eight

Structural Factors in Second Language Phonology

8.1 Introduction

A foreign accent is created when there are phonological mismatches between the learner's native language (L1) and the target language (L2) that is acquired. People with different native languages have remarkably different productions in their pronunciations in a given foreign language. It is common to hear comments such as "Spanish speakers say it as ____, but Japanese speakers say it as ____", and so on. Such clear differences are not restricted to languages that are unrelated to one another such as Spanish and Japanese, but also are observable between speakers of languages that are closely related. For example, as will be clear later in this chapter, it is a rather simple task to differentiate between a speaker of Portuguese and a speaker of Spanish by their pronunciations of English. The reason for this is that the mismatches existing between Spanish (L1) and English (L2) are very different from the ones existing between Portuguese (L1) and English (L2), and result in different resolutions of the conflicts, which create different foreign accents.

Learners' renditions of English targets are governed in part by their native language sound patterns. The terms 'interference' or 'transfer' have long been used to designate the influence of the native language on the target patterns. Mismatches between the target and the native language may take different forms. One common situation is represented by the lack of the target sound in the native language. For example, the interdental fricatives of English, /θ/ and /ð/, are absent in many of the world's languages; these are usually substituted for by /s, z/ or /t, d/ respectively. Another frequently attested mismatch between L1 and L2 is created by under-differentiation of the phonemic distinctions of the target language. For example, as noted earlier, the English contrast between /tʃ/ and /t/ (e.g. <u>chip</u> – <u>tip</u>) is not patterned in the same way in Portuguese; rather, these two sounds are the allophones of one and the same phoneme, /t/. The Portuguese production of the phoneme /t/ is [tʃ] before /i/. Thus, it is only to be expected that speakers of Portuguese pronounce the target word <u>teacher</u> [titʃɚ] as [tʃitʃɚ] via a Portuguese filter.

Applied English Phonology, Second Edition Mehmet Yavaş
 ISBN: 978-1-444-33322-0

The foreign accent is not always due to a complete lack of the target phoneme, nor is it always because of the under-differentiation of target phonemic distinctions. Rather, the culprit is often the phonetic differences between identically defined targets and native sounds. For example, liquids present a good case for this. The phonetic quality of the non-lateral liquid of American English is very different than other r-sounds (taps, trills) found in a great many languages. Another such example is provided by the lateral liquids across languages. Differences between the so-called 'clear' and 'dark' laterals are easily observable, as shown by the cognate word animal in English [ænɪməl] (with a final 'dark l') and Spanish [animal] (with final 'clear l'). While the substitutions of these phonetically different sounds between the native and the target languages may not create a breakdown in communication by changing the word meaning (e.g. tip – chip), they do create a *very recognizable* foreign accent.

Mismatches in phonotactic (sequential) patterns also create significant problems. For example, while English allows up to triple onsets and triple codas, a language such as Japanese has no clusters. Such a mismatch between these two languages expectedly creates tremendous problems for Japanese speakers learning English. In addition, the number of onset or coda members is not the only problem; often the problem is created because of the type(s) of sound(s) and/or combinations demanded by the L2 not matching with what is allowed by the L1. For example, while double onsets are allowed both in English and Spanish, the variety of the combinations in English is much larger. Predictably, such a situation creates difficulties for the speakers of Spanish. Specific examples regarding the above points will be given in the following section on contrastive patterns.

Besides the segmental and sequential mismatches discussed above, there may be suprasegmental (prosodic) mismatches that make foreign accent obvious. The effects of rhythmic differences between the two languages considered, involving stress and intonation, are well known. It is also worth mentioning that the stress-timed versus syllable-timed nature of two languages produces noticeable non-native productions.

The observation of such clashes between L1 and L2 resulting in foreign accent created a huge industry of contrastive phonological studies in the 1950s and 1960s, which provided invaluable material for teachers and remediators.

In the following section, we will present a number of mini contrastive phonological structures with English as the target language and different languages as native languages, and point to the insights that can be gained from such analyses. It is important to stress the 'mini' character of these analyses, as each of these comparisons could be a book-length project that could be dealt with in a semester. Our aim here is simply to make the case in a thought-provoking manner and to stimulate the student and/or practitioner to make more detailed investigations.

The difference between contrastive phonologies and contrastive analyses in other domains (e.g. syntax) lies in a speaker's ability to communicate. While it is common to observe native language interference in syntax (e.g. "I have twenty-five years", instead of the native English "I am twenty-five years old", uttered by a Spanish or Portuguese speaker, is clearly a direct translation from L1), problems in several aspects of the syntactic domain may not be apparent all the time. For example, if a learner does not have sufficient knowledge of the differences in the uses of the 'simple past' and the 'past perfect', she or he can paraphrase things and get by with the use of the 'simple past' alone. To give another example, we can look at the modal verbs of 'obligation'. While English possesses a plethora of forms (e.g. must, have to, should, ought to) with certain nuances, several other languages deal with the corresponding situations with one or, at the most, two forms. Thus, when speakers of such languages learn English, they encounter a problem. A learner who does not master the nuances among multiple English forms (let us say that she or he has limited competence for ought to) can get by perfectly without using ought to once; nobody will stop and remind him or her that ought to was required in one of the utterances he or she made and that he or she therefore sounded non-native. When we look at the phonology of L2, however, we realize that such evasions are not possible. A learner who has a problem with the interdental fricatives of English cannot simply utilize a strategy of avoiding in his or her speech words containing /θ/ or /ð/. The frequency of /ð/ in grammatical morphemes such as the definite article, the, the demonstrative pronouns (e.g. this, that, etc.), the case forms of the personal pronouns (e.g. them), and some common adverbs (e.g. then, thus) is more than enough to create a disastrous situation.

All the above make a special case for contrastive phonology in that, unlike in other domains of language, in interlanguage phonology the learner is in an exposed state, with nowhere to hide his or her limitations. Thus, the mismatches that exist between the native and the target languages are very relevant for professionals who deal with remediation. Such factors are especially relevant when we deal with post-pubescent learners for whom the effects of foreign accent are much more obvious and more lasting. I will not go into details of the age factor in L2 phonology learning, but simply present a display (figure 8.1) from Scovel (1988), which reveals the differences between pre-pubescent and post-pubescent learners unambiguously.

8.2 Mini Contrastive Analyses

In this section we will look at some contrastive situations that exist between the target language (i.e. English) and ten different first languages. As stated earlier, these are not exhaustive descriptions but rather summary statements. To clarify the purpose of the section, our first example, Spanish–English, will be a little more detailed; the remaining examples from nine other languages will be presented in a briefer manner.

8.2.1 Spanish–English

We start our description by giving the phonemic inventory of the L1 consonants and vowels.

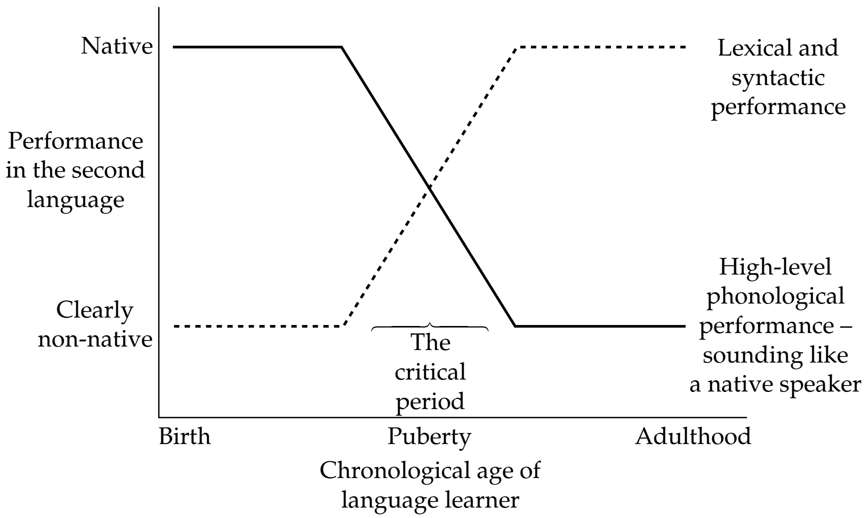


Figure 8.1 Contrasts in success between phonological learning and other linguistic skills, viewed chronologically
(*Source*: T. Scovel (1988) *A Time to Speak: A Psycholinguistic Inquiry into Critical Period for Human Speech*. Reproduced by permission of the author.)

Consonants of Spanish

	Bilabial	Labio-dental	Dental	Alveolar	Palatal	Velar	Glottal
Stop	p b		t d			k g	
Fricative		f		s		x	h
Affricate					tʃ		
Nasal	m			n	ɲ		
Liquid				l r ɾ	ʎ		
Glide					j	w	

Vowels of Spanish

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Before we go into the mismatches, we should mention some facts about Spanish. While the status of the vowels is rather consistent across varieties of Spanish, consonants show considerable variation. For example, /θ/, which is not included in the above table, is used only in dialects in Spain. Voiceless velar and glottal fricatives are encircled to indicate that either one or the other, not both, occurs in a given variety. Also noteworthy is the fact that the palatal lateral liquid /ʎ/, which is in contrast with the alveolar lateral /l/ in some varieties, is gradually being lost.

The inventory of the L1 (Spanish) given above is useful for depicting the target English phonemes that are missing. Accordingly, we can easily see that the targets /v, θ, ð, z, ʃ, ʒ, dʒ, ŋ/ will be problematic for learners, as Spanish does not have these phonemes (phonetically [ŋ] occurs in cinco, but in Spanish, unlike in English, it does not contrast with other phonemes). That these predictions are correct can be shown by the following frequently attested examples, where the missing targets are replaced by the closest sounds that are available in the native L1 inventory, resulting in several phonemic violations.

