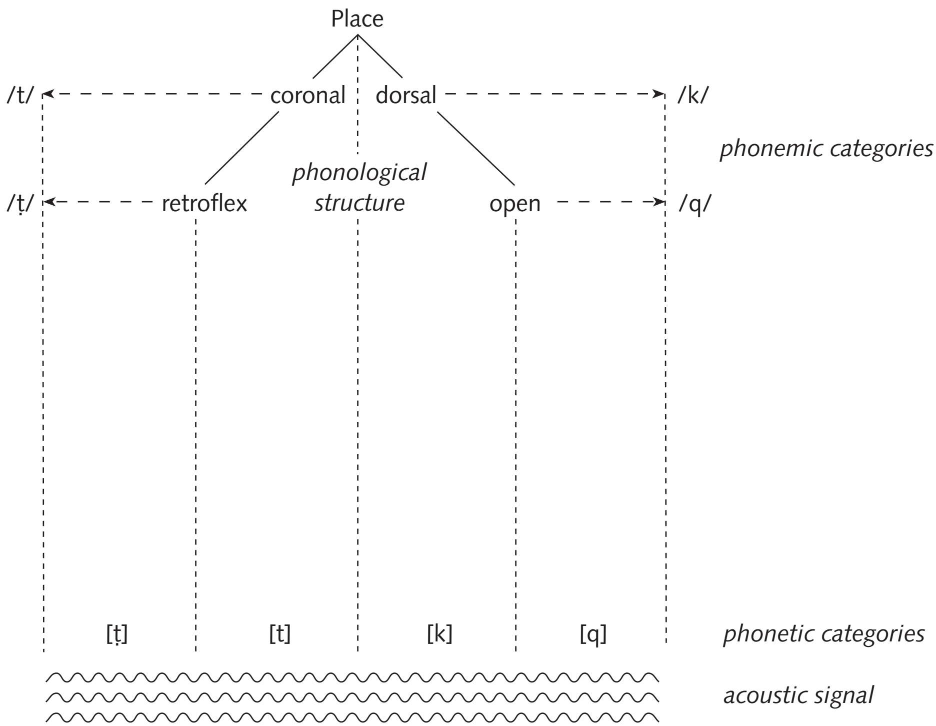


Figure 1.5 Mediation of speech perception by phonological structure

that the acoustic signal cannot be characterized in terms of abstract categories, such as a phoneme, but does correlate to properties of the gesture (e.g., place of articulation features correspond to spectral peaks in release bursts and to formant frequencies). The schematized diagram in figure 1.5 (taken from Brown, 1993a, 1998) illustrates the role of the intervening layer of phonological structure and how this level, in effect, funnels the acoustic signal into the phonemic categories of the speaker's language. The feature geometry depicted is from a hypothetical language in which /t/, /ṭ/, /q/ and /k/ are distinct phonemes; we will only consider the Coronal and Dorsal nodes (and their dependents) of the geometry.

Starting from the bottom of the diagram, the acoustic signal is first broken down into phonetic categories. At this level, the acoustic signals for an alveolar [t] and a retroflex [ṭ] remain distinct. These stimuli then pass to the second level which consists of a speaker's feature geometry. This phonological structure serves to further categorize the phonetic stimuli into phonemic categories which are then fed into the language processor. Because this language exploits a dependent feature of the Coronal node, the phonetic signals for [t] and [ṭ] are channeled into the distinct phonemic categories /t/ and /ṭ/. The acoustic signals for [q] and [k] are processed in the same way. This model of

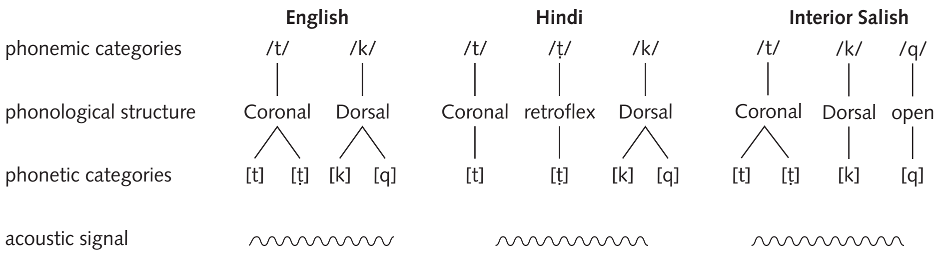


Figure 1.6 Cross-language speech perception

speech perception is supported by research by Werker and Logan (1985), who found evidence for three distinct levels of processing: depending on the length of the interval placed between the stimuli (and hence on the memory load required to perform the task), subjects exhibit perception at either the auditory, phonetic, or phonemic level. In particular, these researchers showed that, under certain conditions, English speakers are able to acoustically discriminate the Hindi /t–ṭ/ contrast more accurately than predicted by chance.

According to the model in figure 1.5, the acoustic signal will first be divided into distinct phonetic categories, which are only subsequently categorized into native phonemic categories. Thus, regardless of the phonological system of a speaker, non-native contrasts are distinct at some level and may be discriminated under certain controlled conditions, as Werker and Logan have demonstrated. In other words, the "loss" of sensitivity observed in young infants is not really a loss at all, but rather is the result of perceptual reorganization – reorganization, I would like to suggest, that reflects the hierarchical organization of the feature geometry in the speaker's grammar.

Figure 1.6 illustrates, in slightly more abstract terms, how the speech perception of English, Hindi and Interior Salish (Nthlakampx) speakers differs from one another.

In English, the Coronal node serves to distinguish coronals from non-coronals (e.g., /t/ vs. /p/), but no distinction is made within the coronal place of articulation (e.g., /t/ vs. /ṭ/). Thus, the English feature geometry does not contain the feature [retroflex]. As a result, (all) coronal sounds, regardless of their distinct acoustic signals, are perceived as a single phonemic category. Likewise, English makes no phonemic distinction between velar and uvular sounds; therefore, the Dorsal node has no dependents and velar and uvular sounds will be perceived as the English phoneme /k/. The feature geometry of Hindi also lacks the dorsal dependent [open], so that perception of [k] and [q] as /k/ is the same as for English speakers. Unlike English, however, the Hindi feature geometry contains both the Coronal node and its dependent [retroflex]. Thus, all coronal sounds will not be funneled into one phonemic category; the two features in the geometry ensure that /t/ and /ṭ/ will be perceived as distinct phonemes. In Interior Salish, the situation is just the reverse: /k/ and /q/ are perceived as distinct phonemes (due to the presence of the feature [open] as a dependent of the Dorsal node) while the acoustic signals for /t/ and /ṭ/ are perceived as the single phoneme /t/ (due to the absence of Coronal node dependents).

That the native system operates in this way is not accidental: perceiving speech in terms of phonemic categories undoubtedly aids processing and facilitates comprehension of the linguistic signal. Native speakers are continually faced with variable realizations of segments, due to coarticulation, sloppy articulation or inter-speaker variability. By filtering out this irrelevant "noise" in the acoustic signal, the memory load put on the auditory system is greatly reduced and processing can proceed more quickly. Those variations in the acoustic signal that do not contribute to differences in meaning are simply not perceived by the listener. Yet, although categorical perception aids processing of one's native language, it can be a barrier to correctly perceiving and processing a foreign language: variation in the acoustic signal which is filtered out by the native phonological system (i.e., is treated as *intra*-category variation) may, in fact, contribute to differences in meaning in the foreign language (i.e., actually constitute *inter*-category variation). Thus, the influence of the mature phonological system on the perception of foreign sounds is an artifact of how speech perception functions in general. To summarize thus far, I have suggested that a learner's developing feature geometry causes the gradual decline in the ability to acoustically discriminate non-native contrasts and then continues to mediate between the acoustic signal and the linguistic processing system. The next section outlines the predictions this proposal makes for L2 acquisition of phonology.

Implications for L2 phonological acquisition

Establishing a link between a learner's phonological development and his or her speech perception has important implications for the acquisition of non-native contrasts by second language learners. In particular, this proposal suggests that the learner's native grammar constrains which non-native contrasts he or she will accurately perceive and, therefore, limits which non-native contrasts that learner will successfully acquire.

A speaker's phonological knowledge consists of phonemic representations as well as the features that comprise those representations. The position that the features exist in grammar (somewhere) independent of the segments they define is an assumption at this point in the discussion; however, we will see experimental support for this claim in experiment 3 below. A priori, either of these levels of knowledge (i.e., featural or segmental) could potentially impinge upon the L2 acquisition process. According to the theory of phonological interference outlined here, however, it is the *features* contained in the learner's native grammar, not the phonological representations themselves, which constrain perception. The prediction of this position is that if a speaker's grammar lacks the feature that differentiates a given phonological contrast, then he or she will be unable to accurately perceive that contrast; conversely, the presence of the contrasting feature in the native grammar will facilitate perception of that non-native contrast, regardless of whether the particular segment is part of the inventory. That is, despite a lack of acoustic, phonetic or phonemic experience with a *particular* non-native contrast, a speaker's experience perceiving native phonemic contrasts along an acoustic dimension defined by a given underlying feature (for example, voicing) permits him or her to accurately discriminate *any* non-native contrast that differs along that same dimension. (This *is* a strong claim, but one I would like to maintain until empirical data force me to a weaker position.) Thus, perception of certain non-native contrasts is possible by virtue of the fact that the phonological feature that underlies that particular acoustic dimension exists independently in the learner's native grammar.[10]

This position does not, however, entail that the phonological categories themselves play no role whatsoever in perception; while it is claimed that it is the features that determine the perceptual sensitivities (and that guide the mapping of the acoustic signal onto perceptual categories), it is still the existing phoneme categories which the incoming acoustic stimuli are sorted into, at least initially. As we will see from the experimental studies reported below, the effects of the L1 phonological categories will be most apparent in the initial stages of acquisition, before new L2 categories have been established. Nevertheless, it is the features of the L1 which ultimately enable or prevent the construction of these new L2 categories.

If the native phonological system affects perception of non-native contrasts in this way, either preventing or facilitating accurate perception, what are the consequences for the acquisition of these contrasts by L2 learners? Recall from our discussion of L1 phoneme acquisition, that acquisition of the relevant phonological structure is triggered by the learner's detection that the two sounds are used contrastively in the language (i.e., that they correspond to separate phonemes). For example, if the learner is to acquire the phonological structure required to differentiate /l/ and /r/ in his or her grammar, then he or she must notice that minimal pairs, such as *right* and *light*, are distinct words. In short, accurate perception of a phonemic contrast is necessary for successful acquisition of that contrast. It follows, then, that L2 learners will acquire only those non-native phonemic contrasts that they perceive as distinct sounds. If an L2 learner detects that two segments are used contrastively in the foreign language, then acquisition of the novel representations will be triggered.[11] On the other hand, if a contrast between two foreign sounds is not perceived (i.e., both sounds are perceived as belonging to the same phonological category), then acquisition will not be triggered and the L2 learner will fail to distinguish those segments in his or her interlanguage grammar. Put slightly differently, if the L2 input continues to be (inaccurately) mapped to L1 representations, there will be no impetus for acquisition.

Experimental Evidence

The following three experimental studies investigate how the grammars of Japanese speakers, Korean speakers and Mandarin Chinese speakers affect their acquisition of English contrasts and whether, given the necessary conditions, novel segmental representations can be constructed. These studies were designed to explore three related issues: the acquisition of a range of phonemic contrasts by a single group of speakers (experiment 1), the acquisition of a particular contrast across different groups of speakers (experiment 2) and whether the nature of L1 phonological influence changes over the course of L2 development (experiment 3). As the studies examine the acquisition of different subsets of English contrasts, the representations of all of the segments under investigation, as well as the phonological properties of the three L1s, will be discussed together, prior to the description of the individual experiments.

Contrasts investigated

The English /l–r/, /b–v/, /p–f/, /f–v/ and /s–θ/ contrasts were chosen to test the proposed model of phonological interference because these pairs are not contrastive in Japanese, Korean or Chinese; furthermore, since these contrasts are distinguished by different phonological features, each contrast could potentially cause a differing degree of difficulty for these groups of learners (both with respect to the various contrasts and with respect to the various L1 groups). The internal structure of each pair is given in figure 1.7; note that these representations are for the segments as they occur in English, a language in which they are contrastive. The phonological feature that distinguishes each contrast is given to the right (the superordinate Supralaryngeal and Laryngeal components are not relevant for this discussion and have been omitted for ease of exposition).

The representations in figure 1.7 are what the learner must acquire in order to distinguish these phonemes in his or her interlanguage grammar. The important thing to note is that each pair of phonemes is minimally differentiated by the presence of a single phonological feature. Without this contrasting feature in the representation of one of the sounds, the two segments will not be distinguished in the learner's interlanguage grammar. The /l–r/ contrast is distinguished by the feature [coronal], the /b–v/ and /p–f/ contrasts by the feature [continuant], while the /f–v/ and /s–θ/ contrasts are distinguished by the features [voice] and [distributed], respectively. Whether an L2 learner will successfully acquire each of these contrasts depends entirely on the presence or absence of the contrasting feature in his or her native grammar.