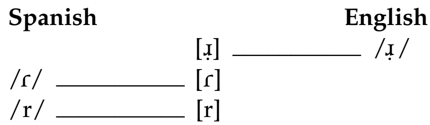
/θ/ → [t/s] (e.g. thin/tin → [tɪn], or [sɪn])
/ð/ → [d/z] (e.g. they/day → [de])
/v/ → [b] (e.g. vowel/bowel → [baʊl])
/z/ → [s] (e.g. zeal/seal → [sil])
/ʃ/ → [tʃ] (e.g. shop/chop → [tʃɑp])

It should be mentioned that one of the target English sounds above, [ð], is different from the others; this sound is phonetically present in both languages but has different phonemic mappings. As mentioned earlier (chapter 2), it is a separate phoneme in English and contrasts with /d/ (e.g. they vs. day); in Spanish, however, [d] and [ð] are allophones of the same phoneme.

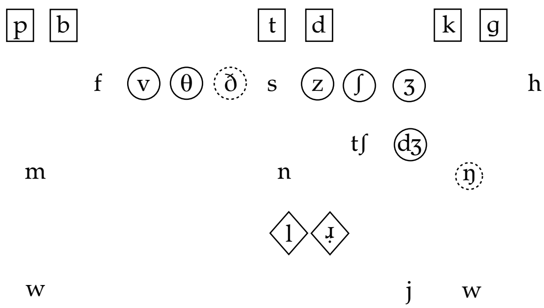
Although the inventory is capable of showing the above-mentioned problems, it is rather limited in scope, as different allophonic rules of identically described phonemes in two languages are also responsible for foreign accents. For example, despite the fact that the two languages in question have the same number of stop phonemes, these are far from being problem-free. Voiceless stops are always unaspirated in Spanish, whereas they are contextually (at the beginning of a stressed syllable) aspirated in English. Thus, in their production of English, Spanish speakers are expected to produce unaspirated stops in, for example, ton, pay, car. Also, voiced stops, /b, d, g/, of Spanish have fricative allophones, [β, ð, ɣ] respectively. Stop variants occur after pauses, after nasals, and after /l/; the fricative variants occur in other environments. Thus, Spanish speakers may produce fricatives for target voiced stops in adore, aboard, and so on.

Distributional restrictions are also the cause of problems in L2 phonology. Spanish has rather severe restrictions with respect to final consonants. Since the language allows only /s, n, r, l/ (and maybe /d/) to occur in final position, we might encounter several instances of final consonant deletion because English can demand that all consonants (except /h/) occur in this position. Similarly, since the only nasal that can occur finally is /n/ in Spanish, a target such as from with a bilabial nasal may be realized with a final [n] instead.

Another source of a foreign accent is salient phonetic dissimilarities in certain sounds between the two languages. This is nowhere more obvious than in a comparison of the liquids. While both Spanish and English have lateral and non-lateral liquids and can employ them in the same word positions, their clearly identifiable phonetic differences in the two languages produce easily detectable foreign accents. The alveolar lateral is always realized as 'clear l' (i.e. non-velarized) in Spanish, whereas the American English counterpart is produced mostly as shades of 'dark l' (i.e. velarized). The non-lateral liquids (i.e. r-sounds) of the two languages also exemplify considerable phonetic dissimilarity. The American English r is a retroflex approximant, while the two r-sounds of Spanish are a trill and a flap. Thus, we have the following mismatch:



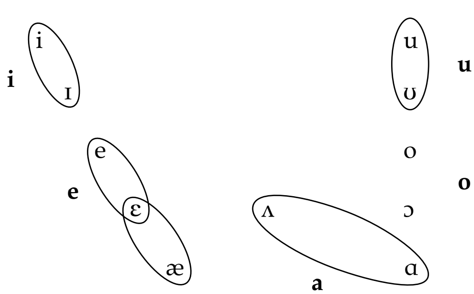
We can summarize the above in the following overlay of the L1 inventory onto the target English inventory (Spanish phonemes that have no relevance to the mismatches, such as /ʎ, x, ɲ/, are not considered here):



The following symbolizations are used throughout the comparisons between the target language (English) and the various first languages:

- ○ missing target phoneme in L1;
- ◌ sound existing only as an allophone of another phoneme in L1;
- □ different allophonic/distributional patterns in L1 and L2;
- ◇ salient phonetic difference between the target and the L1 counterpart.

The comparison of the vowel systems also makes certain problematic aspects rather obvious. Although there are no distributional problems in vowels (i.e. Spanish vowels can occur in all word positions), Spanish has a far smaller number (five) of vowels than English, and this proves to be an important and frequent source of insufficient separation (i.e. under-differentiation) of target phonemic distinctions. The frequently attested lack of contrasts (i.e. homophonies) that results includes /i/ – /ɪ/ (e.g. greed – grid), /u/ – /ʊ/ (e.g. fool – full), /ʌ/ – /ɑ/ (e.g. buddy – body), /ɛ/ – /æ/ (e.g. mess – mass). The following chart summarizes these potential confusions:



The following conventions are used throughout the comparisons with vowel systems of L1 and L2:

○	Circling of vowels indicates that these target contrasts are overlooked by learners coming from a specific L1.
i a ɛ	Bold-type vowel symbols indicate the expected native language vowels used in the rendition of targets.

It should be pointed out that the use of identical phonetic symbols for the bold-type L1 vowel does not imply that it is phonetically identical to any of the L2 (English) targets. For example, we use /i/ and /ɪ/ for the English high front vowels. Spanish speakers' rendition of /**i**/ does not mean that they are successful for English /i/ and unsuccessful for /ɪ/. Spanish substitution of /**i**/ is not identical to either English vowel. In almost all the languages we compare, the symbols /**i, e, o, u**/ indicate phonetically simple (not long and diphthongized) vowels. Similarly, the use of other symbols (e.g. /**ɛ**, **ɔ**/) does not make a claim that the phonetic qualities of these vowels are identical to those of English. The reader should keep these facts in mind when examining the vowel charts throughout.

The diphthongs are not expected to create problems for Spanish speakers as the language has a wide variety of diphthongs including all of those occurring in English.

Phonotactics (i.e. sequential patterning) is another aspect to consider in the comparison. In the present case, we see that Spanish and English are rather disparate:

Syllable structure

(L1) Spanish	(L2) English
(C) (C) V (C) (C*)	(C) (C) (C) V (C) (C) (C) (C**)

* possible only if syllable-final within word as stop/sonorant + /s/
** possible only if an affix

While English allows triple onsets and triple codas, the maximum number of consonants in Spanish in these positions is two. The number of consonants in clusters can tell only part of the whole story. The disparities are greater once we examine the other relevant dimension, namely the possible combinations. For example, English has a wide variety of double codas (see chapter 6), whereas Spanish has very limited combinations (stop/sonorant + /s/) only in word-internal position. There are differences for the double onsets too. The variety of combinations Spanish allows is limited to stop//f/ + liquid; any English target cluster other than these (there is a multiplicity of cases) can create significant trouble for Spanish speakers learning English.

Finally, mention should be made of the suprasegmental effects. Firstly, we can mention the stress-timed (English) versus syllable-timed (Spanish) difference. A rather obvious consequence of this difference is seen in rhythm because of the lack of vowel reductions, which are mandatory in English. Another aspect of the prosodic differences is related to different stress patterns. Such mismatches are especially dangerous in the case of cognates. Learners may (and indeed do) fall into the 'same/similar form and meaning' trap between the two languages. This is especially true when Spanish words have the stress on the final syllable, which English avoids. Here are some examples of such conflicts:

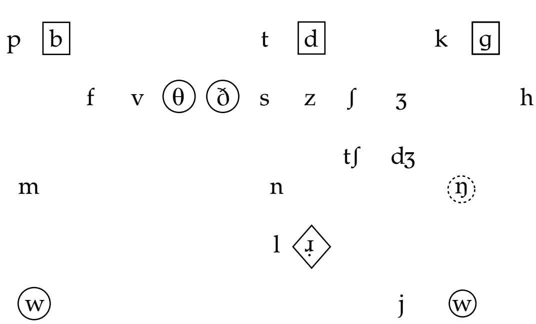
- **disyllabics:** ult in Spanish vs. penult in English: color, labor, honor, fatal, accion/action;
- **trisyllabics:** ult in Spanish vs. antepenult in English: animal, general, cultural, natural;
  ult in Spanish vs. penult in English: decision, informal, profesor/professor;
- **four syllables:** ult in Spanish vs. penult in English: artificial, horizontal, education/educacion;
  ult in Spanish vs. antepenult in English: particular, original, opinion (or on pre-antepenult in English because the antepenult has an [ə], which is unstressable: calculador/calculator, operador/operator, navegador/navigator).

The following summarizes the major trouble spots:

- entirely missing targets: /v/ → [b], /θ/ → [t], /ð/ → [d], /ʃ/ → [tʃ], /z/ → [s];
- distribution: only /s, n, l, ɾ/ occur finally in L1;
- aspiration of target /p, t, k/;
- fricative variants of L1 voiced stops intervocalically (e.g. adore → [aðɔɾ]);
- significant phonetic violations: liquids;
- consonant clusters;
- insufficient separation of several target vowel contrasts;
- stress;
- rhythm.

8.2.2 Turkish–English

The overlay of the native language consonantal system onto the target English inventory results in the following:

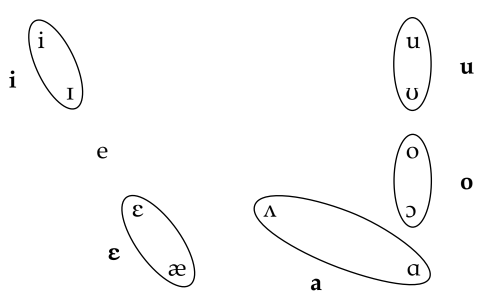


From the consonantal inventory, we can easily see potential troubles for the missing English targets /θ, ð, w, ŋ/, which manifest themselves in thin → [tin], they → [de]. Although there is also no /ŋ/ in the consonantal inventory of Turkish, [ŋ] is phonetically present in Turkish before velar stops, as in Ankara [aŋkaɾa], banka "bank" [baŋka]. Also, while [v] is used for the missing target /w/ (e.g. well → [vɛl]) this problem is not present intervocalically, as /v/ has a [w] allophone in this position.

Non-continuant obstruents (i.e. stops and affricates) have voicing contrasts in initial and medial positions; in final position, however, we find only the voiceless members of these pairs. This is the source of substitutions, for example, in bag → [bɛk], bid → [bit]. The progressive ending -ing [ɪŋ] creates a problem, which can be accounted for in two steps. Since [ŋ] in Turkish requires the presence of a following velar stop, and the velar stop in final position cannot be anything other than the voiceless variant, the rendition of -ing [ɪŋ] is [ɪŋk], as in going [goɪŋk].

Significant phonetic differences are relevant to liquids, especially for the non-lateral target retroflex approximant. The Turkish r-sound is an alveolar tap, /ɾ/. In addition, it is produced voiceless (and with friction) in final position (e.g. [kaɾ̥] "snow"). The alveolar lateral has both the 'clear' and the 'dark' variants, although their distributions are different from those of English and create mismatches. All word-initial laterals and all coda laterals after front vowels are 'clear' (cf. English 'dark' realizations in lawn and sell).

The conflicts in vowels involve several insufficient separations of contrasting English pairs /i/ – /ɪ/ (e.g. peach – pitch), /ɛ/ – /æ/ (e.g. mess – mass), /ʌ/ – /ɑ/ (e.g. buddy – body), /u/ – /ʊ/ (e.g. fool – full), which are summarized in the following chart:



Other Turkish vowels not relevant for mismatches are /y, ø, ɯ/.

The syllable structure of Turkish can be described as (C) V (C) (C). There are no initial clusters. The language does allow certain double codas, which can be described as "C1 = sonorant and C2 = obstruent, or C1 = fricative and C2 = stop". Because of great differences between these clusters and those of English, all target onset clusters, all triple codas, and several double codas expectedly create problems.

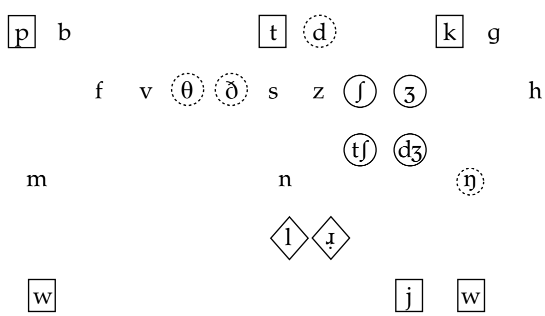
Coming from a syllable-timed language, Turkish speakers are expected to have difficulties with English vowel reductions and with rhythm. In addition, the stress patterns of the two languages are significantly different and prove to be sources of difficulty.

The following summarizes the major trouble spots:

- missing target phonemes: /θ/ → [t], /ð/ → [d], /w/ → [v] (except in V_V), /ŋ/→ [n] (except before a velar stop);
- final devoicing of non-continuant obstruents /b, d, g, dʒ/;
- significant phonetic differences: liquids, especially the non-lateral;
- under-differentiation of certain target vowel contrasts;
- onset and coda clusters;
- stress;
- rhythm.

8.2.3 Greek–English

The overlay of the L1 consonants onto the English targets results in the following picture:



Other Greek phonemes not relevant for mismatches are /x, ɣ/.

Starting with the targets missing in the L1 inventory, we note the lack of palato-alveolar fricatives /ʃ/ and /ʒ/; these tend to be replaced by the alveolar fricatives with their combinations with [j] as [sj] and [zj] respectively. Also lacking in Greek are the palato-alveolar affricates /tʃ/ and /dʒ/, which are replaced by the closest native alveolars, /ts/ and /dz/, respectively.

Although circled as a missing target phoneme, /ŋ/ is a little different from the others, because [ŋ] is an allophone of /n/ in Greek occurring before velar obstruents. Thus, problems are expected only in its occurrences in English with no adjacent velar stops.

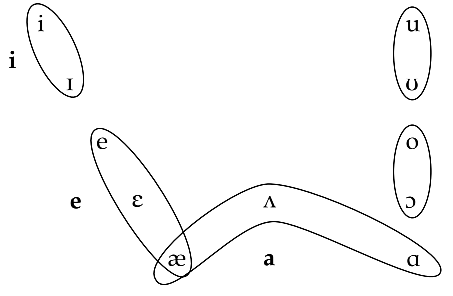
Voiceless stops in Greek are always unaspirated. Thus, problems are expected for the English targets in the beginning of stressed syllables.

As far as the salient phonetic differences are concerned, we need to highlight the liquids. The Greek alveolar lateral is always 'clear' and exemplifies a salient phonetic difference; the r-sound in Greek is also noticeably different, as it is an alveolar flap or trill. A minor difference can be cited between the /t/ and /d/ phonemes in the two languages; while these two are alveolars in English, they are dentals in Greek.

Positional/distributional restrictions are also sources of difficulty. All Greek consonants can occur initially and medially, and all except /d, θ, ð/ (among the relevant ones) occur finally. Thus, English targets with the above three in final position may cause problems.

Glides /w, j/ can create problems between the two languages, as Greek learners of English tend to hear and pronounce these glides as high vowels /u/ and /i/ respectively. While there is sufficient phonetic similarity between the glides and the corresponding high vowels, pronouncing them as vowels will give the impression to the English native speaker that there are separate syllables.

Vowel mismatches create the following insufficient separations for the target distinctions:



Greek has no diphthongs, but two vowel sounds can occur in sequence, and thus learners can handle the target English diphthongs.

Major problems with phonotactics are associated with the final clusters, which are non-existent in Greek. Thus, in addition to some simple codas pointed out earlier, Greek speakers will have problems with all the complex codas of English.

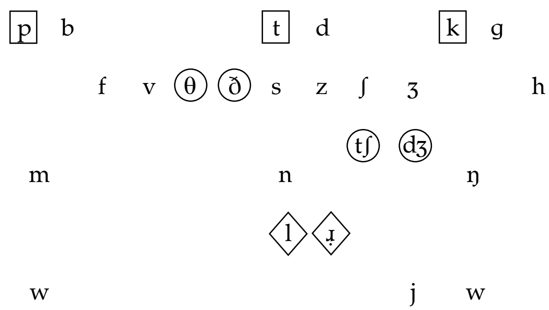
Greek is another syllable-timed language in our list, and expectedly has no vowel reduction. This results in considerable difficulties in learning the rhythm of English. In addition, different lexical stresses in the two languages are sources of problems.

The following summarizes the major trouble spots:

- missing target phonemes: /ʃ, ʒ, tʃ, dʒ/;
- aspiration;
- distributional restrictions: /θ, ð, d/;
- salient phonetic differences: approximants;
- insufficient separation of target vowels;
- stress;
- rhythm.

8.2.4 French–English

The overlay of the native phonemes onto the target English inventory gives us the following picture:



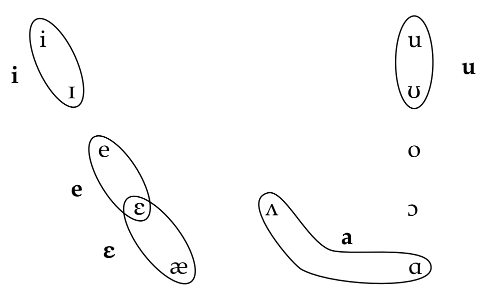
French consonants that are not relevant for the discussion are /ɲ, ɥ/.

Missing target phonemes include the interdentals /θ/ and /ð/ (which are rendered as [s] and [z] respectively, giving rise to mispronunciations such as think [sɪŋk], that [zæt]) and affricates /tʃ/ and /dʒ/ (which are rendered as [ʃ] and [ʒ] respectively). The status of /ŋ/ is different in that while it does not occur in native French words, one does find it in final position in borrowed words.

Allophonic differences may be observed in /p, t, k/ regarding aspiration. As in other Romance languages, voiceless stops are unaspirated in French, leading to mispronunciations of English targets.

Salient phonetic differences belong, once again, to the realm of liquids. The lateral in French is always ‘clear’, and the non-lateral is either an alveolar trill, /r/, or the uvular fricative/approximant, /ʁ/, and these are consistently used to substitute for the English liquid targets. The sounds /t, d/ present minor phonetic differences, as these are dental in French.

The mismatches and the under-differentiations regarding the target vowel contrasts are highlighted in the following diagram:



French vowels that are not relevant are /y, ø, œ/ and the nasal vowels /ɛ̃, ɑ̃, ɔ̃, œ̃/.

The syllable structure of French, which can be described as (C) (C) V (C) (C), allows a maximum of double onsets and codas. In addition, the combinations allowed by these double onsets (basically, C_1 = /f, v/ or stop, C_2 = liquid), and codas (basically C_1 = liquid, C_2 = stop) are more limited than those of English. Thus, some problems are expected in these mismatches.

Although French is classified as a syllable-timed language, it does not have the typical ‘staccato’ (or ‘machine-gun’) rhythm, and has reduced vowels.

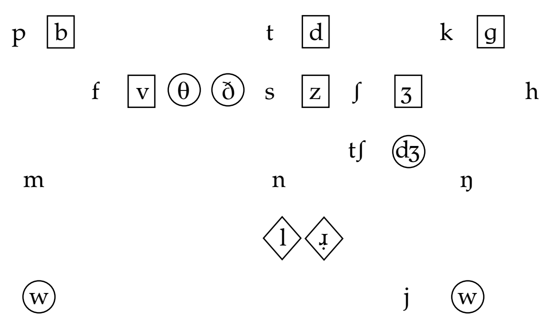
Despite this, the rhythm is quite different than that of English. In an English rhythm group, the first syllable is stressed and its pitch is higher than the other unstressed syllables. In French, on the other hand, the final syllable of each rhythmic group is lengthened and its pitch is leveled to half way before it is lowered. Thus, learners have considerable problems with English stress and rhythm.

The following summarizes the major trouble spots:

- missing target phonemes: /θ, ð, tʃ, dʒ, (ŋ)/;
- aspiration;
- salient phonetic differences: liquids;
- certain onset and coda clusters;
- insufficient separation of several target vowel contrasts;
- stress;
- rhythm.