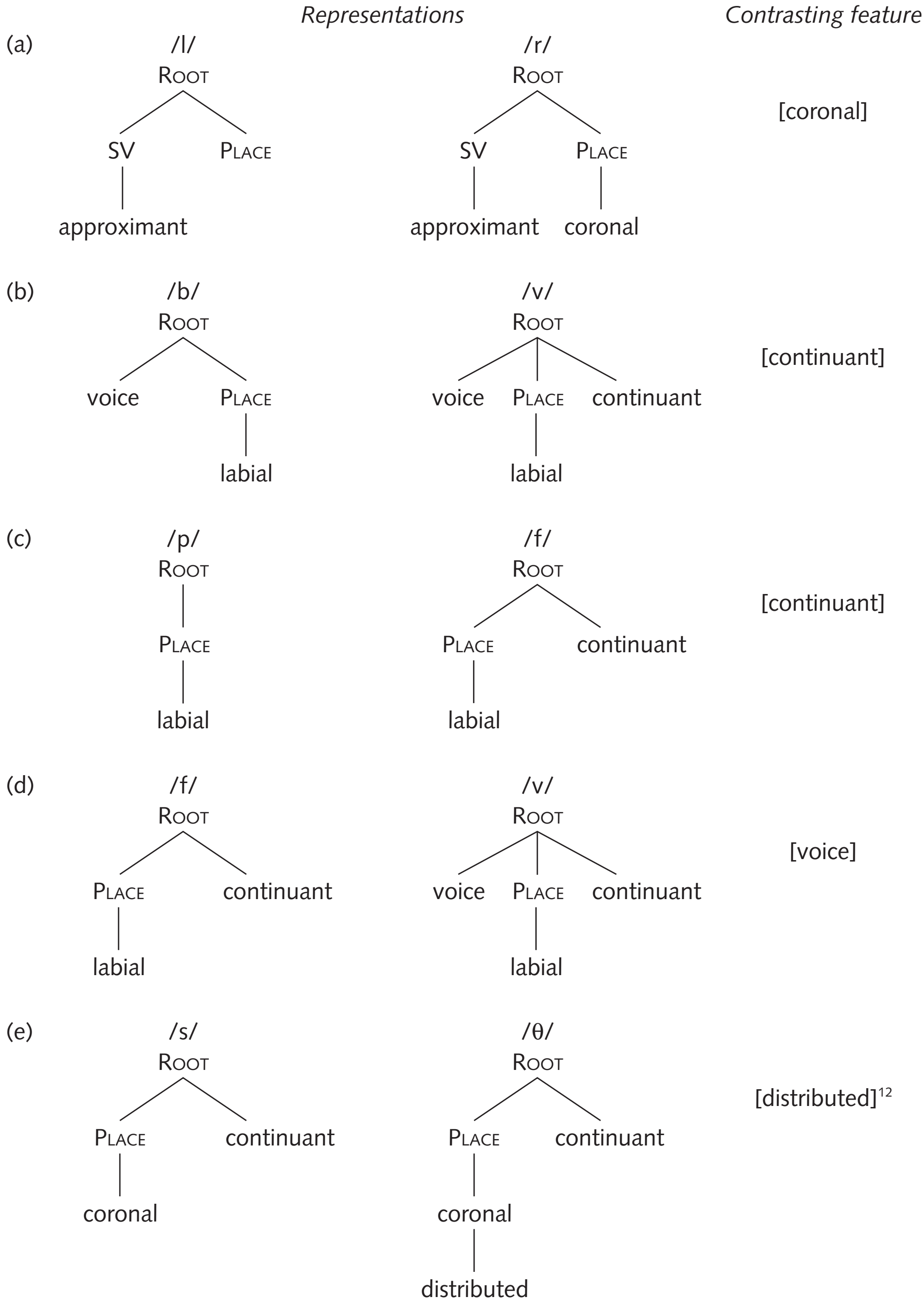


Figure 1.7 Representation of contrasts under investigation

Japanese Inventory[13]

p	t		k	
b	d		g	
		ʧ		
		ʤ		
	s			h
	z			
m	n			
	ɾ			
w		j		

Mandarin Chinese Inventory[14]

p	t		k	
p^h	t^h		k^h	
	ts	ṭṣ		
f	s	ṣ		h
		ẓ		
m	n			
	l			
w	j			

Korean Inventory

p	t		k	
p^h	t^h		k^h	
p′	t′		k′	
		ʧ		
		ʧ′		
	s			h
	s͈			
m	n			
	l			
w	j			

Figure 1.8 Phoneme inventories of L1 groups under investigation

Phonological properties of Japanese, Korean, and Mandarin Chinese

Let us now examine the consonant phoneme inventories of Japanese, Korean and Chinese, given in figure 1.8, in order to ascertain whether the features that distinguish the English contrasts are contained in the mental grammar of these speakers (from Maddieson, 1984, and Vance, 1987). From this, we see that each of the five non-native contrasts we are interested in has a slightly different status *vis-à-vis* the inventories; the status of these contrasts is summarized in table 1.1.

It should be noted that "corresponds to a native segment" means here that the surface realization of a given segment (in Japanese, Chinese or Korean) could reasonably be assumed to be a surface realization corresponding to the underlying representation of an English phoneme. For example, while Japanese does not contain a labiodental fricative (/f/), the Japanese bilabial fricative ([Φ]) can be considered to "correspond" to the underlying representation of the labiodental fricative given in figure 1.7 (c). Thus, while Japanese, Chinese or Korean may not have a given English segment in its inventory, it may contain a very similar sound that could potentially factor into the acquisition of that English segment. For example, considering the Japanese bilabial fricative [Φ] again, the Japanese speaker may successfully perceive the English /p–f/ contrast by virtue of the fact that both English sounds can be categorized

Table 1.1 Status of English contrasts in Japanese, Chinese and Korean

	English Contrasts									
	/b/ vs. /v/		/p/ vs. /f/		/f/ vs. /v/		/s/ vs. /θ/		/r/ vs. /l/	
Status in Japanese Inventory										
is a native segment	√		√				√			
corresponds to a native segment				√[15]	√				√[16]	
does not correspond to a native segment		√				√		√		√
Status in Chinese Inventory										
is a native segment			√				√			√
corresponds to a native segment	√			√	√					
does not correspond to a native segment		√				√		√	√	
Status in Korean Inventory										
is a native segment			√				√			√
corresponds to a native segment	√									
does not correspond to a native segment		√		√	√	√		√	√[17]	

as different segments in the Japanese mental grammar (i.e., English /p/ will be funneled into the Japanese category /p/ and English /f/ will be funneled into the Japanese category [Φ]). In other words, even though a particular learner might not have the exact surface segment in his or her native language grammar, having a native segment that "corresponds to" the non-native segment may be advantageous in the acquisition of the non-native contrast. Notice that, with respect to phoneme inventories, the status of these English contrasts is similar for Chinese and Korean, which would lead us to expect to find comparable patterns of acquisition across these two groups of learners, if it is the phonemes of the L1 which constrain perception.

We can also use the consonant phoneme inventories in figure 1.8, along with our theory of underspecification, to determine which phonological features are used contrastively in these three languages; the adult feature geometries in figure 1.9 illustrate which features are manipulated.

The phonological features that we are interested in (i.e., the ones that differentiate the English contrasts at hand) appear bolded and underlined. For our immediate purposes, it is only the *features* (and not the entire adult *feature geometry* given in figure 1.9) that are relevant for determining whether a non-native contrast will be acquired. Nevertheless, I have included the entire adult geometry to underscore the claim made in this chapter that the full geometry is, in fact, a property of the speaker's mental grammar and, as such, reflects the organization of the perceptual system. Recall that it is this feature geometry that will operate during the process of speech perception

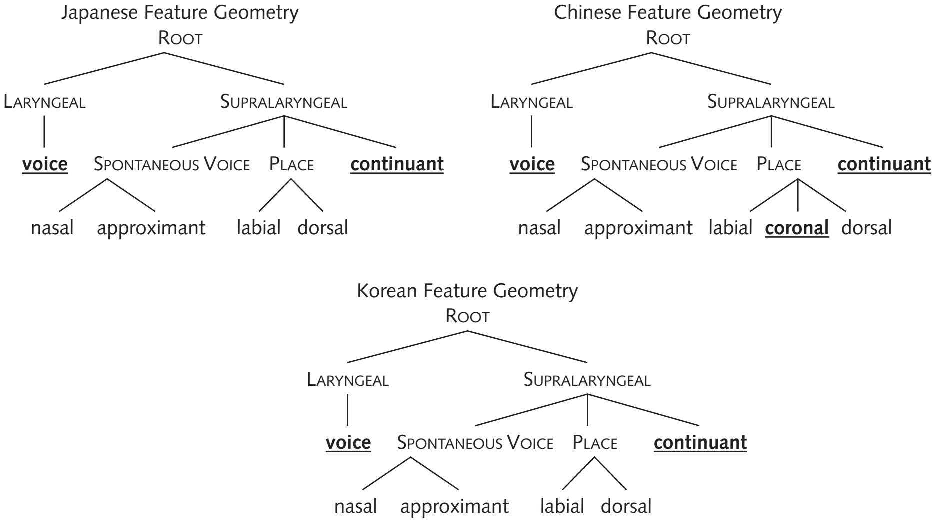


Figure 1.9 Feature inventories of L1 groups under investigation

to map the incoming continuous stimuli onto discrete perceptual categories (e.g., refer back to figure 1.5).

Notice that the features [continuant], which distinguishes the /b–v/ and /p–f/ contrasts, and [voice], which distinguishes the /f–v/ contrast, are present in the grammar of all of the three languages. Even though /b/, /v/, /f/ and /p/ are themselves not contrastive in these languages, other native segments differentiated by these particular features are contrastive. For example, the feature [continuant] is required in the Japanese grammar to differentiate native stop–continuant contrasts, such as the /t–s/ and /d–z/ contrasts, while the feature [voice] is present in the grammar in order to represent native voicing contrasts, such as /t–d/ or /s–z/. Thus, the feature that distinguishes /b–v/, /p–f/ and /f–v/ exists in the grammar for independent reasons.

However, the feature that distinguishes the /l–r/ contrast ([coronal]) is present only in the Chinese grammar (to distinguish the native alveolar /s/ and the retroflex /ʂ/); [coronal] is not present in the Japanese or Korean grammar as there are no consonants in either language that are distinguished from each other by this feature.[18] Finally, the feature that distinguishes the /s–θ/ contrast ([distributed]) is not utilized in any of the three grammars. Thus, in terms of features (as opposed to segments), Korean is more similar to Japanese than either is to Chinese.

Predictions

Recall that according to the theory of phonological interference being pursued here, it is the status of the contrasting *feature(s)* in the learner's native grammar (i.e., presence or absence) that determines the perception and subsequent acquisition of non-native contrasts. Of particular interest is the difference between the acquisition of a non-native contrast when the second language learner's native grammar does NOT contain the feature that distinguishes the segments and the acquisition of a non-native contrast when the learner's native grammar DOES contain the distinguishing feature. Taking Japanese learners of English as an example, three of the contrasts under consideration, /b–v/, /p–f/ and /f–v/, are distinguished by a feature present in the Japanese grammar, whereas /l–r/ and /s–θ/ are not. Thus, we would expect the former three contrasts to pattern together in acquisition, to the exclusion of the latter two.

More specifically, since perception of a non-native contrast is facilitated by the presence of the relevant feature in the learner's grammar, Japanese speakers should accurately perceive that /b/, /v/, /p/ and /f/ are distinct segments. This is by virtue of the fact that the feature [continuant] operates in the mental grammar of Japanese speakers, functioning to sort acoustic stimuli that differ along this dimension. Likewise, Japanese speakers should accurately perceive /f/ and /v/ as distinct segments, in this case, because the feature [voice] exists in their grammar. Since accurate perception of these two non-native contrasts is facilitated by the learner's native grammar, the learner will detect that these sounds are contrastive in English and acquisition of phonological representations for these segments will be triggered. On the other hand, since perception of a non-native contrast is blocked by the absence of the relevant feature from the learner's grammar, Japanese speakers will be unable to accurately perceive a contrast between /l/ and /r/ or between /s/ and /θ/. Lacking the features [coronal] and [distributed], the phonological system of the Japanese speaker's grammar will funnel the distinct acoustic stimuli for /l/ and /r/ into one perceptual category and for /s/ and /θ/ into another. Consequently, Japanese speakers will perceive instances of /l/ and /r/ as the same sound (likewise for /s/ and /θ/). Unable to perceive that they are distinct segments, the learner will not detect contrastive use of these sounds and, as a result, novel representations will not be acquired; consequently, these segments will not be distinguished in the learner's interlanguage grammar.[19]

A summary of these predictions for the acquisition of the English contrasts by Japanese, Korean, and Chinese learners is given in table 1.2. In addition to whether each language contains the relevant feature for a particular non-native contrast (and the resulting prediction for perception and acquisition), information regarding the segments themselves is also included: in particular, whether both members of the English contrast correspond to distinct segments in the L1 system (and could thus potentially be discriminated by the learner on the basis of those L1 categories). This information enables us to directly compare the predictions made by the "feature hypothesis" advocated here and the alternative "phoneme hypothesis" in which accurate perception and acquisition of non-native contrasts are thought to derive from the status

Table 1.2 Predictions for acquisition of English contrasts

English Contrasts	Japanese speakers			Chinese speakers			Korean speakers		
	Correspond to distinct segments in the L1	L1 grammar contains relevant feature	Learners will perceive & acquire	Correspond to distinct segments in the L1	L1 grammar contains relevant feature	Learners will perceive & acquire	Correspond to distinct segments in the L1	L1 grammar contains relevant feature	Learners will perceive & acquire
/b/ vs. /v/	No	continuant	Yes	No	continuant	Yes	No	continuant	Yes
/p/ vs. /f/	Yes	continuant	Yes	Yes	continuant	Yes	No	continuant	Yes
/f/ vs. /v/	No	voice	Yes	No	voice	Yes	No	voice	Yes
/s/ vs. /θ/	No		No	No		No	No		No
/l/ vs. /r/	No		No	No	coronal	Yes	No		No

of the *segments* themselves. In this regard, the most informative cases (i.e., the cases where the two hypotheses make opposite predictions) are the English /f–v/ contrast for all three language groups, the /l–r/ contrast for the Chinese speakers versus the Japanese speakers and the Korean speakers, the /p–f/ contrast for the Korean speakers versus the Japanese speakers and the Chinese speakers, and the /b–v/ contrast versus the /l–r/ contrast for the Japanese and Korean speakers.

Three Experimental Studies

Three experimental studies that test the predictions in table 1.2 are reported in the following section. We will find ample evidence that the critical factor predicting success is indeed the status of the L1 features, not the L1 segments.

Experiment 1

Experiment 1 was designed to test whether the theory of phonological interference outlined in this chapter could accurately account for variation in the acquisition of several different contrasts by learners with the same L1.[20]

Subjects

The experimental group consisted of 15 Japanese speakers, ranging in age from 20 to 32 years, who had learned English as their only second language. Each of these subjects was raised in Japan and had come to North America to study in an undergraduate or graduate program at McGill University in Montreal, Canada, where the testing was conducted. The control group consisted of 15 monolingual native speakers of English, who ranged in age from 15 years to 54 years. Their background information is summarized in table 1.3.

Contrasts investigated and hypotheses

The experimental contrasts were /l–r/, /b–v/ and /f–v/. A native contrast, /p–b/, was also included, as a control item for statistical comparison with the

Table 1.3 Experiment 1: Subject information

Group	Mean age at testing	Mean age of exposure	Mean years studied	Mean years in N. America
Japanese	24.5	9	8	3.5
Controls	25	–	–	–

non-native contrasts. As outlined above, our hypothesis is that the Japanese speakers will accurately perceive the /b–v/ and /f–v/ contrasts, but not the /l–r/ contrast. Moreover, they should perceive these two non-native contrasts and the native /p–b/ contrast equally well. Again, successful acquisition of these two non-native contrasts, to the exclusion of /l–r/, is predicted to follow.

Tasks and materials

An AX Discrimination task was used to assess the subjects' ability to acoustically discriminate (i.e., perceive) the English contrasts. In this task, subjects hear a minimal pair (one item containing, for example, an /l/ and the other item containing an /r/ in onset position) and are asked to indicate whether the words are the same or different (e.g., *rip / lip*).[21] The items used in the test were natural tokens of real English monosyllabic words. These tokens were spoken by a man with a standard American English accent and taped so that the stimuli were identical for every subject.

Two types of foils were also included in the test materials. One foil type, which consisted of native contrasts, was included as a means of checking that poor performance on the task was not due to difficulty with the task itself; any difficulty with the task would be reflected in poor performance on the native contrasts as well as the non-native contrasts. This native-contrast foil comprised 25 percent of the entire stimuli set (i.e., equal to the number of stimuli for each experimental contrast). A second type of foil, which consisted of identical pairs of words, was included to detect any response biases. Since each experimental minimal pair differs with respect to some consonant, the correct response for every trial is that the words are different. This set of foils ensured that any response bias or strategy toward responding that all of the pairs were different would result in inaccurate performance. If a subject responded that these same-word stimuli were different, his or her data were discarded. This type of foil comprised 20 percent of the stimuli for each individual contrast (i.e., three of the stimuli for each contrast were same-word and 12 were different-word).

The second aspect investigated is whether those subjects who can acoustically discriminate the non-native contrast are able to acquire the phonological structure necessary to distinguish the two sounds phonologically. Based on the arguments in Brown and Matthews (1997), data from comprehension tasks are assumed to be a more accurate indication of the learner's underlying competence than production abilities. Since production involves several peripheral mechanisms, such as motor control, relying on production data may lead us to underestimate or (particularly in the case of adults) overestimate the learner's underlying phonological competence. Several researchers have, in fact, demonstrated that L2 learners may be able to accurately produce a non-native contrast even though the same learners are unable to distinguish the two sounds perceptually (Brière, 1966; Flege, Takagi and Mann, 1995;

Goto, 1971; Sheldon and Strange, 1982).[22] Thus, if we rely on production data we may falsely attribute more segmental structure to a learner's underlying phonological competence than he or she actually has. On the other hand, some L2 learners are more like young children in that they are unable to correctly produce a novel contrast despite their ability to perceive and accurately distinguish that contrast phonologically in comprehension, in which case we would underestimate the learner's competence. Of course, this is not to deny the operation of peripheral mechanisms in comprehension (it too is performance), nor would I claim that comprehension can be equated with underlying competence. However, it is reasonable to assume that performance on comprehension tasks (where all context has been removed and the subject must access his or her mental representations in order to perform the task) provides a more accurate picture of grammatical competence than does performance on production tasks and is, therefore, to be preferred.

Subjects were given a Forced Choice Picture Selection task (modified from Brown and Matthews, 1993, 1997), in which the subject is presented with two pictures and a verbal cue that corresponds to one of the pictures.[23] For example, the subject would see a picture of a *rake* on the left side of the page and a picture of a *lake* on the right side. At the same time, the subject would hear the word *lake*. The subject's task is to indicate which of the pictures the verbal cue names. In order to successfully complete this task, the learner must refer to his or her internal phonological representations of the pictured objects and determine which lexical representation corresponds to the verbal stimulus. If the subject's lexical representations of the pictured objects are identical (i.e., if they do not have the necessary phonological structure to contrast /l/ and /r/), then he or she will be unable to determine to which picture the verbal cue corresponds and should perform the task with chance accuracy. Successful completion of this task indicates that the subject has acquired the non-native contrast. The monosyllabic words used in this task were the same as those used in the AX Discrimination task. Both tasks were administered on the same day, with a short break between tasks.

Results and discussion

For the auditory discrimination task, a response that the two words in the minimal pair were "different" was counted as correct and a "same" response was counted as an error. Performance scores on each of the contrasts were tabulated separately for statistical analysis. The graph in figure 1.10 reports the mean performance scores of both groups on each of the contrasts.

From the Japanese subjects' near perfect performance on the native /p–b/ pairs, it is clear that the task itself does not pose any difficulty for the learners. Thus, performance on the non-native contrasts can be interpreted to reflect properties of the speakers' interlanguage grammar. As can readily be seen from the graph in figure 1.10, the Japanese speakers were significantly poorer

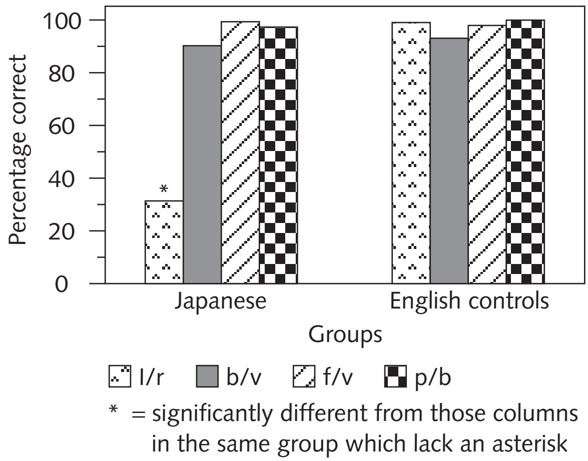


Figure 1.10 Experiment 1: Overall auditory performance by group

than the English controls at discriminating the /l–r/ contrast [t (28) = –16.16, p = .0001]. Yet, there was no statistical difference between the two groups in their ability to perceive the other contrasts; the Japanese speakers discriminated each of these English contrasts as accurately as the native controls [/b–v/ contrast: t (28) = –1.28, p = .21; /f–v/ contrast: t (28) = 1.87, p = .08; /p–b/ contrast: t (28) = –1.46, p = .15]. The Japanese speakers' performance on the /b–v/ and /f–v/ contrasts is quite striking: despite the fact that these are both non-native contrasts, they are very good (in fact, native-like) at discriminating each of them.[24] This suggests that these subjects perceived /b/, /v/ and /f/ as distinct speech sounds.

In order to evaluate performance on each contrast relative to the other contrasts, additional analyses were carried out separately on the two groups.[25] Beginning with the Japanese group, we find that their performance on the /l–r/ contrast is significantly worse than their performance on the other three contrasts [F (14, 45) = 119.85, p = .0001]; however, their performance on the other three contrasts (/b–v/, /f–v/, /p–b/) was not significantly different from one another. Thus, the Japanese speakers are unable to discriminate /l/ from /r/, perceiving them, instead, as a single category.

One might be tempted to surmise that the Japanese speakers are unable to discriminate the /l–r/ contrast precisely because they have allophonic experience with these segments: given the allophonic variation in the native language, we could imagine that these speakers have been "trained" to ignore these variations (recognizing them both as instantiations of the same underlying phoneme). Under this account, the allophonic exposure, rather than the absence of the contrasting feature, would be responsible for their lack of perceptual sensitivity to this contrast. However, this allophonic explanation cannot be correct because /b/ also varies allophonically in Japanese, with the voiced bilabial fricative [β], which shares acoustic and phonological properties with English /v/ (Kawakami, 1977).[26] Thus, if allophonic variation were the cause of the learners' inability to accurately perceive certain non-native contrasts, we would (incorrectly) expect perception of the /l–r/ and /b–v/ contrasts to be similarly impaired, which, as we see in this experiment, is certainly not the case.

In contrast to /l–r/, though, not only do the Japanese speakers perceive the non-native /b–v/ and /f–v/ contrasts with native-like accuracy, as predicted, but they perceive them equally well. This is what we would expect, given that they are both distinguished by a feature in the Japanese grammar. Note that uniform performance is not predicted if the aspect of the native grammar responsible for filtering non-native sounds is the phonemic representations themselves (rather than the features). Recall from table 1.1, that the members of the /b–v/ and /f–v/ contrasts have a different status *vis-à-vis* the Japanese phoneme inventory. Hence, if the phonemic representations constrain perception, we might expect differential perception of these two pairs, since the segment /b/, but not /f/, occurs in Japanese.

Moreover, that these non-native contrasts are discriminated as well as the native /p–b/ contrast suggests that perception of non-native sounds operates in the same manner as perception of native sounds. Although the native controls' performance on the /b–v/ contrast appears to be depressed relative to the other contrasts, there is, in fact, no statistical difference between the four contrasts [$F\ (14, 45) = 1.44$, $p = .25$]. Thus, we can regard the Japanese speakers' performance, in light of the native speaker data, as a true reflection of their perceptual capabilities.

To summarize, then, we have found a difference in the Japanese speakers' ability to perceive non-native contrasts, depending on whether the feature that distinguishes a given contrast exists in their grammar: the /l–r/ contrast (whose contrasting feature is absent from the Japanese grammar) is not accurately perceived, whereas the /b–v/ and /f–v/ contrasts (whose contrasting features are contained in the Japanese grammar) are accurately perceived, in a native-like manner (with respect both to the English controls and to the native Japanese contrast).

For the picture identification task, selection of the target picture was counted as a correct response, and selection of the contrast picture was counted as an error. The groups' overall performances are compared in figure 1.11.

Near perfect performance was attained by the Japanese group on the control items in this task. As in the auditory task, the Japanese speakers were significantly poorer than the English controls at differentiating the /l–r/ contrast [$t\ (28) = -9.73$, $p = .0001$]. There was, however, no statistical difference between the two groups in their ability to discriminate the other contrasts [/b–v/ contrast, $t\ (28) = -1.8$, $p = .08$; /f–v/ contrast, $t\ (28) = -.32$, $p = .75$; /p–b/ contrast, $t\ (28) = -1.27$, $p = .22$]. When shown two pictures that constituted a minimal pair (e.g., *rake*, *lake*), the Japanese subjects were unable to correctly choose the one that corresponded to the verbal cue. Yet

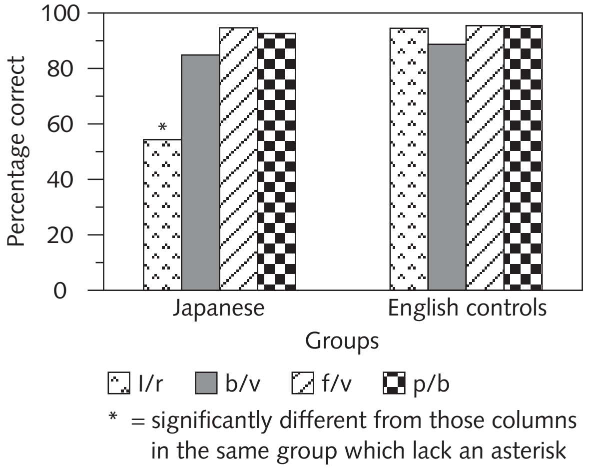


Figure 1.11 Experiment 1: Overall picture performance by group

these subjects performed this task with native-like accuracy when the pair of pictures differed by /b–v/ (e.g., *boat*, *vote*) or /f–v/ (e.g., *fan*, *van*).

Let us now compare performance on each of the contrasts relative to each other. In this analysis, too, the Japanese speakers' performance on the /l–r/ contrast was significantly worse than performance on the other three contrasts, while performance on the /b–v/, /f–v/ and /p–b/ contrasts was uniform [$F(14, 45) = 57.65$, $p = .0001$]. This pattern of performance is confirmed by an examination of individual scores. Looking at the native controls, their mean performance on each of the contrasts was not significantly different from each other [$F(14, 45) = 1.56$, $p = .21$]. Thus, as with the auditory task, the performance of the Japanese subjects can be taken to accurately reflect their underlying phonological competence.

Although the performance of the Japanese subjects on the /l–r/ contrast was lower than their performance on the other contrasts, their accuracy rate (almost 60 percent) would seem to indicate that these learners have some knowledge of the /l–r/ contrast. However, in order to correctly interpret these results, it is necessary to consider the expected baseline performance on this task, that is, what chance performance would be. Suppose that a learner has no phonological knowledge of the /l–r/ contrast and is, therefore, unable to distinguish /l/ from /r/ in lexical representations. When that subject is presented with two pictures and a single verbal cue, he or she simply will be unable to decide which picture corresponds to the cue (i.e., since the representation for both items is the same, both correspond to the verbal cue). With a choice between two pictures, this subject has a 50 percent chance of choosing the correct one, just by guessing.[27] The observed performance, then, at 60 percent, is not significantly different from chance. We can infer with reasonable confidence from the Japanese speakers' performance on this task, that /l/ and /r/ are not differentiated in their grammars. Performance on the other contrasts, on the other hand, indicates that both the /b–v/ and /f–v/ contrasts are differentiated. In other words, the phonological structure that represents the /b–v/ and /f–v/ contrasts has successfully been acquired by these learners.

The hypothesis guiding this experiment was that perception of non-native contrasts is constrained by the phonological features manipulated in the native grammar of the learner. This led us to predict that Japanese learners of English would accurately perceive the /b–v/ and /f–v/ contrasts, as these two pairs are differentiated by features already present in the Japanese grammar, but that accurate perception of the /l–r/ contrast would be blocked by the absence of the relevant feature from the Japanese grammar. Each of these predictions was borne out by the data. The Japanese speakers' inability to perceive /l/ and /r/ as distinct phonemes can be understood as a direct consequence of the influence of the native grammar on the operation of the speech perception mechanism. The Japanese speakers' perception of /b/, /f/ and /v/ as distinct phonemes likewise provides experimental support for the model of phonological interference outlined in this chapter. In a similar vein, since acquisition of a phonemic contrast is dependent upon accurate perception of that contrast, we predicted that Japanese learners would successfully acquire the /b–v/ and /f–v/ contrasts, but would fail to acquire the /l–r/ contrast. These predictions, too, were confirmed by the data. The Japanese learners successfully acquired only those non-native contrasts which they accurately perceived. The finding that Japanese speakers do not accurately perceive the difference between /l/ and /r/ is not particularly surprising, given the large body of literature that reports this observation (Goto, 1971; Miyawaki et al., 1975; Sheldon and Strange, 1982; Strange and Dittmann, 1984; Yamada, 1995). However, this current research is the first to also examine and compare Japanese speakers' perception of additional English contrasts; it is also the first to investigate these speakers' acquisition of the /l/ and /r/ feature geometric representations.

The theory of phonological interference that has been tested in experiment 1 has correctly accounted for differences in Japanese learners' abilities to acquire different English phonemic contrasts. The theory further predicts that speakers of different L1s which differ in the features that they utilize will exhibit differing success rates of acquiring various contrasts. The following experiment tests whether this model of phonological interference can account for *cross-language* differences in L2 phonological acquisition.

Experiment 2

The purpose of this experiment was to examine differences in the acquisition of English contrasts by speakers of different native languages, and to replicate the findings in experiment 1 for Japanese speakers.

Table 1.4 Experiment 2: Subject information

Group	Mean age at testing	Mean age of exposure	Mean years studied
Japanese	20.3	11.8	8
Chinese	30.7	12.6	10.4
Korean	30	12.8	9.9
Controls	34.3	–	–

Subjects

A total of 51 subjects, divided into one control group and three experimental groups, participated in this study. One experimental group consisted of 15 undergraduate Japanese speakers who were learning English at Hokkaido University, Sapporo, Japan and had never lived in an English-speaking country. The second experimental group consisted of 15 native Mandarin Chinese speakers who were enrolled in graduate programs at Hokkaido University (and, therefore, were proficient in Japanese). Eleven native speakers of Korean (also proficient in Japanese) comprised the final experimental group; these subjects were also graduate students at Hokkaido University. Neither the Chinese native speakers nor Korean native speakers were enrolled in English classes at the time of testing. The control group consisted of ten native monolingual speakers of American and British English, who teach English at universities in Sapporo, Japan. Table 1.4 summarizes the background information for each of the four groups.