8.2.5 German–English

The overlay of the native phonemes onto the target English inventory reveals the following:



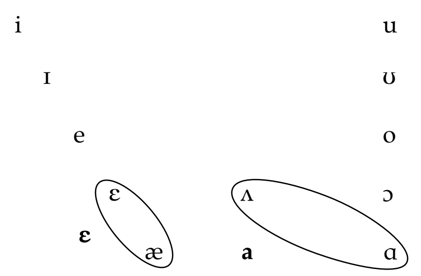
Other German consonants are /χ, ç, t^s, p^f/.

Missing targets include /θ, ð, dʒ, w/, which are commonly rendered as [s, z, tʃ, v] respectively.

Voiced obstruents /b, d, g, v, z, ʒ/, although shared by the two languages, do present problems in final position, as they are rendered voiceless in German.

Salient phonetic differences, once again, are related to the liquids. The German lateral is 'clear', and the r-sound is a uvular fricative. It is also worth mentioning that /ʁ/ is normally an approximant intervocalically; after voiceless obstruents it is voiceless (e.g. trat [tχat] "kicked"); post-vocalically before a consonant or word-finally, it is vocalized to [ɐ]. All these variations are sources of the problems learners face when dealing with the target English retroflex approximant /ɹ̣/. It may also be worth mentioning a slightly different phonetic realization of German /j/ in that it is produced with friction.

Vowel mismatches are depicted in the following chart:



Other German vowels are /y, ʏ, ø, œ, ɛ:, a:/.

The German tense vowels /i, e, o, u/ are longer but lack the diphthongal characteristics of the ones in English. This presents a slight phonetic mismatch.

German syllable structure, which can be described as (C) (C) (C) V (C) (C) (C), is as complex as that of English, although the specific combinations allowed may not be identical. Thus, any difficulty that may be observed will be due not to the number of consonants but rather to mismatches of the combinations of the types of sounds.

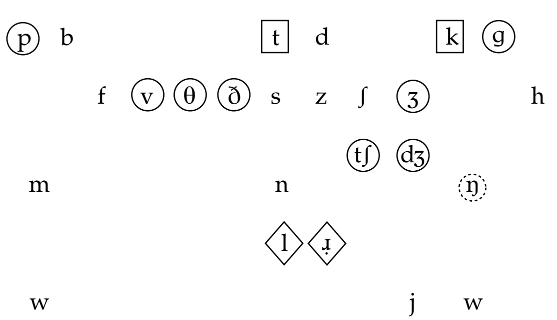
Being stress-timed languages, English and German share many characteristics in stress and rhythm. Thus, these areas are not expected to create problems for learners.

The following summarizes the major trouble spots:

- missing target phonemes: /θ, ð, dʒ, w/;
- distributional restrictions: voiced obstruents;
- salient phonetic differences: liquids;
- insufficient separation of target vowel distinctions.

8.2.6 Arabic–English

The overlay of the Arabic consonantal phonemes onto the target English inventory reveals the following:



Other Arabic consonants are /x, ɣ, ħ, ʕ/ and the pharyngealized (emphatic) consonants /tˤ, dˤ, sˤ, lˤ, ʕˤ/.

Missing target phonemes /p, g, v, θ, ð, ʒ, tʃ, (dʒ)/ are responsible for the following phonemic clashes:

/p/ → [b] pan – [ban]
/f/ → [v] fan – [van]
/θ/ → [s]/[t] thin – [sin] [tin]
/ð/ → [z]/[d] breathe – [breeze] [breed]
/tʃ/ → [ʃ] chin – [shin]

The occurrence of /θ/ and /ð/ in classical Arabic complicates the problem, giving the impression that the learner should not have problems with these targets in English, because she or he has been exposed to these sounds in the study of Arabic. This, however, does not translate into reality and learners have serious problems with respect to English interdentals.

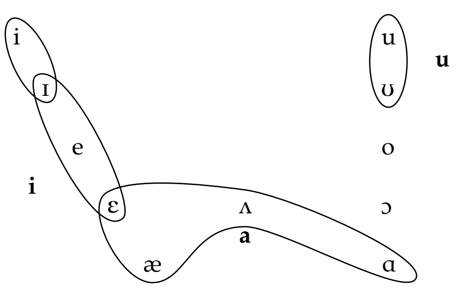
The sound /dʒ/, although present in some dialects of Arabic, was lost in Egyptian Arabic; also noteworthy is the questionable status of /ʒ/.

The case of /ŋ/ is similar to those of Turkish and Greek, in that this sound occurs as an allophone of /n/ before a velar stop, but cannot stand alone. Thus, while <u>finger</u> [fɪŋgɚ] may not be problematic, because [ŋ] is followed by a velar stop, <u>sing</u> [sɪŋ] and <u>singer</u> [sɪŋɚ] will be (i.e. the expected productions are [sɪŋg] and [sɪŋgɚ]).

The two voiceless stops of Arabic /t, k/ are unaspirated and are expected to be problematic.

Salient phonetic differences are related to liquids once again. The Arabic lateral is 'clear', and the r-sound is an alveolar apical trill. In addition, both liquids of Arabic have voiceless allophones pre-pausally following voiceless obstruents. All these result in obvious foreign accents in their English productions. Slight phonetic differences are observed in /t, d/ because they are dental in Arabic.

Mismatches and the resulting insufficient separation of English vowel contrasts are depicted in the following:



Arabic syllable structure, (C) V (C) (C), clashes considerably with that of English. Having no onset clusters and allowing only very limited double codas result in an epenthetic vowel to break up complex English targets.

Although Arabic is a stress-timed language, vowel reductions do not follow English patterns, and this results in some differences in rhythm.

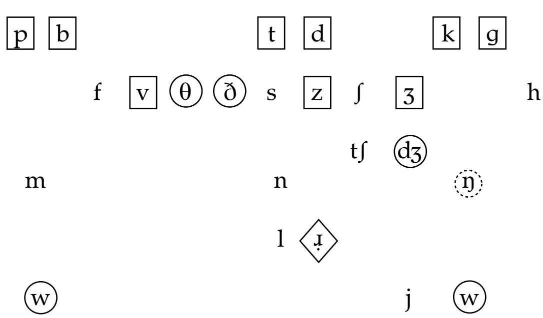
Word stress is fairly regular in Arabic; it falls on the final heavy syllable (one with either a long vowel or a VCC rhyme) of a morpheme. This is responsible for the commonly observed errors (stress on the final syllable as opposed to the native English pattern of initial stress) in difficult, expert, narrowest, institute, where the first three words have VCC rhymes, and the last word has a long vowel in the final syllable.

The following is a summary of the major trouble spots:

- missing target phonemes: /p, g, v, θ, ð, ʒ, tʃ, (dʒ), (ŋ)/;
- aspiration;
- salient phonetic differences: liquids;
- insufficient separation of several target vowel contrasts;
- onset and coda clusters;
- stress;
- rhythm.

8.2.7 Russian–English

The overlay of the native language phonemes onto the target English inventory reveals the following:



Other Russian consonants are /p^j, b^j, t^j, d^j, k^j, g^j, f^j, v^j, s^j, z^j, x, x^j, ts, m^j, n^j, r^j, l^j/.

Missing target phonemes in L1 include /θ, ð, dʒ, ŋ, w/, which have the substitutes [t, d, tʃ, ŋ, v], respectively.

Notable distributional and/or allophonic mismatches concern the following:

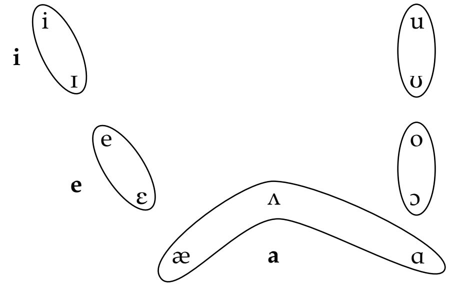
- All voiced obstruents are devoiced in final position, resulting in commonly observed homophonies neutralizing the target contrasts, such as bag – back, cab – cap, bed – bet, save – safe, in favor of the voiceless member.
- Voiceless stops, /p, t, k/, unlike in English, are unaspirated and provide another source for observable foreign accent.
- While the lateral liquid is quite similar to that of English (i.e. 'dark'), we have a different situation with the non-lateral. The Russian r-sound is an alveolar trill and this gives rise to a distinct foreign accent. There are some cases that provide minor phonetic differences. Among these are /t, d, n/,

which are dental in Russian, and /tʃ/, which is slightly more palatalized in Russian.

As the list of Russian phonemes below the diagram demonstrates, Russian has several palatalized consonants, and learners may use the palatalized sound when English targets occur in environments conducive to palatalization, such as before a high front vowel or /j/.

Similar to the situation mentioned for Turkish, /ŋ/ targets in final position undergo a two-step process. First is the insertion of the velar support, /g/, and then the subsequent devoicing of it to [k], yielding productions such as going [goɪŋk].

The clashes in the vowel systems of L1 and L2 result in the following under-differentiations of the target distinctions:



The limited five-vowel system of Russian is reduced to three, [i, a, ə] in unstressed syllables. Although both English and Russian are stress-timed languages, vowel reductions work differently; in Russian, [ə] never occurs immediately before the stressed vowel, and this results in non-reduction in many pretonic syllables of English target words. Also, Russian words contain only one stress; thus learners will tend to stress only the syllable with the tonic accent.

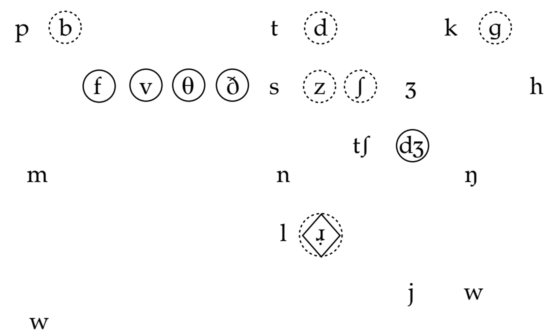
The syllable structure of Russian, which can be described as (C) (C) (C) V (C) (C) (C), is comparable in its complexity to that of English, and thus, this area is not expected to be problematic for the learners.

The following summarizes the major trouble spots:

- missing target phonemes: /θ, ð, dʒ, ŋ, w/;
- aspiration;
- final devoicing of the obstruents;
- salient phonetic differences: non-lateral liquid;
- insufficient separation of target vowel contrasts;
- stress;
- rhythm.