Contrasts investigated and hypotheses

Acquisition of the following contrasts was investigated: /p–f/, /f–v/, /s–θ/ and /l–r/, with the /p–t/ contrast (a native contrast for all groups) serving as a control item. Speakers of Chinese, Korean, and Japanese were chosen for comparison because the grammars of these languages differ in interesting, theoretically relevant ways. In particular, when we compare the status of English contrasts in L1 phoneme inventories (table 1.1), Chinese and Korean appear to be more similar to one another, which might lead us to expect that speakers of these two languages would have the same difficulty (or success) acquiring the English contrasts under investigation. However, when the features employed in each of these grammars are compared (table 1.2), Korean and Japanese are more similar to one another. Thus, examining acquisition by all three groups should provide evidence as to which level of phonological knowledge is responsible for L1 interference. To briefly review the predictions set out in table 1.2, according to the theory of phonological interference adopted in this study, speakers of all three languages should accurately perceive and have successfully acquired the /p–f/ and /f–v/ contrasts since each of the L1 grammars utilizes the contrasting features ([continuant] and [voice], respectively) to distinguish native segments. Likewise, as the three L1 grammars lack the feature that contrasts /s/ and /θ/ ([distributed]), we predict that these two segments will be perceived (inaccurately) as a single category by speakers of all three languages; unable to hear a contrast between the two segments, they will also fail to acquire the feature geometric structure necessary to distinguish them phonologically in their interlanguage grammars. Finally, the three language groups should differ with respect to their ability to perceive and acquire the /l–r/ contrast: Chinese speakers, whose L1 contains the feature [coronal], will accurately perceive and, therefore, acquire this contrast; whereas, the adult feature geometry of Japanese speakers and Korean speakers will fail to sort the acoustic signal for these two sounds into distinct perceptual categories and their acquisition of the novel segmental representations will be prevented.

Tasks and materials

Phonological competence was assessed with the same Forced Choice Selection task used in experiment 1 above. A 4IAX Discrimination task, rather than an AX task, was used to assess perception in this study. In the 4IAX task, each trial consists of two pairs of words (Pisoni, 1971); in one of those pairs, the two words will be different (i.e., a minimal pair), and the other pair of words will be the same (e.g., *ra/ra, ra/la*). The subject's task is to indicate which of the two pairs of words is different. This task is becoming increasingly employed in speech perception research, since the AX task has been argued to bias the subject to respond "same" when discrimination is difficult (Beddor and Gottfried, 1995). The 4IAX task avoids this response bias since the subject knows that one of the pairs is, in fact, different and must simply determine which one.

The stimuli for this task, again in contrast to experiment 1, were nonwords in order to prevent the subjects' perception from being influenced by their familiarity with particular lexical items (Yamada, Kobayashi and Tohkura, in press). The "same" pairs consisted of two instances of a CV (consonant–vowel) syllable whose onset consonants were members of the same phonemic and phonetic category (e.g., aspirated [p^ha]), but which were not physically identical. Thus, subjects could not accurately choose the "same" pair (and thereby determine the "different" pair) simply by comparing physical objects and attending to non-linguistic acoustic variations, such as amplitude or speed. The "different" pairs consisted of two CV syllables whose onset consonants were members of different phonemic categories. Given these type of stimuli, accurate performance on this task requires the learner to filter out irrelevant variations across the segments and respond to higher-order phonological information. Stimuli were recorded by a male speaker of standard American English onto a Sony DAT Workstation and then arranged temporally by computer to create uniform intervals of 1000 milliseconds between members of a pair, 1800 millisecond intervals between pairs in a trial and 3000 millisecond intervals between trials. These time intervals were chosen following Werker and Logan (1985) to ensure phonemic processing of the stimuli.

Results and discussion

For each 4IAX trial, selection of the pair whose members were from different phonemic categories was counted as correct and selection of the pair whose members were from the same phonemic category was counted as an error. The mean performance scores of all groups on each of the contrasts are reported in figure 1.12.

The comparison that we are primarily interested in here is between the performance of the three language groups on particular/individual contrasts. As can be seen from the graph, the Japanese and Korean speakers are not as good as the Chinese speakers at discriminating the /l–r/ contrast. Statistical analyses reveal two distinct perceptual patterns: the Chinese speakers' performance is not significantly different from the native controls' performance, while the performance of the Japanese and Korean speakers is significantly worse than the Chinese speakers and native controls, but they are not significantly different from each other [$F(3, 47) = 16.39$, $p = .0001$]. As we saw in experiment 1, the Japanese speakers are unable to distinguish spoken tokens of /l/ and /r/; we see that Korean speakers, too, perceive these segments as a single category (confirming Borden, Gerber and Milsark's 1983 findings), whereas the Chinese speakers have no problem discriminating this contrast.

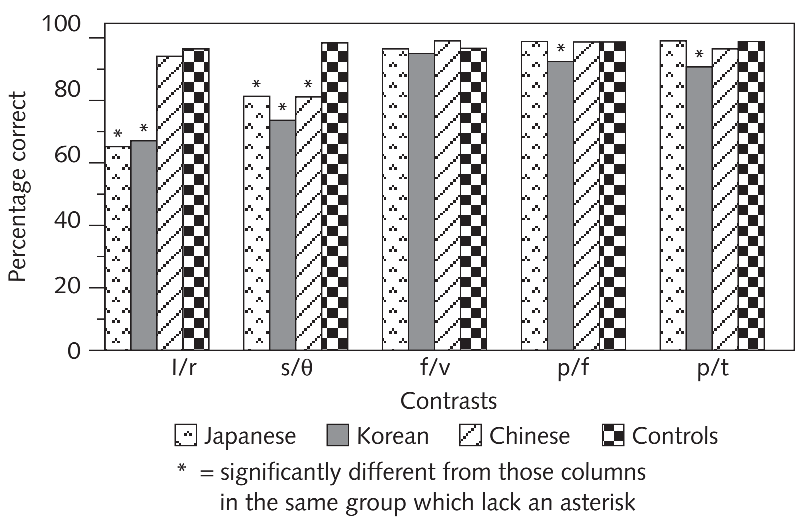


Figure 1.12 Experiment 2: Overall auditory performance by contrast

Turning to the /s–θ/ contrast, we find that the Japanese, Chinese and Korean speakers all discriminated this contrast equally poorly; they are significantly worse than the native controls, but not significantly different from each other [$F(3, 47) = 3.8$, $p = .016$]. The acoustic signals for these two sounds are funneled into the same perceptual category by speakers of all three language groups. With respect to the /f–v/ contrast, we also find consistent performance but, in this case, the groups are equally good, and not significantly different from the native controls [$F(3, 47) = 1.49$, $p = .23$]. The feature [voice] in the L1 grammar serves to separate, and keep distinct, the acoustic signals for these two sounds as they are processed.[28] Performance on /p–f/, the other non-native contrast that is distinguished by an L1 feature manipulated by all three L1s, is roughly uniform for all groups, although we do find a small statistical difference between the experimental groups [$F(3, 47) = 2.93$, $p = .04$].

The Japanese and Chinese speakers are able to discriminate these two sounds as accurately as the native controls, though the Korean speakers are significantly worse than both experimental groups and the native controls. However, the Korean speakers' performance, at 90 percent accuracy, is still well above chance and can be considered accurate, albeit not native-like. Finally, all of the groups are able to accurately discriminate the native /p–t/ contrast, though small differences among groups do approach statistical significance [$F(3, 47) = 2.6$, $p = .06$]. Once again, the performance of the Korean speakers (91 percent) is a bit lower than the other groups. Given that /p–t/ is a native contrast for these speakers, this result is somewhat surprising and suggests that the Korean speakers' performance on all of the contrasts in this task is slightly depressed. Overall, the performance of the Japanese, Korean and Chinese speakers on the /f–v/ and /p–f/ contrasts is quite remarkable: despite the fact all three languages lack these two contrasts, these learners perceive them with native-like accuracy. Their ability to perceive these contrasts is particularly striking in light of their inability to discriminate the /s–θ/ contrast. Moreover, the differing ability of the Chinese speakers, on the one hand, and Japanese and Korean speakers, on the other, to accurately perceive the /l–r/ contrast indicates that the ability (or inability) to perceive non-native contrasts is linked to phonological properties of those contrasts and the L1 grammars, not to acoustic properties of the sounds themselves.

Although our main interest in this study is in differences between groups, it is still informative to consider performance on each of the contrasts relative to the others. In order to make such comparisons, additional statistical analyses were carried out separately for each group. It should be kept in mind that baseline performance on the 4IAX task is different than on the AX task. Recall that in the AX task the subject's decision is whether the two words are the same or different. If a subject cannot hear a difference between two sounds, then he or she will respond "same". In this case, performance would theoretically be 0 percent accuracy. In other words, given the influence of the native grammar, the probability of responding "same" or responding "different" is not 50 percent. In contrast, the subject's decision in the 4IAX task is which pair of sounds is different. If the subject's L1 grammar causes him or her to hear both pairs of words as being the same, the choice is still between "first pair" and "second pair" and the probability of randomly choosing either one is 50 percent. Thus, an inability to perceive a contrast in the AX task would result in 0 percent accuracy, whereas an inability to perceive a contrast in the 4IAX task would result in 50 percent accuracy. A consequence of this is that comparing performance on the 4IAX task across different phonemic contrasts is more difficult since differences between performance on contrasts that are perceived accurately and those that are not will be smaller (50 percent–100 percent; cf. AX task: 0 percent–100 percent); for the same reason, scores from the two tasks cannot be directly compared.

Starting with the Chinese group, we find they are equally good at discriminating the /p–f/, /f–v/ and /l–r/ contrasts, and with the same accuracy with which they distinguish their native /p–t/ contrast; they discriminate all of these contrasts significantly better than they do the /s–θ/ contrast [$F\ (14, 60) = 8.55$, $p = .0001$]. This is what we would expect, given that the former non-native contrasts are distinguished by a feature contained in the Chinese grammar, whereas the latter is not. The Japanese group, too, discriminate the /p–f/ and /f–v/ contrast with the same accuracy that they discriminate their native contrast, and they are significantly better at perceiving these contrasts than they are the /s–θ/ or /l–r/ contrasts [$F\ (14, 60) = 29.78$, $p = .0001$]. These speakers do not, however, perceive the /s–θ/ and /l–r/ contrasts equally poorly; their discrimination of /l–r/ is worse than their discrimination of /s–θ/. It appears that, even though the acoustic signals for both sets of sounds will each be funneled into their respective category and perceived as the same sound, in a temporally adjacent presentation, the Japanese speakers are able to distinguish /s/ and /θ/ (but not /l/ and /r/) with higher accuracy than would be predicted by chance (possible reasons for this difference are discussed below). We find a similar pattern with the Korean speakers: /l/ and /r/ are discriminated less accurately than /s/ and /θ/ and performance on both of these contrasts is significantly worse than on the other contrasts [$F\ (10, 44) = 9.06$, $p = .0001$].

The most important thing to note from these perceptual data is that Japanese speakers and Korean speakers differ from Chinese speakers in their ability to discriminate /l/ and /r/. This difference between the language groups might seem surprising given that all three languages lack this phonemic contrast. However, it can be properly understood as a direct consequence of the influence of the phonological features in their respective native grammars: the presence of the feature [coronal] in the grammar of Chinese speakers ensures that acoustic stimuli which differ on this dimension will be perceived as distinct, whereas the absence of the feature from the Japanese and Korean grammars causes the acoustic signal for these two sounds to be funneled into

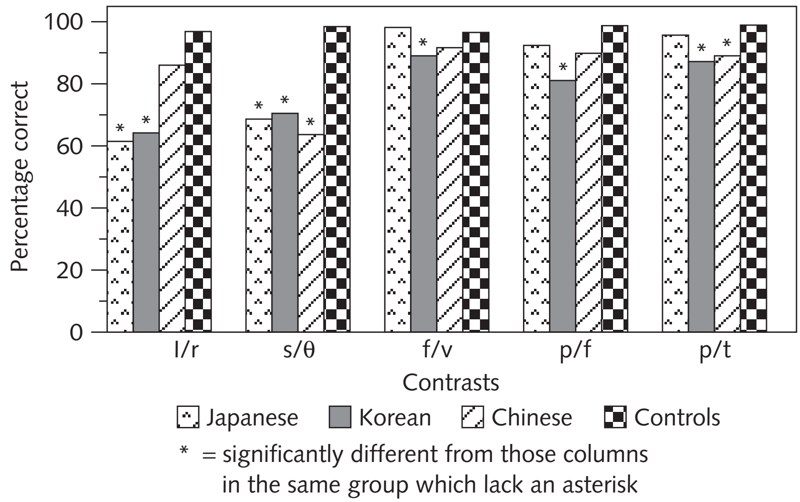


Figure 1.13 Experiment 2: Overall picture performance by contrast

a single perceptual category. The three language groups do not differ in their ability to accurately discriminate those contrasts which are distinguished by a feature that exists in all three L1s or in their inability to perceive those contrasts which are distinguished by a feature not utilized in their native grammars.

Figure 1.13 compares the groups' overall performance on the picture selection task. With respect to the /l–r/ contrast, we find that, as in the auditory task, the Chinese speakers perform more accurately than the Korean speakers and the Japanese speakers [$F\ (3, 47) = 21.35$, $p = .0001$]; in fact, they perform as well as the native controls. Chinese speakers have no problem choosing between two pictures that constitute a minimal /l–r/ pair, indicating that these two phonemes have distinct representations in their interlanguage grammars. The Japanese and Korean speakers, however, were significantly worse on this contrast than the Chinese speakers and native controls, though not different from each other; thus, neither of these two groups of speakers distinguishes /l/ and /r/ phonologically.

All three groups of learners were unable to perform this task accurately when the lexical items differed by /s–θ/; there was no difference between experimental groups, and their performance was significantly lower than the controls' performance [$F\ (3, 47) = 11.53$, $p = .0001$]. This indicates that a new segmental representation for /θ/ has not been acquired by the learners, as we predicted, so /s/ and /θ/ are represented by the same geometric structure in their interlanguage grammars. In contrast, lexical items containing /f/ and /v/ are distinguished phonologically by all three groups of learners, as indicated by their high performance levels, though the Korean speakers' performance, at 89 percent accuracy, is slightly worse than the native controls' performance

[F (3, 47) = 3.21, p = .03]. Similarly, on the /p–f/ contrast, the Japanese and Chinese speakers are as accurate as the controls, while the performance of the Korean speakers, though significantly lower, is still well above chance (83 percent) [F (3, 47) = 5.9, p = .002]. With respect to the native /p–t/ contrast, all language groups distinguish /p/ and /t/ in their interlanguage grammars. Japanese speakers distinguished these two sounds in a native-like fashion; however, both the Chinese and Korean speakers were just slightly less accurate than the native speakers [F (3, 47) = 3.22, p = .03]. Overall, then, we see that the learners in all three groups have distinct segmental representations for /p/, /f/ and /v/, while the Chinese speakers also have distinct representations for /l/ and /r/, and none of the learners have distinct representations for /s/ and /θ/. Let's now compare performance on the different contrasts by each group individually to confirm these acquisition patterns.

A separate analysis of the Chinese group reveals that the /s–θ/ contrast is distinguished much more poorly than the other contrasts, including the native /p–t/ contrast [F (14, 60) = 13.4, p = .0001]; the /l–r/, /p–f/ and /f–v/ contrasts, however, are distinguished equally well and as accurately as the native /p–t/ contrast. This means that /l/, /r/, /p/, /f/ and /v/ each have a distinct segmental representation in the Chinese speakers' interlanguage grammars; /s/ and /θ/, on the other hand, will correspond to the same phonological structure and, therefore, will not be distinguished in these learners' interlanguage grammars.

Analysis of the Japanese data also confirms two distinct acquisition patterns [F (14, 60) = 34.99, p = .0001]. These learners represent the /p–f/ and /f–v/ contrasts in their interlanguage grammars in the same way that they represent the native /p–t/ contrast; there is no statistical difference between their (equally good) performance on these three types of contrasts. There is also no difference in their (in)ability to phonologically distinguish the /l–r/ and /s–θ/ contrasts: they are equally poor. This finding is especially interesting, given the difference we found between these two contrasts on the auditory discrimination task. Despite the slight advantage in perceiving /s/ and /θ/, it is not sufficient to trigger acquisition. Neither the segmental representation for /l/ nor for /θ/ has been acquired by these learners.

Finally, we turn to the Korean group, who have the identical pattern of acquisition as the Japanese speakers: the /l–r/ and /s–θ/ contrasts are not distinguished in lexical items, whereas the /p–f/ and /f–v/ contrasts are [F (10, 44) = 12.8, p = .0001]. In fact, although we saw above that the performance of the Korean speakers was slightly depressed on the /p–f/ contrast relative to the other language groups, it is not significantly different from their performance on their native contrast. Thus, it appears that whatever is causing the lower performance on the non-native contrast is not due to the non-native nature of the contrast, but rather to some more general performance factor. Nevertheless, additional studies examining Korean speakers' auditory and phonological discrimination are clearly required to establish their perceptual and linguistic abilities conclusively.

In summary, this experiment was conducted in order to determine whether our theory of phonological interference could account for the acquisition of English phonemes by speakers of different languages. Assuming that perception of non-native contrasts is constrained by the phonological features manipulated in the learner's native grammar and that languages differ as to the features they manipulate, we would expect learners with different L1s to differ in their ability to acquire particular non-native contrasts. Japanese, Korean, and Chinese differ in just this way.[29] Specifically, the grammar of Chinese contains the feature [coronal], whereas the grammars of Japanese and Korean lack this feature. Given this, speakers of Japanese and Korean, on the one hand, and Chinese, on the other, should differ in their ability to acquire the /l–r/ contrast, which relies on the feature [coronal]. This is, indeed, what we found. Chinese speakers accurately perceive this contrast and, therefore, successfully acquire it. Japanese and Korean speakers are unable to acquire this contrast since they do not perceive /l/ and /r/ as different segments.

These three groups were fairly evenly matched for age of exposure, education and years spent studying English. Therefore, we can be confident that the differential performance of these groups stems from their respective L1s. Simply comparing the phoneme inventories of these three languages, however, does not allow us to explain why Chinese speakers accurately perceive and acquire the /l–r/ contrast, but that both the Japanese and Korean speakers do not. By the same token, it is only by considering the features utilized by the L1s that we can adequately explain why all three language groups were able to perceive and acquire the /p–f/ and /f–v/ contrasts, which are distinguished by features that exist in Japanese, Chinese, and Korean.

Likewise, the absence of the relevant feature from all three L1s accounts for their uniform inability to acquire the /s–θ/ contrast. Thus, the differential abilities of the Japanese and Chinese speakers lends support to our theory of phonological interference: not only can we account for disparate acquisition of non-native contrasts by speakers of a single language, we can also explain disparate acquisition of a particular non-native contrast by speakers of different languages.[30]

Now that we have seen how the native grammar can affect perception and acquisition of non-native contrasts, a question that naturally arises is whether the native grammar always constrains phonological acquisition in this way or whether its effect changes over time, as the learner progresses. The Japanese learners tested in experiment 1 were relatively advanced, living in North America and receiving abundant natural English input. Since these learners had already acquired two of the three non-native contrasts, we found no evidence for any stages of acquisition. The learners tested in experiment 2 had never lived in an English-speaking environment and were receiving minimal to no aural English input at the time of testing; but, although these learners were not always as accurate as the native speaker controls on those contrasts they had acquired, we still did not observe distinct stages of acquisition. Since the first two experiments were not longitudinal and also did not compare learners with differing levels of L2 proficiency, we have no data to determine whether there is any change in learners' perceptual capacities. Does perception of non-native contrasts improve over time? Is there any effect of increased linguistic input? Is there evidence for stages of acquisition? These questions were addressed in the following experiment, which investigated the acquisition of English contrasts by low proficiency and higher proficiency Japanese learners of English.

Experiment 3

This experiment was conducted in order to determine whether the influence of the native grammar on the perception of non-native contrasts changes over time as the L2 learner progresses.

Subjects

The subjects for this experiment were 35 native speakers of Japanese and 10 native speakers of English. The control group comprised American, British, and Canadian English teachers at Hokkaido University and Hokkai Gakuen University, in Sapporo, Japan. The Japanese subjects were learning English as a foreign language at Hokkaido University and had never lived in an English-speaking country. Based on teacher interview assessment of their overall proficiency in English, the Japanese speakers were divided into two experimental groups: Low-level (n = 20) and High-level (n = 15). The relevant background data are given in table 1.5.

Contrasts investigated and hypotheses

Two experimental contrasts were tested in this experiment, /l–r/ and /b–v/; the native /p–b/ contrast was also included as a control item. If perception and acquisition of non-native contrasts is constrained by the features of the native grammar, then, since both beginner Japanese learners of English and

Table 1.5 Experiment 3: Subject information

Group	Mean age at testing	Mean age of exposure	Mean years studied
Low-level	19	11.7	7.6
High-level	24.5	12	11.5
Controls	35	–	–

more advanced Japanese learners of English have the same native grammar, they should both be able to perceive the /b–v/ contrast, yet unable to perceive the /l–r/ contrast.

Tasks and materials

The tasks and materials used in this experiment were the same as those used in experiment 1; an AX Discrimination task was used to assess perception and a Forced Choice Picture Selection task was used to assess phonological competence.

Results and discussion

On the auditory discrimination task, a response that the two words in the minimal pair were "different" was counted as correct and a "same" response was counted as an error. Performance scores on each of the contrasts were tabulated separately for statistical analysis. Figure 1.14 reports the mean performance scores on each of the contrasts for each group.

From both the Low-level and High-level groups' near perfect performance on the control items (i.e., native /p–b/ pairs), it is clear that the task itself does not pose any difficulty for the learners. Furthermore, the control group performed as expected, accurately discriminating each of the three contrasts, with no significant difference between contrasts [F (9, 20) = 1.09, p = .36].

As figure 1.14 illustrates, both groups of Japanese speakers were significantly worse than the English controls at discriminating the /l–r/ contrast [F (2, 44) = 74.49, p = .0001]. However, there was no difference between the Low-level and High-level groups in their ability to discriminate this contrast; learners in both groups were unable to perceive the difference between /l/ and

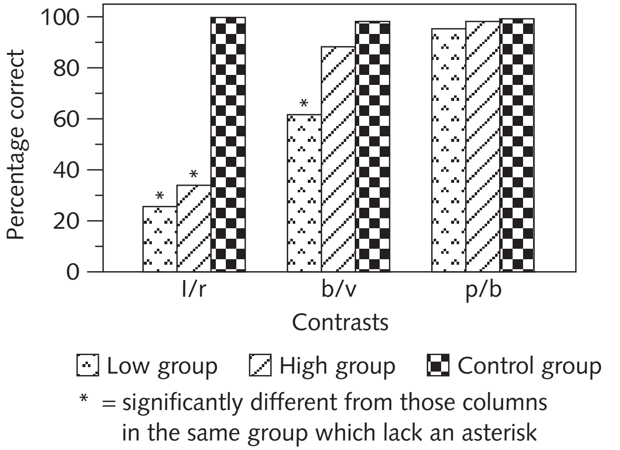


Figure 1.14 Experiment 3: Overall auditory performance by contrast

/r/. Thus, an increase in English proficiency does not appear to affect perception of this non-native contrast. Accurate perception is blocked by the native grammar in the earliest stages of acquisition and continues to prevent perception even as the learner progresses.

The situation is slightly different with respect to the other non-native contrast. While the learners in the Low-level group were not as accurate at discriminating the /b–v/ contrast as the learners in the High-level group, there was no difference between the High-level and the control groups' performance on this contrast [$F(2, 44) = 9.79$, $p = .0003$]. Thus, there was improvement in the Japanese speakers' ability to perceive this non-native contrast. We must keep in mind, though, that the Low-level group's somewhat poorer ability to discriminate /b/ and /v/ is still much better than either Japanese group's ability to distinguish /l/ from /r/. Finally, there was no statistical difference between the three groups in their ability to perceive the native /p–b/ contrast: all Japanese speakers discriminated this contrast as well as the native controls did [$F(2, 44) = 1.08$, $p = .35$]. In short, whereas the ability to accurately perceive the /l–r/ contrast does not improve over time, the ability to perceive the /b–v/ contrast does improve, from being fairly good to being native-like.

We can now evaluate the relative effect of the native grammar on each of the contrasts at different stages of acquisition by examining the performance of each group individually. Beginning with the Low-level group, we find that their performance on each of the contrasts is significantly different from each other [$F(19, 40) = 73.53$, $p = .001$]. That is, performance on the /p–b/ contrast, which is native-like, is better than performance on the /b–v/ contrast, which is better than performance on the /l–r/ contrast. However, there was no difference in the High-level learners' ability to discriminate the /b–v/ and /p–b/ contrasts; both were perceived equally well and more accurately than the /l–r/ contrast [$F(14, 30) = 91.75$, $p = .001$]. These data show that at both stages of acquisition, the Japanese speakers are unable to discriminate /l/ and /r/, perceiving them, instead, as members of a single category. However, they differ in their ability to distinguish /b/ from /v/, indicating that the influence of the native grammar is not static, but changes as the learner's interlanguage grammar develops.

To summarize, these data allow us to see the influence of the native grammar at different stages of acquisition. We found that the ability to discriminate the /l–r/ contrast does not change over time, whereas learners do improve in their ability to perceive the /b–v/ contrast. We might be tempted to conclude from this that the influence of the native grammar simply changes over time, constraining perception more tightly in the early stages of acquisition but gradually weakening as the learner's interlanguage grammar develops. However, the situation is a bit more complex. We know that the native grammar influences perception of non-native sounds in two ways: it may either *block* perception or *facilitate* perception, depending on whether the relevant feature

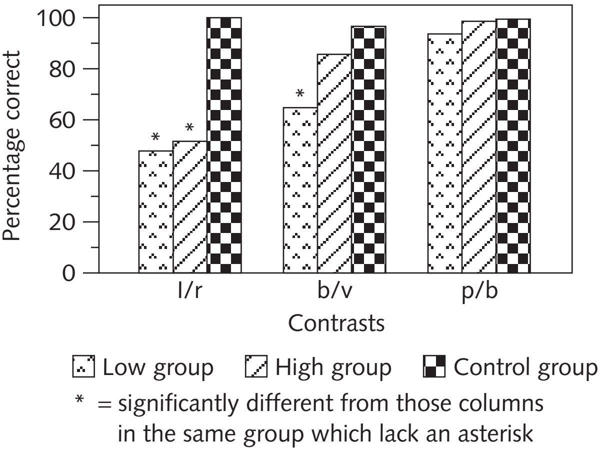


Figure 1.15 Experiment 4: Overall picture performance by contrast

is present or absent in the L1 grammatical system. Looking at the data again, we see that perception of the /l–r/ contrast does not improve; it is only the perception of /b–v/ which improves. Thus, when the relevant feature is absent from the native grammar, as it is in the case of /l–r/, and perception is blocked, the effect of the grammar remains constant. However, if the relevant feature is present in the native grammar, as it is for /b–v/, then the effect of the grammar may change. In other words, the negative influence of the native grammar on perception is absolute, but the positive influence of the native grammar is enhanced as the learner progresses.

For the picture identification task, selection of the target picture was counted as a correct response and selection of the contrast picture was counted as an error. The groups' overall performance is compared in figure 1.15.

Near perfect performance was attained by the Japanese group on the control items in this task. Again, the performance of the control subjects – accurate and with no differences between contrasts – ensures that our task and materials are reliable [F (9, 20) = 2.39, p = .12].

The pattern of performance on this picture task is very similar to that on the auditory task. Both groups of Japanese speakers were significantly worse than the English controls at distinguishing lexical items that differed by /l/ or /r/ [F (2, 44) = 35.20, p = .0001]. Yet, there was no difference between the Low-level and High-level groups in their ability (or inability) to distinguish this contrast; learners in both groups were unable to discriminate /l/ and /r/ phonologically. This indicates that neither the beginner learners nor the more advanced learners have acquired the phonological structure necessary to differentiate these segments in their interlanguage grammars. This is not the case, though, with the /b–v/ contrast. While the learners in the Low-level group were not as accurate as the learners in the High-level group at distinguishing items that differed by /b/ or /v/, there was no difference between the High-level and the control groups' performance on this contrast [F (2, 44) = 5.43, p = .007]. Thus, duplicating the results from the auditory task, the ability to differentiate /b/ and /v/ in one's interlanguage grammar appears to develop over time. This suggests that there are, in fact, stages of phoneme acquisition. With respect to the native /p–b/ contrast, there was no statistical difference between the three groups: all Japanese speakers discriminated this contrast as well as the native controls did [F (2, 44) = 1.94, p = .15].

Looking at each group individually, we find that the performance of the Low-level group on each of the contrasts is significantly different from the others [F (19, 40) = 47.81, p = .0001]. We find the same pattern of performance by the High-level group [F (14, 30) = 31.82, p = .0001]. Both groups are better at distinguishing the /b–v/ contrast than they are the /l–r/ contrast, but they are still not as good at distinguishing the /b–v/ contrast as they are their own native /p–b/ contrast. It is clear from the data that the learners do not differentiate /l/ and /r/ in lexical items (i.e., the same structure is used to represent both segments). It is also clear that they do have distinct representations for /p/ and /b/.