8.2.8 Korean–English

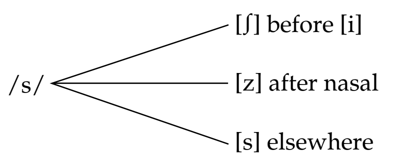
The overlay of the L1 consonant phonemes onto the target English inventory results in the following:



Other Korean phonemes are /pʰ, p′, tʰ, t′, kʰ, k′, tʃʰ, tʃ′, s′/.

Target phonemes that are completely missing in L1 include /f, v, θ, ð, dʒ/, and they are rendered as [p, b, t, d, tʃ] respectively in target English words. Although /b, d, g/ are not in the Korean phonemic inventory, [b, d, g] are present as allophones of /p, t, k/ between two voiced sounds. As a result, we expect difficulties in English /b, d, g/ targets when they are not in between two voiced sounds (e.g. book, cab, dog). Equally problematic are the /p, t, k/ targets when between two voiced sounds, as exemplified in the following erroneous productions: apart [əbɑɹ̥t], attack [ədæk], mocha [mogə].

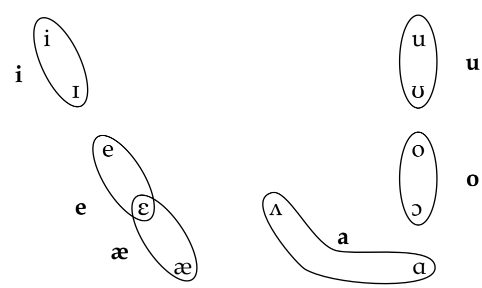
As mentioned in chapter 2, [z] and [ʃ] exist in Korean as allophones of /s/. We repeat the distributional requirements here for convenience:



Consequently, we expect the target sea shells [si ʃɛlz] to be rendered as [ʃi sɛls].

Liquids present both phonemic and phonetic problems. The Korean r-sound is a flap and is in complementary distribution with the lateral; [ɾ] occurs intervocalically and [l] elsewhere, thus giving rise to failures to distinguish between target pairs such as feeling – fearing, soul – sore.

The mismatches between the vowel systems of L1 and L2 result in the following under-differentiations:



Other Korean vowels are /y, ø, ɯ/.

The syllable structure of Korean, which is described as (C) V (C) (C), is much simpler than that of English. Although the above formulation allows double codas, the actual combinations are very limited. As a result, the wide variety of double and triple onsets and codas in target English words are broken up by vowel insertions.

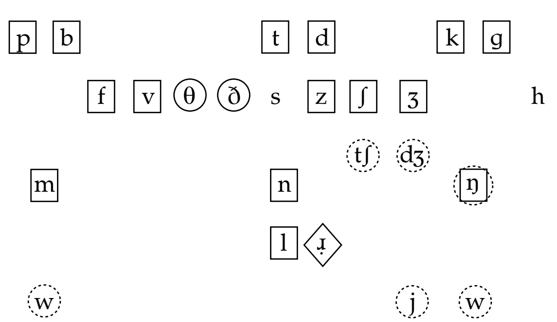
Korean stress patterns are quite different from those of English, mainly manifested as the rise in pitch on the initial syllable of the word or phrase. In addition to the mismatches in stress, Korean, as a typical syllable-timed language, does not have vowel reductions and this results in a clearly different rhythm.

The following summarizes the major trouble spots:

- missing target phonemes: /f, v, θ, ð, dʒ/;
- sounds existing as allophones;
- salient phonetic differences;
- insufficient separation of target vowel contrasts;
- onset and coda clusters;
- stress;
- rhythm.

8.2.9 Portuguese–English

The overlay of the L1 consonant phonemes onto the target English inventory results in the following:



Other Portuguese phonemes are /ʎ, ɲ/.

Missing target phonemes in L1 include /θ, ð, tʃ, dʒ, ŋ/. Of these, the first two have the substitutes [t, d] respectively. The sounds [tʃ] and [dʒ] exist in Portuguese as the allophones of /t/ and /d/ respectively before /i/. Thus, we can expect problems when the targets /t/ and /d/ occur before /i/, as in <u>teacher</u>, <u>difficult</u>, where the common renditions are [tʃ] and [dʒ], respectively, for the initial sounds. The sound [ŋ] is phonetically present before a velar stop.

As in other Romance languages, Portuguese voiceless stops are unaspirated and create problems for learners in dealing with English aspirated targets. In addition, no obstruent of Portuguese, except /s/, can occur in syllable-/word-final position. Consequently, English words with such demands receive an epenthetic vowel.

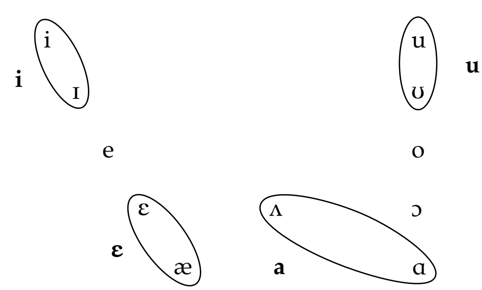
Nasals do not occur in final position either. The result is the nasalization of the previous vowel in English targets (e.g. from [fɾã]).

Liquids present both phonetic and distributional challenges. The lateral, /l/, is phonetically not very different from that of English (i.e. it is 'dark'); in syllable-final position, however, it is very much vocalized and becomes a [w] (e.g. Brazil [braziw]). The target word from a brand name of an analgesic Advil puts together three pattern clashes between L1 and L2. The typical rendition of this word as [adʒiviw] is easily explainable: since /d/ is not allowed in syllable-final position, an epenthetic vowel [i] is inserted; now that /d/ is followed by an [i] it turns into the appropriate allophone [dʒ]; the final [w] is accounted for by the above-mentioned allophonic rule of the lateral.

The two r-sounds of Portuguese, alveolar tap [ɾ] and velar/uvular fricative [x/χ], are phonetically very different than that of English /ɹ̣/. Substitutions of English targets vary depending on the word position dictated by L1 (i.e. [x/χ] in initial position, [ɾ] otherwise).

Glides /w, j/ create problems similar to those we observed in Greek–English mismatches; they are produced as high vowels /u/ and /i/ respectively, and give the impression of separate syllables.

The vowel mismatches between the two languages are shown below.



The syllable structure of Portuguese, which can be described as (C) (C) V (C), can match English demands in a limited fashion. A final single coda is possible only if the consonant is /s/ or a liquid. Also, the double onsets can only have the following structure: C_1 = stop or /f/, C_2 = liquid. Any other English target onset predictably suffers a modification.

Portuguese stress tends to go on the penult; thus anything different demanded by English may prove difficult for learners.

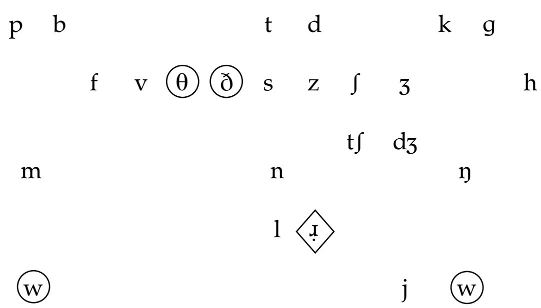
Although Portuguese leans more toward the 'stress-timed' pattern (Brazilian Portuguese less than European Portuguese), it does not have the same vowel reductions as those of English. This, coupled with the different lexical stress, results in difficulties in target rhythmic patterns.

The following summarizes the major trouble spots:

- missing target phonemes /θ, ð, tʃ, dʒ, ŋ/;
- different allophonic/distributional restrictions;
- aspiration;
- salient phonetic differences: non-lateral liquid;
- insufficient separation of target vowel contrasts;
- onset/coda clusters;
- stress;
- rhythm.

8.2.10 Persian (Farsi)–English

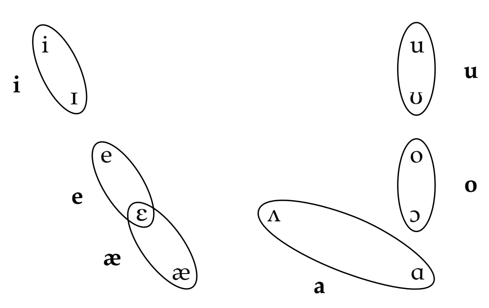
The overlay of the L1 consonant phonemes onto the target English inventory results in the following:



Missing target phonemes in L1 include /θ, ð/, which are substituted for by [t, d] respectively. Persian also lacks /w/; although several manuals suggest the rendition as [v], it actually is a frictionless approximant [ʋ].

The r-sound presents a salient phonetic difference, as it is an alveolar trill /r/ in Persian, with its allophones of a voiceless trill [r] in final position, and the tap [ɾ] intervocalically. The result is a clear foreign-accented English target /ɹ̣/.

Vowel mismatches creating under-differentiations are shown in the following:



As in many other languages, the Persian vowels replacing the targets do not have the English distinctions of tense/lax; rather, the quality of the vowels is in between.

The syllable structure of Persian, which can be described as (C) V (C) (C), is responsible for the difficulties experienced with the target double and triple onsets of English. Epenthetic (prothetic in the case of s-clusters) vowels are used to break up the impermissible clusters. Triple codas are problematic, as they do not exist in L1. Also, although Persian allows double codas, the combinations are more limited than those demanded by English; thus learners may experience difficulties with certain targets.

Since Persian stress is generally on the ult, there is considerable difficulty with English stress patterns. Combined with the difficulties in lexical stress, the syllable-timed characteristic of Persian, which does not allow any vowel reduction, may lead to a very different rhythmic pattern than that of English.

The following summarizes the major trouble spots:

- missing target phonemes: /θ, ð, w/;
- salient phonetic differences: r-sounds;
- insufficient separation of target vowel contrasts;
- onset and coda clusters;
- stress;
- rhythm.

The comparisons between English as L2 and several languages as L1 we have looked at repeatedly highlighted certain problematic areas for learners. Table 8.1 summarizes these important targets that create phonemic as well as some significant phonetic clashes (the 15 languages include the 10 we looked at and another 5).