What, then, is the status of the /b–v/ contrast, which seems to fall somewhere between the other contrasts; has it been acquired or not? I think the answer to this question is different for the two groups. In the case of the Low-level group, it appears that the new representations have not been acquired. This is not so surprising, given their perception of the /b–v/ contrast, which while quite good is not native-like. It is possible that these learners have not yet detected contrastive use of these segments in English and, as a result, have not yet acquired the new representations. In the case of the High-level group, however, I think we can be confident that they have acquired the new representations. Importantly, their perception of /b–v/ is native-like; thus a necessary condition for proper acquisition has been met. Moreover, their ability to distinguish /b/ and /v/ in this task (although poorer than their ability to distinguish the /p–b/ contrast) is as good as the native speakers' ability to distinguish /b/ and /v/, who undoubtedly differentiate these two sounds in their grammars.[31]

The research question we attempted to answer in this experiment was whether the influence of the native grammar on the perception and acquisition of non-native contrasts changes over time as the L2 learner progresses. The data demonstrate that there is not one answer to this question. The effects of the grammar, either to *block* or to *facilitate* perception and acquisition are differentially altered by the learner's development. If the feature underlying a non-native phonemic contrast is absent from the native grammar, then the native phonological system will continue to funnel the distinct acoustic signals for those sounds into a single perceptual category throughout the learner's development; perception of these non-native contrasts will not improve. In this case, the influence of the native grammar is rigid, immutable by increased exposure to a second language. If, however, the feature underlying a non-native phonemic contrast is present in the native grammar, the capacity of the native phonological system to use this feature in the processing of non-native sounds will be enhanced over the course of the learner's development; perception of these non-native contrasts will improve. In this case, increased exposure to a second language will actually strengthen the facilitative influence of the native grammar.

General Discussion and Conclusions

The goal of the research program presented in this chapter is to develop a comprehensive theory-driven model of L2 phoneme acquisition that accounts for the interrelation between perception and phonological acquisition and explains how and why this interrelation affects L2 phonological acquisition. It was proposed that the monotonic acquisition of feature geometric structure by young children reduces their perceptual sensitivity to particular non-native contrasts and that this adult feature geometry continues to mediate between the acoustic signal and the linguistic processor in adult speech perception, constraining which non-native contrasts adult learners will be sensitive to in the L2 input and, therefore, capable of acquiring. Thus, in order to fully understand why the L1 grammar exerts such a profound influence on the perception and acquisition of non-native phonemic contrasts, it is important that we understand the development of these systems in first language acquisition and their operation in the mature speaker.

Having determined how the interrelation between speech perception and the phonological system originates, we are in a better position to capture the nature of the mechanism that maps the L2 input onto L1 phonological categories; utilizing the tools of Feature Geometry theory enables us to formally articulate this mapping process. The central claim of the theory of phonological interference developed here is that the L1 influence found in L2 phonological acquisition is a consequence of how the speech perception mechanism operates in the native speaker. Based on the proposal that the decline in infants' ability to discriminate non-native contrasts is caused by the acquisition of phonological structure, speech perception in the native speaker will continue to be constrained by phonological properties of his or her native language throughout adulthood; more specifically, all speech sounds (native and non-native) will be perceived in terms of the features exploited by that particular language.

The experimental studies reported above have demonstrated that not all non-native contrasts are created equal: learners with the same L1 have more difficulty perceiving and acquiring some non-native contrasts than they do others; likewise, certain non-native contrasts are easily perceived and acquired by speakers of some languages, while those same contrasts will not be perceived or acquired by speakers of other languages. These differences, both between contrasts and between speakers of different languages, were argued to follow directly from the status of the relevant distinctive feature in the learner's L1 grammar: presence of the contrastive features in the grammar serves to sort the acoustic signal along that particular dimension, mapping the signals for two segments onto distinct phonological categories, whereas absence of the contrastive feature entails that the acoustic signals for the phonemes be mapped onto a single phonological category.

We saw in experiment 1 that Japanese learners' perception and acquisition of various English contrasts differed in exactly this respect: they were able to perceive and, therefore, acquire those contrasts which are distinguished by a feature that their native grammar employs for independent reasons (e.g., /b–v/ and /f–v/), but were unable to perceive that contrast that is distinguished by a feature not utilized in the L1 (e.g., /l–r/). Similarly, experiment 2 provided evidence that Japanese, Korean, and Chinese speakers differ in their acquisition of particular English contrasts just as the model predicts: speakers of all three languages were unable to perceive the /s–θ/ contrast, which relies on a feature absent from all three L1s, yet were able to perceive those contrasts distinguished by features present in all three L1s (i.e., /p–f/ and /f–v/); most importantly, speakers of these three languages differed in their ability to perceive and acquire precisely that contrast, /l–r/, which is distinguished by a feature whose status differs among these languages.

While the discussion of experimental results has not focused on individual differences, some inter-subject variability does exist, and so it is necessary at this point to address how such variability fits into my model. It might seem that the model – with performance so heavily influenced by the speaker's L1 grammar – would allow for no variability between subjects having the same L1 background. In fact, though, we might expect two different kinds of inter-subject variability: "grammar-driven" variability and "test-strategy" variability. "Grammar-driven" variability would be those differences in subjects' performance that directly reflect properties of their interlanguage. This kind of variability will be restricted to those contrasts distinguished by a feature utilized in the L1. Since it is precisely these contrasts that are acquirable, we should find subjects at different stages of development: a learner in the early stages of L2 acquisition will assimilate the non-native segments to native categories to a greater degree (and thus have lower performance) than a more advanced learner whose grammar has developed the novel L2 categories. This situation is indeed what we observed in experiment 3 regarding the /b–v/ contrast – considerable (albeit, not wild) variability across the two levels of Japanese learners. Thus, although my model would predict that all Japanese speakers, for example, will eventually perceive this contrast in a native-like fashion, it does allow, even expect, that these speakers will vary somewhat in their accuracy according to their particular point in development. This "grammar-driven" variability will typically constitute a range of good performance on a task (e.g., 75 percent vs. 96 percent).

The other kind of variability – "test-strategy" variability – we would expect to find for those contrasts that are distinguished by a feature *not* present in the L1. As mentioned previously, in an AX task we would theoretically expect 0 percent accuracy since the L1 grammar funnels both L2 sounds into a single L1 category (i.e., the subject's answer will always be "same"). However, in practice subjects often surmise that some trials will be the same and some will be different and so adopt a strategy of simply guessing between "same" and "different". This strategy results in chance performance of approximately 50 percent. Thus, in the testing situation subjects may respond on the basis of their grammar, they may adopt a guessing strategy, or some combination of both. This would give us accuracy scores ranging from 0 percent to roughly 50 percent. In contrast to the "grammar-driven" variability, this "test-strategy" variability will always fall within the range of poor performance (e.g., 9 percent vs. 33 percent correct). Although "test-strategy" variability is not explicitly predicted to occur by my model it also does not compromise the model, providing that all scores are within chance levels (e.g., we should never find that some Japanese speakers perceive the English /l–r/ contrast with native-like accuracy and that other Japanese speakers do not).

If we expect to find some inter-subject variability for both kinds of non-native contrasts, those distinguished by a feature present in the L1 grammar ("grammar-driven" variability) and those by a feature absent in the L1 grammar ("test-strategy" variability), is there any empirical difference between the two? We can in fact make an interesting prediction regarding these two types of variability. Since "test-strategy" variability does not derive from the subjects' interlanguage, we should find this type *across* all levels of learners; in other words, we should find inter-subject variability for a contrast such as /l–r/ for both beginner and advanced Japanese or Korean learners of English. "Grammar-driven" variability, on the other hand, would be limited to variation *between* the different levels of learners since this type of variability directly reflects properties of the learners' grammars.

One implication of the model I have outlined is that prior to the development of a phonological system, infants should be able to perceive contrasts which they will fail to perceive as adults. With respect to Japanese speakers, if their difficulty discriminating /l/ and /r/ does indeed stem from the interference of their phonological system (rather than to, say, some genetic property of Japanese speakers), then before that system is in place, accurate perception of those sounds should be possible. Japanese infants have, in fact, been shown to perceptually distinguish /l/ from /r/ (Tsushima, Takizawa, Sasaki, Shiraki, Nishi, Kohno, Menyuk, and Best, 1994). Thus, the inability of Japanese speakers to discriminate /l/ and /r/ as adults does indeed appear to be a consequence of language development.[32] Moreover, an early study by Miyawaki et al. (1975) shows that the difficulties Japanese speakers have discriminating /l/ and /r/ are due to specific properties of their perception of *speech*, not to deficiencies in their basic auditory mechanisms. These researchers found that adult Japanese speakers accurately discriminate /l/ and /r/ when they are presented in a "non-speech mode".[33] In other words, Japanese speakers are able to discriminate /l/ and /r/ when the acoustic signal is processed directly by the auditory system, rather than the linguistic module.

We know from first language acquisition research that the development of segmental structure involves the interaction of Universal Grammar and the learner's detection of phonemic contrasts in the input. Thus, successful acquisition of novel phonemes by L2 learners depends not only on the availability of UG, but, importantly, on adequate intake to the language acquisition device. By demonstrating that some L2 learners do not perceive the L2 input correctly (in fact, precisely those learners who are unable to acquire the given contrast), this research strongly suggests that the inability of some L2 learners to acquire novel phonemic contrasts is due to the lack of proper input, rather than the unavailability of UG. Thus, the failure of L2 learners to acquire novel phonemes should not necessarily be taken as evidence that UG is not available in L2 acquisition. In fact, these results demonstrate that if L2 learners are able to perceive a non-native contrast, they are able to acquire that contrast, suggesting that the mechanism for constructing novel segmental representations (which is arguably part of UG) is still operative in L2 acquisition.

These findings fit in nicely with recent trends in second language acquisition theory which suggest that differences in L1 and L2 acquisition (as well as differences across learners in L2 acquisition) stem not from the unavailability of Universal Grammar but rather from the initial state of acquisition (papers in Schwartz and Eubank, 1996). The goal of this new line of research is to define the initial state of L2 acquisition and, thereby, explain the development of the L2 grammar. Differences in L2 acquisition of a particular language that co-vary with learners' native language are now assumed to be a result of the L2 initial state. The research agenda, then, is to define this initial state. Research on the acquisition of syntactic properties of the L2 has been used to support several hypotheses regarding the linguistic content of the initial state (see Schwartz and Eubank, 1996, for Schwartz and Sprouse's *Full Transfer/Full Access Model*, Vainikka and Young-Scholten's *Minimal Tree Hypothesis*, and Eubank's *Weak Transfer Hypothesis*). The findings from experimental research on the L2 acquisition of phonemes seem to be most consistent with Schwartz and Sprouse's hypothesis that the entire L1 system forms the initial basis of L2 acquisition: all of the data indicate that in the earliest stages of L2 acquisition, L2 phonemes are mapped according to the L1 feature geometry onto L1 phonemic categories; only subsequently are new L2 categories acquired.

The claim that the entire L1 phonological system constitutes the initial state for L2 phoneme acquisition raises an interesting question: If the acoustic signal is perceived in terms of the learner's L1 phonemic categories, how can the learner accurately perceive non-native sounds in order for new phonemic categories to be established? In other words, how can the input be mapped by the adult feature geometry onto new L2 categories when those categories don't yet exist? The answer to this question, I believe, depends on whether it is the *phonemes* or the *features* of the L1 which constrain perception. If the acoustic signal is mapped onto L1 phonemes, then it would seem that it is the phonemic level which impinges upon L2 acquisition. Yet, I have argued that it is the featural level which is relevant.

A closer examination of the data from low and high proficiency learners presented in experiment 3 suggests how these two positions might be reconciled. In particular, these data suggest that initially, in the earliest stages of acquisition, the phonemes of the L1 have a profound influence on the perception of non-native contrasts. In an attempt to understand the L2 input, and in the absence of new phonological categories, the L2 input is fitted into the L1 system any way it can be (often by brute force, ignoring variations that the system senses but cannot yet deal with appropriately). For example, the acoustic signals for both /b/ and /v/ will be mapped by Japanese speakers onto the L1 phoneme /b/ and the acoustic signals for /l/ and /r/ will be mapped onto /ɾ/. This will be the initial stage of acquisition and, incidentally, the situation for loanword phonology.[34]

However, despite the initial attempt of the L1 system to accommodate all of the input within L1 structures, portions of the L2 input will not map adequately to the L1 system, as a result of the presence of the relevant contrasting feature. Taking our example again, the English /b/ will map completely onto the Japanese phoneme /b/, but the English /v/ will not map precisely to the Japanese category.[35] Thus, although both are perceived as a single category in the early stages of acquisition, the presence of the feature [continuant] ensures that the acoustic signal for /v/ does not correspond exactly to a Japanese category; this slight mismatch between the L2 input and the L1 structures will cause perceptual reorganization (the beginner learners in experiment 3). Over the course of development, and with increased exposure, a new phonological category will be established; following the establishment of this new category, the original native category will be bypassed entirely in perception, and perception of those contrasts will be native-like (the advanced learners in experiment 3).

If, however, the feature that distinguishes a given non-native contrast is absent from the L1 grammar, then the L2 input will map perfectly onto an existing L1 category and there will be no trigger for acquisition, as was the case with both our beginner and advanced learners for the /l–r/ contrast. Thus, while the input is initially sorted in terms of L1 phonemes, it is the L1 features which guide this mapping process and, therefore, determine to what extent the L2 input can be accommodated by existing phonological structure; in this way the features also constrain which non-native contrasts will be acquired by the learner. This picture receives empirical support from recent research by Matthews and Brown (1998) who, using measures of reaction time, demonstrate that non-native contrasts that are distinguished by an L1 feature are processed differently on-line than those contrasts distinguished by a feature not used by the L1 grammar; they also show that the on-line processing of non-native segments changes over time as novel phonemic categories emerge. This exciting new avenue of research provides additional support for the model of speech perception outlined here and, in particular, for the claim that the organization of feature geometry in a speaker's mental grammar operates in the on-line processing of speech sounds to map the incoming acoustic stimuli onto discrete perceptual categories, giving rise to categorical perception and, thus, making segmentation of the speech stream possible.

Notes

1 Although the term "second language acquisition" (SLA) technically refers to the acquisition of a second language by either an adult or a child, it is typically used to denote acquisition by post-pubescent learners. In this chapter, we will only consider SLA by adults; however, the claims made here may be extended to L2 learners of all ages.

2 Note that the introduction of Optimality Theory (Prince and Smolensky, 1993; McCarthy and Prince, 1993) and the concomitant shift away from concern about the structure of representations, including the internal structure of phonemes, does not negate the insights captured by the theory of Feature Geometry assumed here; any theory must capture the fundamental dependency and constituency relations that exist between phonological features. As it is not the goal of this chapter to argue in favor of Feature Geometry over Optimality Theory, I will only direct the reader to some relevant papers on this issue (see Cole and Kisseberth, 1994; Padgett, 1994; Pulleyblank, 1997; see Brown, 1997: 291–317, for arguments that segment-internal structure must be maintained in underlying representations; these arguments take the form of a demonstration that the type of speech perception results discussed in this chapter cannot be captured easily, or possibly at all, in terms of constraint ranking).

3 I will assume, along with a growing number of researchers, that features are monovalent and that it is the mere presence of a feature in the representation of a segment that designates the active involvement of its corresponding articulator; likewise, the absence of a feature entails that the corresponding articulator is not active for a given segment (e.g., Anderson and Ewen, 1987; Avery and Rice, 1989; van der Hulst, 1989). For example, the voiceless segment /t/ will simply not contain the feature [voice] in its representation (that alone ensures that the vocal cords are not active for this segment), whereas the phoneme /d/ will be specified for the feature [voice].