8.3 Differential Treatment of Mismatches

In the previous sections we observed, besides many phonetic mismatches, several examples of phonemic mismatches between a learner's L1 and L2. Although the difficulties resulting from these mismatches are real, there seem to be differences in quality among them, and consequently, degrees of difficulty created by different types of mismatches.

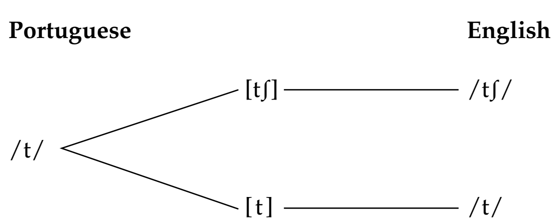
One type of phonemic mismatch between two systems was a result of a situation in which the two sounds that were in contrast in L2 were non-existent in L1. This was exemplified by the /θ/ – /ð/ contrast of English (e.g. <u>ether</u> [iθɚ] vs. <u>either</u> [iðɚ]). As we saw above, many languages, including Arabic, French, German, Korean, Turkish, Persian, Portuguese, and Russian, lack these completely, and the likely substitutions created violations of target contrasts.

The second mismatch that resulted in phonemic violations occurred when two sounds that were in contrast in L2 were present as the allophones of a single phoneme in L1. As mentioned earlier, the English contrast between /t/ and /tʃ/ (e.g. <u>tip</u> [tɪp] vs. <u>chip</u> [tʃɪp]) is under-differentiated by learners whose

Table 8.1 Significant phonemic and phonetic conflicts between English and several other languages

	Arabic	French	German	Greek	Hindi	Italian	Japanese	Korean	Mandarin	Persian	Russian	Portuguese	Spanish	Turkish	Vietnamese
θ ð → t d / s z	✔	✔	✔		✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
v → b					✔	✔	✔						✔		
w vs. v			✔		✔				✔	✔	✔			✔	
Onset/coda CC	✔			✔	✔	✔	✔	✔	✔	✔		✔	✔	✔	✔
Fin. C. devoic.			✔	✔				✔	✔		✔			✔	✔
i vs. ɪ	✔	✔		✔		✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
u vs. ʊ	✔	✔		✔		✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
ɛ vs. æ	✔	✔	✔		✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
ʌ vs. ɑ	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Aspiration	✔	✔		✔		✔					✔	✔	✔		
ɹ	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Stress	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔
Rhythm	✔	✔		✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔

L1 is Portuguese, because the two sounds are allophones of the same phoneme in their L1, as shown in the following:



The first situation involves acquiring new phoneme(s), and the second type is a creation of a phonemic split from an existing allophonic variation in the native language. While one may be inclined to think that acquiring new phoneme(s) will be more difficult than rearranging the two existing sounds from allophones of the same phoneme to separate phonemes, research has proven otherwise. It has been shown that learning becomes more difficult when the structures or sounds are similar in L1 and L2 than when they are dissimilar (Oller and Ziahosseyni 1970; Flege 1987, 1990; Major and Kim 1999).

8.3.1 Basic vs. derived context

The level of difficulty of going to a phonemic split for L2 from an existing allophonic variation in L1 has a correlation with various contexts. In a thorough examination of this issue, Eckman et al. (2003) state that whenever there is a conflict between L1 and L2 in the above manner (i.e. two sounds are in contrast in L2 but are the allophones of a single phoneme in L1), the target language phonemic contrast will be acquired first in basic (tautomorphemic) contexts, then in derived (heteromorphemic) contexts. Going back to the Portuguese–English mismatch regarding the /t/ – /tʃ/ contrast of English, the expected rendition of target English words tea and betting in the earlier stages of learning will be [tʃi] and [bɛtʃɪŋ] respectively, which fails in regard to the target contrast. Eckman et al. describe this as "stage I" ("no contrast"), where the native language allophonic rule applies in both the basic context (i.e. tautomorphemic situation where the affected sound /t/ becoming [tʃ], and the relevant environment – following vowel /i/ – are in the same morpheme), and in a derived context (in heteromorphemic context). Eckman et al. predict the next stage ("stage II") will show a partial contrast in the way the native rule applies only in the derived (heteromorphemic) context (i.e. the sound affected, /t/ becoming [tʃ] before high vowels taking place in betting realized as [bɛtʃɪŋ], while tea is realized as [ti] and not as *[tʃi]). The last stage ("stage III") is the one where the target contrast is acquired in both the basic and the derived context (i.e. in both tautomorphemic and heteromorphemic contexts). Thus, the following implicational relationship holds: *if a target pattern is acquired in a heteromorphemic context, it implies that the same is acquired in a tautomorphemic context, but not vice versa.*

This excludes a situation where the learner is successful in a derived context (e.g. betting pronounced as [bɛtɪŋ]) but will fail in the basic context (e.g. tea pronounced as [tʃi]). This hypothesis receives support from studies where the acquisition patterns reflect such an order (i.e. learning is earlier in basic contexts than in derived contexts).

8.3.2 Deflected contrast

Eckman et al. point out some situations where some phonemic mismatches between L1 and L2 result in an intersection of two interlanguage substitutions, and that one of these substitutions is systematically blocked. The rendition of English interdentals /θ, ð/ by Portuguese speakers provides a good case for this. The typical substitution for the /θ/ target is [t] by the learners (e.g. thank realized as [tænk]). As we saw earlier, Portuguese also under-differentiates the English /t/ – /tʃ/ contrast. Since [t] and [tʃ] are the allophones of a single phoneme in Portuguese, learners pronounce the target tip and chip homophonously. While the learners realize the English target /t/ as [tʃ] before a high front vowel, they do not reveal the same tendency when the target word has /θ/ before a high front vowel. Thus, a word like think [θɪŋk] is not expected to be rendered as [tʃɪŋk], but rather as [tɪŋk]. In other words learners distinguish the fate of two different [t] sounds. While the native allophonic rule converts the /t/ into [tʃ] before high front vowels, the [t] sound that is the substitute for the target /θ/ does not follow the same path. In this way, learners distinguish the three target language phonemes /θ/, /t/, and /tʃ/, and prevent the neutralization of any contrast. Eckman et al. state that their studies with Korean and Japanese speakers also confirm this tendency by maintaining the target contrasts.

8.3.3 Hypercontrast

Language learners may also be found to have difficulties with a newly acquired contrast and substitute the wrong member of the phonemic pair. Eckman et al. call this phenomenon 'hypercontrast' and state that it results from overgeneralization or hypercorrection. It is suggested that hypercontrasts are motivated by speakers' awareness of past errors they have made via L1 interference. For example, Spanish speakers have difficulties in acquiring the English /d/ vs. /ð/ contrast, as they are the allophones of a single phoneme in their L1. Once they acquire the contrast, however, they may produce incorrect [d] for correct [ð] intervocalically. Another example would be the following: a newly learned item with a /d/ target in initial position, which is in accordance with the L1 pattern, may be produced incorrectly as [ð].

8.4 Markedness

The different types of phonemic mismatches discussed above may be helpful in sorting out different degrees of difficulty that learners experience in the acquisition of L2 phonology. They are, however, far from depicting the whole picture. The reason for this is the varying nature of structural elements with respect to their markedness. *Markedness* of a structure is derived from its common occurrence in languages. Simply stated, a structure (constraint) A is more marked than another structure B if cross-linguistically the presence of A in a language implies the presence of B, but not vice versa (Eckman 1977, 1985; Eckman and Iverson 1994). Accordingly, two structures A and B, of which the first is more marked than the second, will present different degrees of difficulty for L2 learners. The classic example frequently discussed consists of the following two identically characterized situations provided by the mismatches of (a) German–English and (b) English–French, with respect to voiced–voiceless contrast in obstruents. The voiced and voiceless stop series /b, d, g/ and /p, t, k/ are part of the inventory of both English and German. While both languages contrast the voiced and voiceless series in word-initial and word-medial positions, the final position contrast is available only in English (e.g. back – bag); German neutralizes the contrast in favor of the voiceless member, and does not allow the voiced member in this position. The mismatch created in this position can easily predict the difficulty that German speakers have in learning English final voiced stops, with commonly observed substitutions such as cab [kæb] → [kæp], bed [bɛd] → [bɛt], and so on.

The second situation that can be described identically is a contrast existing in all word positions in L2 but neutralized in one of the word positions in L1. For this, we will consider the /ʃ/ vs. /ʒ/ contrast in English and in French. While both languages contrast the two sounds in medial and final positions, the initial contrast is available only in French. The prediction from this discrepancy is that speakers of English learning French will have difficulties for the above-mentioned contrast in word-initial position similar to that of German speakers' difficulties for the final voiced stops of English.

	/ʒ/				Voiced stops		
	Init.	Med.	Final		Init.	Med.	Final
L1 (English)	–	+	+	L1 (German)	+	+	–
L2 (French)	+	+	+	L2 (English)	+	+	+

Both cases reveal descriptively identical situations in that L2 has no restrictions of occurrence of the target in any word positions, while L1 has a positional restriction (i.e. English does not have /ʒ/ in initial position, and German does not have voiced stops in final position). Professionals who have observed these two identically describable mismatches would quickly point out that the difficulties experienced in these two situations are very different, and acquisition of the English final voiced stops by German speakers is a much greater challenge than acquisition of French initial /ʒ/ by speakers of English. Although both situations described deal with the voicing contrast in obstruents (/ʃ/ – /ʒ/ in fricatives, /p, t, k/ – /b, d, g/ in stops), acquiring the voicing contrast in final position is a more marked phenomenon than doing the same in initial position. Cross-linguistically, voicing contrast in final position implies the contrast in initial position, but the reverse is not known to be true. Accordingly, the difficulty of acquiring the voiced stops is a result of the more marked nature of voicing contrast in final position. Thus, while simple contrastive analysis can make predictions on the basis of the mismatches between L1 and L2, it cannot go beyond that. It is only by referring to the relative markedness of the structures that we can account for the variable performance of learners for seemingly identical situations.