4 The representations for /l/ and /r/ assumed here differ from standard representations. The phonological feature [lateral] is generally assumed to distinguish laterals from non-laterals, in this case /l/ from /r/. However, Brown (1993b, 1995) argues that [lateral] is not tenable as a phonological feature and that the contrast between /l/ and /r/ is best captured in terms of Place features (see reference for specific theoretical motivation and empirical evidence to support this claim; see also Piggott, 1993, and Spencer, 1984, for this view). Importantly, the representations for /l/ and /r/ given

in figure 1.2 provide an explanation for differential acquisition effects due to a speaker's first language grammar, which is not expected given the standard view of liquids.

5 The Japanese /ɾ/ will be distinguished from other coronal sounds (e.g., /s/) in terms of manner features.

6 This model integrates properties of models proposed by Clements and Hume (1994), Piggott (1992) and Rice and Avery (1991); however, the arguments and findings presented here do not hinge on the correctness of this particular hierarchical organization.

7 Perception research has tended to focus on very young infants (0–14 months) or older children (4–12 years); there is a surprising lack of perceptual data for young children (1–3 years). Thus, while the decline in sensitivity to non-native contrasts has been shown to begin in the first year of life, it has not yet been determined whether this early perceptual reorganization is rigid or remains relatively flexible until the phonological system is firmly in place. For example, it has not yet been determined whether the observed early pattern of perception persists throughout language development or whether a child, if placed in the appropriate language environment once perceptual reorganization has begun, would regain the original sensitivities.

8 See papers in Strange (1995) for reviews of the relevant speech perception data as well as several interesting proposals regarding the relationship between linguistic knowledge and the developing speech perception system.

9 There appears to be a time lag (approximately four to six months) between the age of the perceptual loss and the corresponding phonological development. A possible explanation for this time lag is that there is a confound between lexical development and phonological development, such that segmental representations are integrated into lexical items (which is what Brown and Matthews actually measured) shortly after they are first acquired. If this is indeed the case, acoustic discrimination tasks (specifically, lack of sensitivity) might provide a means of measuring the phonological development of children at even earlier ages than is currently available. This suggests that there may be an inventory of segments that is independent from the lexical items that contain them; this is a possibility that I will leave open for future research.

10 Note that this is true regardless of the actual phonetic realization of a particular contrast. Take, for example, voicing contrasts: although languages may vary as to how they choose to acoustically realize the voicing contrast (i.e., actual Voice Onset Time (VOT) may vary), since the same phonological feature underlies this contrast (i.e., [voice]), the claim is that speakers whose native language exploits this feature will be able to perceive all non-native voicing contrasts. A caveat is necessary here: this claim does not entail that speakers whose L1s contain the feature [voice] will necessarily perceive non-native voicing contrasts with 100 percent accuracy *initially*, nor that their performance will be equally accurate for that contrast in all positions within the syllable; only that there will be a qualitative difference between their perception of such a contrast and a contrast for which their L1 lacks the relevant feature (even initially), and that such a voicing contrast will be *acquirable* for such speakers (and that speakers will attain native-like performance given the appropriate input), even in syllabic positions not allowed in the L1 system.

11 The ability to construct novel segmental representations presumes, of course, that the acquisition device is still operative in L2 acquisition. See White (1989) for arguments regarding the operation of Universal Grammar in L2 acquisition; see Brown (1998) for a discussion of this issue with respect to L2 phonological acquisition.

12 The segments /s/ and /θ/ also differ acoustically in terms of stridency, and some phonologists distinguish them by the feature [strident]; however, following Kenstowicz

(1994: 30), I assume that their phonological representations differ in terms of place features. The predictions for the learners of English will not differ, though, under either analysis, as neither [strident] nor [distributed] is an underlying feature in any of the three languages under investigation.

13 While other segments are realized phonetically in Japanese, such as [Φ], they are derived (i.e., occurring in predictable phonological contexts) and do not, therefore, constitute independent phonemes.

14 The claim that Mandarin Chinese does not contrast /l/ and /r/ phonemically requires some comment. This language contains /l/ and a segment which is transcribed in romanized script as "r"; this transcription gives the impression that there is a contrast between the lateral approximant /l/ and a central approximant /r/. This "r" segment, however, is classified by linguists as a voiced retroflex fricative, /ʐ/. For this study, I follow Maddieson (1984) in treating /ʐ/ as a voiced retroflex fricative and, crucially, not as a retroflex sonorant. But compare Rice (1992) who analyzes this segment as /r/ underlyingly; postulating that it surfaces as a voiceless retroflex fricative [ʐ] in onset position and as rhoticization of the vowel when in coda position [ɚ]. Thus, according to this analysis, /r/ and /l/ do contrast as sonorants in Mandarin. However, it is not clear that this analysis is correct. The coda position in Mandarin is restricted to nasals; thus it is unlikely that the rhoticization of the vowel is from the presence of an approximant in the coda position. Finally, only certain vowels are rhoticized (Chao, 1968; Wu, 1991). This suggests that the rhoticization is a property of the vowel itself, rather than the result of /r/ in the coda position.

15 Japanese contains a bilabial fricative [Φ]; however, this is an allophone of /h/ and is realized before the high back unrounded vowel /ɯ/.

16 Japanese contains one liquid described as a flap [ɾ], which is not identical to the central approximant [r] in English, but is traditionally considered to correspond to English /r/, not /l/. This flap has several variants, which vary freely, one of which is phonetically similar to English [l] (Vance, 1987).

17 In Korean, [ɾ] and [l] are in complementary distribution, with [ɾ] (an apical flap) occurring intervocalically (Jung, 1962).

18 Although there are coronal segments in Japanese (e.g., /t/, /s/, /n/), under a theory of Minimally Contrastive Specification, a feature will only be present in a grammar if that feature is required to contrast segments; accordingly, coronal segments in Japanese will be represented with a bare Place node. However, based on palatal prosody in Japanese mimetics, Mester and Itô (1989) argue that Japanese coronal segments are, in fact, represented with the feature [Coronal]. This specification is proposed in order to account for the fact that all coronal segments *except* /ɾ/ are palatalized (as are non-coronals). By specifying all coronal sounds, other than /ɾ/, with the feature [Coronal], the authors explain why /ɾ/ is not a target of this operation. The same facts, however, can be obtained by assuming (as the authors themselves do to explain why /ɾ/ cannot be geminated) that /ɾ/ is not specified for any Place Node at all, whereas coronal segments are specified for a bare Place Node (with no [Coronal] feature). Lacking a Place Node, /ɾ/ will never be the target of palatalization. This specification would also explain why coronals are the preferred target of this operation, with non-coronals becoming palatalized only in the absence of a coronal: since coronals lack Place features, the palatal morpheme has a free place to dock on these segments, whereas the addition of palatalization to the non-coronal segments creates a less-favored complex structure.

19 While Japanese learners of English receive ample instruction in their language classes regarding the fact that /l/ and /r/ are contrastive in English, this type of explicit input,

due to its very nature, does not feed into the acquisition device and, thus, does not trigger acquisition (Schwartz, 1993).

20 As this study was originally reported in Brown (1993a, 1998) a summary of the methodology and statistical analyses will be given; the reader is referred to the original study for more details.

21 Brown (1993a, 1998) also examines acquisition of the /l–r/ contrast in onset clusters (e.g., glass/grass) and coda position (e.g., ball/bar).

22 Japanese speakers, for example, have been shown to correctly articulate /l/ and /r/, despite their inability to perceive a difference between these two sounds. This is possible since adult learners have a developed motor control system and are able (with practice) to execute the necessary articulations. Once a speaker knows the spelling of a word that contains /l/ or /r/, he or she can accurately produce the correct liquid, thus giving the appearance of having acquired the contrast. To my knowledge, no one has investigated how this knowledge of proper articulation might be encoded into the learner's lexical representation of words. It is not clear whether this knowledge (which is dependent on orthography) is represented in terms of phonological structure.

23 A training book was constructed which included every picture (one to a page) appearing in the experimental test. This book was used to familiarize the subjects with each of the pictures, and corresponding name, to appear in the Picture Selection task. This was done in order to minimize any errors that might be caused by the subjects' unfamiliarity with a particular stimulus item or illustration of an item. The materials used in stimuli preparation for the pictures were adapted from the Bilingual Aphasia Test (Paradis and Libben, 1987).

24 This result is perhaps even more surprising given the tendency of many Japanese learners to substitute /b/ for /v/ in production. But, as pointed out above, there is a well-known dissociation between comprehension and production skills, with comprehension assumed to be a more accurate reflection of the speaker's phonological knowledge.

25 For reasons of space and the goals of this chapter, I will not discuss subjects' individual performances, other than to point out that each of the Japanese subjects (not just the group as a whole) accurately discriminates the non-native /b–v/ and /f–v/ contrasts, but not the /l–r/ contrast. For a more detailed discussion of these individual data, the reader is referred to Brown (1998) in which these subjects' individual performances are analyzed in terms of a standard binomial distribution, showing that the group data are indeed representative of each subject. I will return to how the present model deals with individual differences, more generally, in the concluding section of this chapter.

26 According to Kawakami (1977: 32), the phoneme /b/ is realized as a plosive word-initially, but is often realized as a voiced bilabial fricative word-internally (compare [bareru] "be revealed" with [aβareru] "rampage").

27 Note that baseline performance on the AX Discrimination task is different from the picture task. According to the hypothesis that the Japanese grammar funnels the acoustic signal for both /l/ and /r/ into a single native phonemic category, Japanese speakers should perceive minimal pairs as identical. Thus, we would theoretically expect 0 percent accuracy at discriminating /l/ and /r/. In practice, though, they are able to correctly discriminate pairs more often – perhaps due to variations in duration and amplitude, which were not controlled for in this study. The difference in the subjects' performance on the two tasks (30 percent vs. 60 percent), then, is not indicative of differing abilities to perform each task, but rather reflects the fact that the baseline performance is different for each task.

28 An anonymous reviewer asks how this finding regarding accurate perception of a non-native voicing contrast squares with a previous finding by Jamieson and Morosan (1986) that speakers of Canadian French appear to have difficulty with the English θ/ð voicing contrast, despite the fact that the French grammar contains the feature [voice]. It would indeed be problematic for my proposal if French speakers did have trouble with the English θ/ð voicing contrast. However, a close examination of the Jamieson and Morosan study reveals several factors that undermine this conclusion. First, the stimuli in their study were presented with background cafeteria noise; it is likely that this noise depressed overall perceptual performance, so we cannot conclude that these results indicate that French speakers have any sort of absolute difficulty with the θ/ð contrast. Secondly, there were no native English speaker controls included in the design for comparison; this point is especially important considering that the stimuli were presented with cafeteria noise – we need to know how native speakers would do under these circumstances before we conclude that the French speakers cannot perceive θ/ð. Finally, the subjects were tested on a variety of stimuli spanning the voiceless–voiced continuum: while performance was low on the middle, more ambiguous stimuli (about 50 percent), it is quite good at either end of the continuum (about 80 percent), suggesting an ability to perceive the θ/ð contrast (again it would be useful to have native speaker data on the more ambiguous stimuli for comparison). Jamieson and Morosan's findings are in fact wholly compatible with the model of L1 interference developed here: they show significant improvement in between-category discrimination ability following training (e.g., θ/ð), but no improvement for within-category discrimination (e.g., two instances of θ along the voicing continuum), as well as striking improvement in subjects' ability to correctly identify each segment as voiced or voiceless (e.g., from 48 percent to 96 percent accurate identification). This kind of improvement is exactly what we would expect to find given that the French speakers' L1 grammar contains the feature [voice] – if their grammars lacked this feature the kind of improvement Jamieson and Morosan demonstrate would be impossible. Indeed, we will find a similar pattern of improvement below in experiment 3 where beginner and advanced learners' perception and acquisition of the English /b–v/ contrast are compared.

29 The differential performance of the two language groups also speaks to phonological theory, providing experimental evidence for the representations of /l/ and /r/ assumed in this chapter. According to Brown (1993b, 1995), /l/ and /r/ are distinguished by the presence of Coronal in the representation of /r/. We can interpret the differential performance of the two groups in terms of the presence of this feature in Chinese and the lack of it in Japanese. However, according to Rice and Avery (1991), /l/ and /r/ are differentiated not in terms of place features, but by manner features: /r/ contains a vocalic node whereas /l/ does not. Chinese and Japanese do not differ with respect to this feature, thus Rice and Avery's representations incorrectly predict that Chinese and Japanese speakers should perform similarly.

30 These conclusions are supported by Brown (1998), which compares the auditory and phonological discrimination abilities of Japanese and Chinese speakers living in North America, and Brown (1996), which compares Japanese and Chinese speakers living in Japan.

31 Depressed performance by both the High-level and the native control group on the /b–v/ contrast is likely caused by acoustic properties of this pair of sounds, especially in the environment of high front vowels, which minimize their distinctiveness.

32 Cochrane (1980) demonstrates that preadolescent Japanese children (ages 3–13 years) were no better than adults at perceiving /l–r/ minimal pairs. This finding indicates

that the inability of the Japanese adults to perceive this contrast is a result of a change that occurs very early in language development (i.e., acquisition of phonological structure) and not the result of a more general change that occurs sometime prior to puberty (e.g., lateralization of brain function).

33 In the "non-speech mode," all of the acoustic information that does NOT differentiate the two sounds – namely the first and second formants – was removed from the stimuli, resulting in something that sounds like a high-pitched glissando.

34 Loan words in Japanese are written in katakana, one of the two Japanese syllabary writing systems. When a foreign word containing the segment [v] is borrowed, this segment is traditionally transcribed as one of the kana for [ba], [bi], [be], [bo] or [bu] (i.e., [b] is substituted for [v]). However, within the last five years, a new kana symbol has been introduced by *Monbusho* (Japan's Ministry of Education) to represent the sound [v], ヴ, which is the symbol for the vowel [u], plus a voicing diacritic. Thus, just as it is possible for the learner, who originally maps all L2 sounds onto L1 categories (even those that do not match perfectly), to acquire a new perceptual category for those L2 sounds that do not correspond perfectly to the L1 categories, so too can writing systems be adapted to better represent the original pronunciations of loanwords. Words that were borrowed into Japanese before the introduction of this new symbol for [v] continue to be written as though they contained a [b] (e.g., "boɾiboɾu" for volleyball), but words that have been borrowed after the introduction of this symbol are written to accurately reflect the language of origin's pronunciation (e.g., "Bon Jovi" for Bon Jovi). It is quite interesting (and not accidental, I think) that a new katakana symbol has been introduced for [v] (an L2 phoneme which Japanese speakers have been shown above to accurately perceive), but not for [l] (an L2 phoneme which these speakers do not accurately perceive). That the writing system has been modified to accommodate [v], but not [l], reflects, I think, the increasing perceptual awareness of Japanese speakers that [v] does not adequately correspond to any native Japanese phonemes.

35 This position predicts that the time required to process English /b/ and /v/ by Japanese speakers, for example, will differ. Since English /b/ maps exactly to the Japanese category, it should be identified as /b/ more quickly than /v/ is identified (as /b/ or /v).