Digging further into the markedness relations, we can discover other factors that are relevant for remediation. For example, it has been observed that learners have greater difficulty in acquiring the voicing contrast with velars (i.e. /k/ vs. /g/) than with alveolars (i.e. /t/ vs. /d/); bilabials are the least difficult. That is, the tendency to neutralize the contrast by devoicing is greater as the place of articulation moves further back. There is an aerodynamic explanation for such differences based on the place of articulation. The larger the supraglottal area for a stop, the better it can accommodate glottal flow for some time before oral pressure exceeds subglottal pressure and stops the vocal cord vibration. Since the cavity size gets increasingly smaller as we move from bilabial /b/ to alveolar /d/ and then to velar /g/, the velar has the least chance of maintaining the glottal flow and, thus, is more quickly devoiced.

It has also been suggested (Yavaş 1997) that the height of the vowel preceding the final voiced stop may be an important factor for final devoicing. Specifically, increasing the height (i.e. decreasing the sonority index) of the vowel creates a more favorable environment for the devoicing of the final voiced stop target. The reason offered for this is that high vowels (i.e. lower sonority vowels), by raising the tongue and creating more constriction than other vowels, cause higher supraglottal pressure and are more prone to devoicing (Jaeger 1978). This vulnerability to devoicing seems to be carried over to the following final voiced stop. Thus, putting everything together, we might find a variable success rate, for example, for the following different combinations with different degrees of markedness:

pig [pɪg] (velar stop preceded by a high V)	Most marked
bag [bæg] (velar stop preceded by a low V)	↑
bib [bɪb] (bilabial stop preceded by a high V)	↓
cab [kæb] (bilabial stop preceded by a low V)	Least marked

Another example to show the insufficiency of simple contrastive analysis and the necessity of the markedness considerations comes from the coda consonants. While CV is a universally unmarked syllable structure in languages (i.e. no known language lacks CV syllables), any addition to it adds a degree of markedness. A CVC syllable, while not a highly marked structure, may be completely absent from a language, or alternatively may have some restrictions regarding what class of consonants can occupy the coda position. For example, in a language such as Japanese, only /n/ is permitted as a single coda. A simple contrastive analysis will predict that any single coda other than a nasal (i.e. obstruent, liquid) in an English target word would be problematic for a Japanese speaker. While this prediction is accurate in a general sense, the degree of difficulty experienced by learners in different classes of sounds is significantly different; for example, obstruent codas present much greater difficulties than liquid codas. This situation, while inexplicable via contrastive analysis, is actually quite expected if we take into account the relative markedness of certain groups of sounds in coda position. Universally, obstruents are more marked (i.e. less expected) as singleton codas. In a language with CVC syllables, the coda position is most usually occupied by sonorants. There are two patterns that are observed in languages that allow CVC syllables: (a) obstruent and sonorant codas (e.g. English), and (b) only sonorant codas (e.g. Japanese). There is no language that has obstruent codas but lacks sonorant codas; this indicates that sonorants are more natural (unmarked) as codas than are obstruents. Actual examples from L2 learning situations support this view strongly. For example, for speakers of languages in which some obstruents and sonorants are permitted as codas, such as Korean, Japanese, Cantonese (Eckman and Iverson 1994), and Portuguese (Baptista and DaSilva Filho 1997), the difficulty encountered in learning single codas of English reflects the same hierarchy of difficulty, i.e. obstruents are more difficult than sonorants.

Patterns of acquisition of English liquids are also quite revealing with respect to markedness conditions. English makes a contrast between /l/ and /ɹ̣/ in all word positions. A language such as Mandarin restricts its contrasts between the liquids to the onset position; there are no syllabic liquids, and only /r/ is found in coda position. A simple contrastive analysis will predict that Mandarin speakers will be successful in onset position, and the liquid targets of English in other positions will be difficult. Paolillo (1995) examined the rendition of English liquids in five different environments: word-initial (e.g. rain, leaf), postconsonantal (e.g. play, free), intervocalic (e.g. around, polar), syllabic nucleus (e.g. razor, apple), and postvocalic (e.g. fall, cart), and found that there was a hierarchy of environments for successful rendition of the contrasts between the target English liquids. In descending order of favorable environments, it was word-initial, syllabic, intervocalic, postconsonantal, and postvocalic. If learners were not successful in one environment, it implied that they were not successful in the environment(s) that came after in the order. For example, if a learner had a problem in the intervocalic environment, she or he would have a problem in the postconsonantal and postvocalic environments. The explanation comes from the relative markedness of liquids in different environments, which relates to relative acoustic salience in each of these environments. Specifically, the relative salience is higher in initial or syllabic position than in other transitory positions or in clusters. This example shows that learners' difficulties cannot be explained by a simple contrastive analysis mismatch between L1 and L2, and the relative markedness of the targets in different environments should be considered.

For another example of the invaluable insights we can gain from markedness, we turn our attention to the aspirated vs. unaspirated stop mismatches between English and several other languages, which are a significant source of trouble. While English has aspirated stops in syllable-initial position, stops in languages such as Spanish, Portuguese, and so on are not aspirated. Thus, it is commonplace that speakers coming from these languages experience difficulties in their attempts to learn English; they replace the aspirated target stops [p^h, t^h, k^h] with their unaspirated versions [p, t, k]. While a contrastive analysis between L1 and L2 can predict that these mismatches will create difficulties, it cannot say anything about the varying degrees of difficulty among different targets. Several studies (Laeufer 1996; Port and Rotunno 1978; Thurnburg and Ryalls 1998; Major 1987; Yavaş 1996, 2002) found that learners experience less difficulty in acquiring the aspirated stops as we go from bilabial to alveolar and to velar. In other words, we are dealing with the relative markedness among [p^h, t^h, k^h], the first being the most marked and the last being the least marked. The reason for the varying degrees of ease or difficulty (markedness) is related to the degree of abruptness of the pressure drop upon the release of a stop. The more sudden (abrupt) the pressure drop is, the sooner the voicing of the next segment (vowel or liquid) starts. In the case of different places of articulation, differences in the mobility between the articulators involved in occlusion are responsible for the different degrees of abruptness of the pressure drop. The tongue dorsum separates more slowly (i.e. less abruptly from the velum for the velar /k/ than the tongue tip from the alveolar ridge /t/, or the lips /p/). The slower, thus longer, release delays the proper pressure differential to begin voicing for the following segment, hence the longer lag (aspiration) for velars than for alveolars and labials.

It has also been suggested (Weismer 1979; Flege 1991; Klatt 1975; Yavaş 2002) that the sonority of the following segment may influence the degree of aspiration of the stop. An initial stop seems to have a longer lag before a segment that has a narrower opening (i.e. lower sonority index), such as a high vowel, than before another that has a more open articulation (i.e. high sonority index), such as a low vowel. The reason for this is that lower-sonority items (e.g. high vowels) have a more obstructed cavity than high-sonority items (e.g. low vowels). Since the high tongue position that is assumed during the stop closure in anticipation of a subsequent high vowel would result in a less abrupt pressure drop, a stop produced as such will have a longer lag than before a low vowel.

Putting all these together, we can show the relative markedness of the following:

Least marked ↑	kit	(velar with a high vowel)
	cat	(velar with a low vowel)
	tit	(alveolar with a high vowel)
	tat	(alveolar with a low vowel)
↓	pit	(bilabial with a high vowel)
Most marked	pat	(bilabial with a low vowel)

Our final example with respect to markedness comes from a sequential relationship and looks at English double onsets in which the first member is /s/. The possibilities can be described as (a) /s/ + stop (e.g. speak, stop, skip), (b) /s/ + nasal (e.g. small, snail), (c) /s/ + lateral (e.g. sleep), and (d) /s/ + glide (e.g. swim). Several languages that allow double onsets do not have the above combinations, and Spanish is one such language. Thus, it is expected that Spanish speakers will have difficulties with the initial sC (where C = consonant) targets in learning English and, indeed, they do. What is interesting, however, is that the difficulties experienced by the learners are not the same with respect to the different combinations of s-clusters (a), (b), (c), and (d) listed above. A decreasing degree of difficulty has been observed for (a) – (d) in the learning of English: /s/ + stop being the hardest, and /s + w/ being the least difficult.

While a contrastive analysis between the two languages could predict that English initial sC clusters will be difficult for Spanish speakers (because Spanish does not have them), it will have no means of going beyond that to account for the different degrees of difficulty observed. Here, again, the explanation will come from the relative markedness of the targets. As mentioned in chapter 6, the relative naturalness of clusters is closely linked to the principle of sonority sequencing, which dictates that the sonority values should rise as we move from the margin of the syllable to the peak (nucleus). Among the targets in question, one of them, (a) /s/ + stop, violates this principle, because the first member of the onset cluster, /s/, a voiceless fricative, has a higher sonority value, 3, than the second member, /p, t, k/, which has 1. Thus, as we move from C_1 to C_2, a 'fall', rather than the expected 'rise', in sonority takes place. Since this is a highly unexpected (marked) combination in universal terms, it is not surprising that it proves to be a very difficult target to acquire. The remaining targets, (b) /s/ + nasal, (c) /s/ + lateral, and (d) /s + w/, all satisfy the sonority sequencing generalization, because there is a 'rise' in sonority as we move from C_1 to C_2 (/s/ + nasal: 3 to 5; /s/ + lateral: 3 to 6; /s + w/: 3 to 8). As we noted earlier, there was a decreasing degree of difficulty among these three targets, and this also is explainable with reference to their relative naturalness. The fact that laterals are higher in sonority than nasals, and glides are higher than laterals, results in different degrees of sharpness in the sonority jumps between C_1 and C_2, and this seems to be responsible for the greater ease of /sw/ (sonority difference of 5) than /sl/ (sonority difference of 3). Similarly, /sl/ has a bigger difference than /s/ + nasal (sonority difference of 2) and thus, expectedly, provides less difficulty.