References

Abramson, A. and Lisker, L. (1970) Discriminability along the voicing continuum: Cross-language tests. *Proceedings of the Sixth International Congress of Phonetic Sciences*, pp. 563–7. Prague: Academia.

Anderson, J. and Ewen, C. (1987) *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Archangeli, D. (1988) Aspects of Underspecification Theory. *Phonology* 5(2): 183–207.

—— (1984) *Underspecification in Yawelmani Phonology and Morphology*. Doctoral dissertation. Massachusetts Institute of Technology. [Published 1988, Garland, New York.]

Avery, P. and Rice, K. (1989) Segment structure and coronal underspecification. *Phonology* 6(2): 179–200.

Barton, D. (1980) Phonemic perception in children. In G. Yeni-Komshian, J. Kavanagh and C. Ferguson (eds.), *Child Phonology*, vol. 2, pp. 97–116. New York: Academic Press.

Beddor, P. and Gottfried, T. (1995) Methodological issues in cross-language speech perception research with adults. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, pp. 207–32. Baltimore: York Press.

Best, C. (1994) The emergence of native-language phonological influence in infants: A perceptual assimilation model. In H. Nusbaum, J. Goodman and C. Howard (eds.), *The Transition from Speech Sounds to Spoken Words: The Development of Speech Perception*, pp. 167–224. Cambridge, MA: MIT Press.
—— (1993) Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies (ed.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, pp. 289–304. Dordrecht: Kluwer Academic Publishers.
—— and McRoberts, G. (1989) Phonological influences in infants' discrimination of two non-native speech contrasts. Paper presented at the Society for Research in Child Development, Kansas City, Kansas, April 1989.
Bley-Vroman, R. (1989) The logical problem of second language acquisition. In S. Gass and J. Schachter (eds.), *Linguistic Perspectives on Second Language Acquisition*, pp. 41–72. Cambridge: Cambridge University Press.
Borden, G., Gerber, A. and Milsark, G. (1983) Production and perception of the /r/–/l/ contrast in Korean adults learning English. *Language Learning* 33: 499–526.
Brière, E. (1966) An investigation of phonological interferences. *Language* 42: 768–95.
Brown, C. (1998) The role of the L1 grammar in the acquisition of L2 segmental structure. *Second Language Research*.
—— (1997) *The Acquisition of Segmental Structure: Consequences for Speech Perception and Second Language Acquisition*. Unpublished Ph.D. dissertation. McGill University, Montréal, Québec.
—— (1996) English tongues, Japanese ears: A theory of phonological interference. *Language and Culture: Journal of the Institute of Language and Culture Studies* 30: 165–203. Hokkaido University, Sapporo, Japan.
—— (1995) The feature geometry of lateral approximants and lateral fricatives. In van der Hulst, H. and van de Weijer, J. (eds.), *Leiden at Last*, pp. 41–88. University of Leiden, Holland.
—— (1993a) The role of the L1 grammar in the L2 acquisition of segmental structure. *McGill Working Papers in Linguistics* 9: 180–210.
—— (1993b) On the representation of laterality. *McGill Working Papers in Linguistics* 8(2): 109–42.
—— and Matthews, J. (1997) The role of feature geometry in the development of phonemic contrasts. In S. J. Hannahs and M. Young-Scholten (eds.), *Generative Studies in the Acquisition of Phonology*, pp. 67–112. Amsterdam: John Benjamins Publishing Company.
—— and —— (1993) The acquisition of segmental structure. *McGill Working Papers in Linguistics* 9: 46–76.
Burnham, D. (1986) Developmental loss of speech perception: exposure to and experience with a first language. *Applied Psycholinguistics* 7: 207–40.
Chao, Y. R. (1968) *A Grammar of Spoken Chinese*. Berkeley: University of California Press.
Chomsky, N. and Halle, M. (1968) *The Sound Pattern of English*. Cambridge, MA: MIT Press.
Clements, G. N. (1985) The geometry of phonological features. *Phonology Yearbook* 2: 223–50.
—— and Hume, E. (1994) The internal organization of speech sounds. In J. Goldsmith (ed.), *A Handbook in Phonological Theory*, pp. 245–306. Oxford: Blackwell Publishers.
Cochrane, R. (1980) The acquisition of /r/ and /l/ by Japanese children and adults learning English as a second language. *Journal of Multilingual and Multicultural Development* 1: 331–60.

Cole, J. and Kisseberth, C. (1994) An optimal domains theory of harmony. MS., University of Illinois at Urbana-Champaign.

Edwards, M. (1974) Perception and production in child phonology: The testing of four hypotheses. *Journal of Child Language* 2: 205–19.

Eilers, R., Gavin, W. and Oller, D. (1982) Cross-linguistic perception in infancy: Early effects of linguistic experience. *Journal of Child Language* 9: 289–302.

Eimas, P., Siqueland, E., Jusczyk, P. and Vigorito, J. (1971) Speech perception in infants. *Science* 171: 303–6.

Flege, J. (1995) Second language speech learning: Theory, findings, and problems. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, pp. 233–73. Baltimore: York Press.

—— (1992) Speech learning in a second language. In C. A. Ferguson, L. Menn and C. Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications*. pp. 565–604. Baltimore: York Press.

—— (1991) Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Heubner and C. Ferguson (eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theory*, pp. 249–90. Philadelphia: John Benjamins Publishing Company.

—— (1981) The phonological basis of foreign accent. *TESOL Quarterly* 15: 443–55.

——, Takagi, N. and Mann, V. (1995) Japanese adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech* 38: 25–55.

Garnica, O. (1973) The development of phonemic speech perception. In T. Moore (ed.), *Cognitive Development and the Acquisition of Language*, pp. 215–22. New York: Academic Press.

Goto, H. (1971) Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropsychologia* 9: 317–23.

Hancin-Bhatt, B. (1994a) *Phonological Transfer in Second Language Perception and Production*. Unpublished Ph.D. dissertation. University of Illinois at Urbana-Champaign.

—— (1994b) Segment transfer: A consequence of a dynamic system. *Second Language Research* 10(3): 241–69.

Hulst, H. van der (1989) Atoms of segmental structure: Components, gestures and dependency. *Phonology* 6: 253–84.

International Phonetic Association (Principles of the) (1979) University College, London: International Phonetic Association.

Jakobson, R. (1941) *Kindersprache, Aphasie und Allgemeine Lautgesätze*. Uppsala: Almqvist and Wiksell. Citations from the English translation, *Child Language, Aphasia and Phonological Universals*, translated by A. Keiler, 1968. The Hague: Mouton.

Jamieson, D. and Morosan, D. (1986) Training non-native speech contrasts in adults: Acquisition of the English /T/–/D/ contrast by francophones. *Perception and Psychophysics* 40: 205–15.

Jung, M.-W. (1962) *A Contrastive Study of English and Korean Segmental Phonemes with Some Suggestions Toward Pedagogical Application*. MS. dissertation, Georgetown University.

Kawakami, S. (1977) *An Outline of Japanese Sounds [Nihongo Onsei Gaisetsu]*. Tokyo: Ouhussha.

Kenstowicz, M. (1994) *Phonology in Generative Grammar*. Oxford: Blackwell Publishers.

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

Lehn, W. and Slager, W. (1959) A contrastive study of Egyptian Arabic and American English: The segmental phonemes. *Language Learning* 9: 25–33.

Maddieson, I. (1984) *Patterns of Sound*. Cambridge: Cambridge University Press.

Matthews, J. and Brown, C. (1998) Qualitative and quantitative differences in the discrimination of second language speech sounds. *Proceedings of the 22nd Annual Boston University Conference on Language Development*, pp. 499–510. Boston: Cascadilla Press.
Mattingly, I., Liberman, A., Syrdal, A. and Halwes, T. (1971) Discrimination in speech and non-speech modes. *Cognitive Psychology* 2: 131–57.
McCarthy, J. and Prince, A. (1993) Prosodic Morphology I: Constraint Interaction and Satisfaction. MS., University of Massachusetts, Amherst, and Rutgers University, New Brunswick, NJ.
Mehler, J. (1985) Language related dispositions in early infancy. In J. Mehler and R. Fox (eds.), *Neonate Cognition: Beyond the Blooming, Buzzing Confusion*. Hillside, NJ: Erlbaum.
Mester, A. and Itô, J. (1989) Feature predictability and underspecification: Palatal prosody in Japanese mimetics. *Language* 65(2): 258–93.
Michaels, D. (1974) Sound replacements and phonological systems. *Linguistics* 176: 69–81.
—— (1973) Sinhalese sound replacement and feature hierarchies. *Linguistics* 170: 14–22.
Miller, J. and Eimas, P. (1977) Studies on the perception of place and manner of articulation: A comparison of the labial-alveolar and nasal-stop distinctions. *Journal of the Acoustical Society of America* 61: 835–45.
Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J. and Fujimura, O. (1975) An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception & Psychophysics* 18: 331–40.
Munro, M., Flege, J. and MacKay, I. (1996) The effects of age of second language learning on the production of English vowels. *Applied Psycholinguistics* 17: 313–34.
Padgett, J. (1994) Feature classes. In J. Beckman, S. Urbanczyk and L. Walsh (eds.), *Papers in Optimality Theory*, University of Massachusetts Occasional Papers 18: 385–420.
Paradis, M. and Libben, G. (1987) *Assessment of Bilingual Aphasia*. Hillsdale, NJ: Lawrence Erlbaum.
Piggott, G. (1993) The geometry of sonorant features. MS., McGill University, Montreal, Canada.
—— (1992) Variability in Feature Dependency: The case of nasality. *Natural Language and Linguistic Theory* 10: 33–77.
Pisoni, D. (1973) Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics* 13: 253–60.
—— (1971) *On the Nature of Categorical Perception of Speech Sounds*. Unpublished Ph.D. dissertation: The University of Michigan.
Prince, A. and Smolensky, P. (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*, RuCCs Technical Report #2, Rutgers University Center for Cognitive Science, Piscaraway, NJ [in press, Cambridge, MA: MIT Press].
Pulleyblank, D. (1997) Optimality theory and features. In D. Archangeli and D. T. Langendoen (eds.), *Optimality Theory: An Overview*, pp. 59–101. Oxford: Blackwell Publishers.
Repp, B. (1984) Categorical perception: Issues, methods, findings. In N. J. Lass (ed.), *Speech and Language: Advances in Basic Research and Practice*, vol. 10. Orlando: Academic Press.
Rice, K. (1992) On deriving sonority: A structural account of sonority relationships. *Phonology* 9: 61–99.
—— and Avery, P. (1995) Variability in a deterministic model of language acquisition: A theory of segmental acquisition. In J. Archibald (ed.), *Phonological Acquisition and Phonological Theory*, pp. 23–42. Hillsdale, NJ: Lawrence Erlbaum Associates.

—— and —— (1991) On the relationship between laterality and coronality. In C. Paradis and J. F. Prunet (eds.), *Phonetics and Phonology: The Special Status of Coronals*, vol. 2, pp. 101–24. San Diego: Academic Press.

Ritchie, W. (1968) On the explanation of phonic interference. *Language Learning* 18: 183–97.

Sagey, E. (1986) *The Representation of Features and Relations in Non-linear Phonology*, Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Schwartz, B. (in press) On two hypotheses of "transfer" in L2A: Minimal trees and absolute L1 influence. In S. Flynn, G. Martohardjono and W. O'Neil (eds.), *The Generative Study of Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

—— (1993) On explicit and negative data effecting and affecting *competence* and *linguistic behavior*. *Studies in Second Language Acquisition* 15: 147–63.

—— and Eubank, L. (eds.) (1996) Special issue of *Second Language Research* 12(1).

Sheldon, A. and Strange, W. (1982) The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics* 3: 243–61.

Shvachkin, N. Kh. (1948) Razvitiye fonematicheskogo vospriyatiya rechi v rannem vozraste. *Izvestiya Akademii Pedagogicheskikh Nauk RSFSR* 13: 101–32. Citations from the English translation, The development of phonemic speech perception in early childhood. Translated by E. Dernbach and republished (1973) in C. Ferguson and D. Slobin (eds.), *Studies of Child Language Development*, pp. 91–127. New York: Holt, Rinehart and Winston.

Spencer, A. (1984) On eliminating the feature [lateral]. *Journal of Linguistics* 20: 23–43.

Stockwell, R. and Bowen, J. (1965) *The Sounds of English and Spanish*. Chicago: University of Chicago Press.

Strange, W. (ed.) (1995) *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Baltimore: York Press.

—— and Dittmann, S. (1984) Effects of discrimination training on the perception of /r–l/ by Japanese adults learning English. *Perception and Psychophysics* 36: 131–45.

Streeter, L. (1976) Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature* 259: 39–41.

Trehub, S. (1976) The discrimination of foreign speech contrasts by infants and adults. *Child Development* 47: 466–72.

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., and Best, C. (1994) Discrimination of English /r–l/ and /w–y/ by Japanese infants at 6–12 months: Language-specific developmental changes in speech perception abilities. Unpublished manuscript, Boston University.

Vance, T. (1987) *An Introduction to Japanese Phonology*. Albany, NY: State University of New York Press.

Werker, J. and LaLonde, C. (1988) Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology* 24: 672–83.

—— and Logan, J. (1985) Cross-language evidence for three factors in speech perception. *Perception and Psychophysics* 37: 35–44.

—— and Polka, L. (1993) Developmental change in speech perception: New challenges and new directions. *Journal of Phonetics* 21: 83–101.

—— and Tees, R. (1984a) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7: 49–63.

—— and —— (1984b) Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America* 75: 1866–78.

——, Gilbert, J., Humphrey, K. and Tees, R. (1981) Developmental aspects of cross-language speech perception. *Child Development* 52: 349–53.

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

—— (1988) Universal Grammar and language transfer. In J. Pankhurst, M. Sharwood Smith and P. Van Buren (eds.), *Learnability and Second Languages: A Book of Readings*, pp. 36–60. Dordrecht: Foris.

Williams, L. (1977) The perception of stop consonant voicing by Spanish–English bilinguals. *Perception and Psychophysics* 21: 289–97.

Wode, H. (1992) Categorical perception and segmental coding in the ontogeny of sound systems: A universal approach. In C. Ferguson, L. Menn and C. Stoel-Gammon (eds.), *Phonological Development: Models, Research, Implications*, pp. 605–31. Baltimore: York Press.

—— (1978) The beginnings of L2-phonological acquisition. *IRAL* 16: 109–24.

Wu, Y. (1991) Mandarin suffix [R] revisited. *Proceedings of the 1991 Annual Conference of the Canadian Linguistics Association*. University of Toronto.

Yamada, R. (1995) Age and acquisition of second language speech sounds: Perception of American English /ɹ/ and /l/ by native speakers of Japanese. In W. Strange (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-language Research*, pp. 305–20. Baltimore: York Press.

——, Kobayashi, N. and Tohkura, Y. (in press) Effect of word familiarity on non-native phoneme perception: Identification of English /r/, /l/ and /w/ by native speakers of Japanese. In A. James and J. Leather (eds.), *Second Language Speech*. New York: Mouton de Gruyter.