It is also worth mentioning that speakers coming from languages that do not permit any onset clusters reveal different modification patterns with respect to different types of English clusters in contact situations. Error patterns of speakers of Egyptian Arabic, Sindhi, and Bengali (Broselow 1993) show that sonority sequencing-violating /s/ + stop clusters are modified with a prothetic vowel, while the ones that do not violate the sonority sequencing receive an epenthetic vowel, which results in a speedier, native-like pattern:

Egyptian Arabic

street	→ [istirit]	sweater	→ [siwetar]
study	→ [istadi]	slide	→ [silajd]

Sindhi

school	→ [ɪskul]	please	→ [piliz]
spelling	→ [ɪspelɪng]	slipper	→ [siliper]

Bengali

stamp	→ [istamp]	glass	→ [gelas]
school	→ [iskul]	slate	→ [selet]

While, for reasons of space, we will not go on to other examples that demonstrate the importance of markedness, similar examples can easily be multiplied for many other phonological structures. The important message that comes out of all these is to alert remediators about the indispensable nature of such information. The more one can see the highly structured nature of events, the better remediator one can become.

8.5 Ontogeny Phylogeny Model (OPM)

All the above clearly demonstrates that interlanguage phonology is governed by the following three components: L1, L2, and universal principles (markedness).

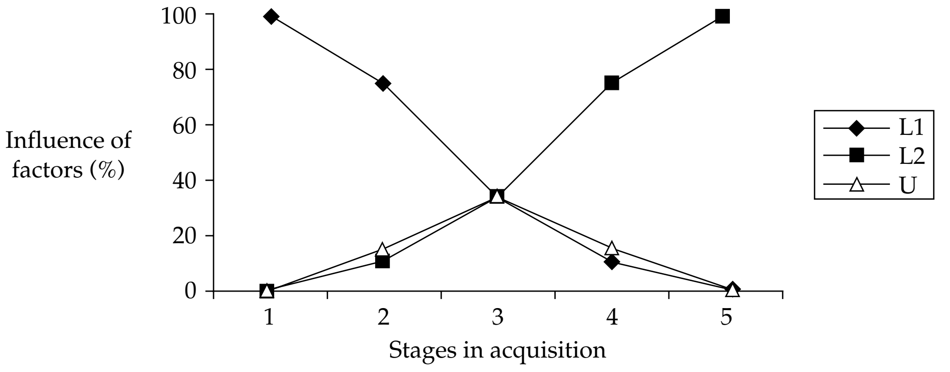


Figure 8.2 The Ontogeny Phylogeny Model: normal phenomena
(*Source*: from R. C. Major (2001) *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology*. Reproduced by permission of Lawrence Erlbaum Associates.)

Although all these factors influence the productions of learners, the role of each may be different at different stages of interlanguage development. The Ontogeny Phylogeny Model (hereafter OPM) proposed by Major (2001) deals with just that and states that in the earlier stages of L2 acquisition, L1 interference is the dominant factor; the role of universals is minimal. Gradually, the influence of L2 and universals increases, and the role of L1 decreases. In later stages of acquisition, the only element on the rise is the influence of L2, with concurrent decline of the role of L1 and universals, as shown in figure 8.2.

Although this general account may be sufficient for the normal phenomena, Major carefully points out that the proportions of the three components will vary, depending on the phenomena under scrutiny. For example, in the similar phenomena, L2 increases in a slower fashion than above and the effects of L1 also decrease slowly. The increase and later decrease of universals are slower as well. To give an example for a similar phenomenon, we can think of the relationship between alveolar stops of English /t, d/ and their slightly fronted counterparts, dental stops in Spanish and Portuguese. Since such minimal distinctions are less likely to be noticed by the learner, a Spanish speaker would be likely to retain the L1 interference longer here in his or her attempts at the target English alveolar stops than, let us say, for his or her substitutions of the alveolar trill for the English target retroflex approximant. Major's account of the similar phenomena is given in figure 8.3.

In the acquisition of the marked phenomena, earlier stages are again dominated by L1 influence, and the acquisition of L2 is again slower than for the normal phenomena. However, the effects of L1 and universals are different in subsequent stages; the effects of L1 decrease faster here and we see a rapid increase in the effects of universals. In later stages, the decrease in L1 and universals is reminiscent of the similar phenomena, i.e. slow. Major shows this as given in figure 8.4.

Major also points out that his OPM model can account for the stylistic variation in interlanguage phonological production. Accordingly, as style becomes

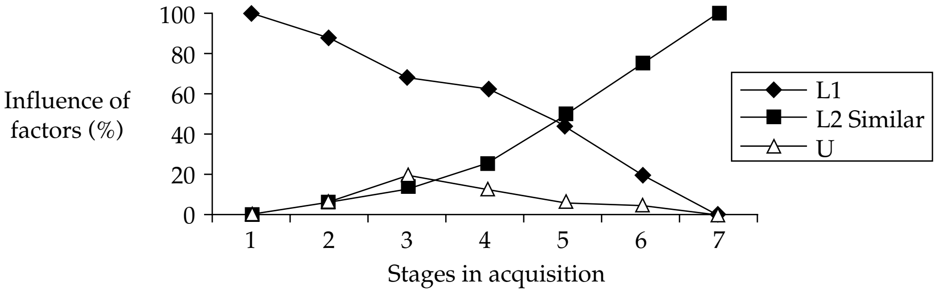


Figure 8.3 The Ontogeny Phylogeny Model: similar phenomena
(*Source*: from R. C. Major (2001) *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology.* Reproduced by permission of Lawrence Erlbaum Associates.)

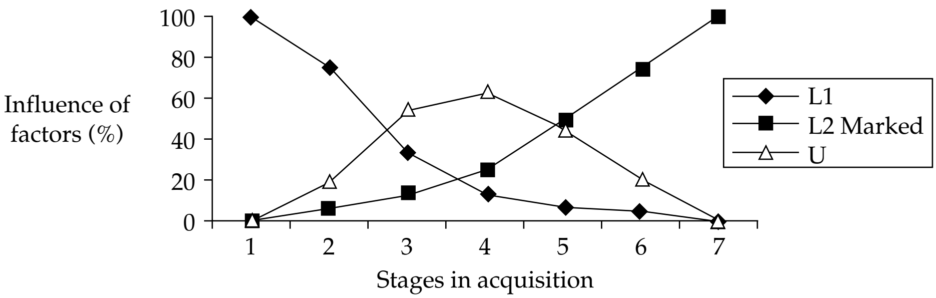


Figure 8.4 The Ontogeny Phylogeny Model: marked phenomena
(*Source*: from R. C. Major (2001) *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology.* Reproduced by permission of Lawrence Erlbaum Associates.)

more formal, L2 increases, L1 decreases, and universals increase then decrease. While this statement is generally true, we are also reminded that, depending on the stage of the learner, the proportion of the different components can vary from speaker to speaker for the same style.

8.6 Optimality Theory (OT)

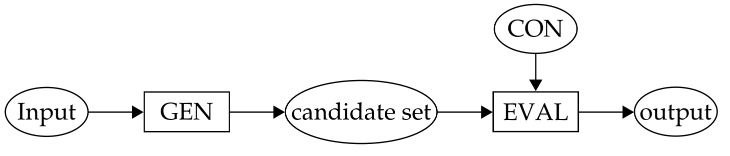
Explanations regarding the interaction of the differential effects of the interlanguage components over time, and the changing nature of the learner's language, have also been analyzed by a recent theoretical approach called Optimality Theory (OT). In the following, we will briefly describe the principles of OT and then give a few examples of its application to L2 phonology.

OT views language as a system of conflicting universal constraints, and different phonological systems as a result of different rankings of these constraints. In other words, languages have different phonologies, because

(a) languages differ in the importance they attach to various constraints (constraint hierarchy), and

(b) constraints may be contradictory, and thus be violated; if two constraints are contradictory, the one that is ranked higher will have priority.

OT has two levels known as the 'input' (underlying form), and 'output' (surface phonetic form). The theory assumes that the possible output forms for a given input are produced by a mechanism called GEN (the 'Generator') and then evaluated by a mechanism called EVAL. An evaluation for the optimal phonetic output is made by screening the candidates through the constraints, and the candidate that violates the fewest constraints is chosen as the correct output. This can be shown in the following diagram (Archangeli 1999):



Constraints are of two conflicting types:

(a) *markedness constraints*, which capture the generalizations on linguistic structures that commonly or uncommonly occur in languages ('unmarked' vs. 'marked'). Unmarked structures are universal and innate and do not have to be learned, while marked features are specific to languages and have to be learned. Sample markedness constraints include "NO CODA. Syllables must not have codas"; "*COMPLEX. No clusters"; "*V NASAL. Vowels must not be nasals".
(b) *faithfulness constraints*, which require that input and output match, so that properties of the input correspond in identity to those of the output. These are of three kinds:

MAX-IO: requires that input segments must correspond to output segments (i.e., the input is maximally represented in the output); thus there should be no deletion.
DEP-IO: requires that output segments must match input segments (i.e. the output must be entirely dependent on the input); thus, there should be no insertion.
IDENT-IO(F): requires that the input representations of place, manner, and voice features should appear in the output; thus, there should be no feature change or substitution.

In all grammars, the constraints are conflicting (Kager 1999), and thus it is not possible to satisfy all constraints simultaneously. The conflict between constraints is resolved by ranking the constraints in a language-specific fashion (constraint hierarchy). For example, one of the markedness constraints, *COMPLEX ONSET, which dictates "no onset clusters", is ranked higher in Turkish, which has no onset clusters, than in English, which allows onset clusters. The optimal output (phonetic form) will be the one that incurs the least serious violations of a set of ranked constraints. Consequently, any output candidate that violates higher ranked constraints will not be the one that will survive.

The expression of ‘domination’ (or ‘precedence’) among the constraints is given in OT by a left-to-right ordering, with the highest ranked constraint being on the left. In prose, the ranking is expressed with the use of double arrowheads: A >> B (constraint A outranks constraint B). We will illustrate these in the following sample tableau, a two-dimensional table in which the constraints are listed across the top line and the candidates down the side.

/plet/ “plate”	MAX	DEP	*COMPLEX
☞ (a) plet			*
(b) pet	*!		
(c) pəlet		*!	

The input that is evaluated is placed at the top left corner. A * in a cell indicates that the form of that row violates the constraint in that column, while *! indicates that such a violation is fatal and thus eliminates that form from further consideration. The optimal (winning) form is marked with a little hand, ☞. In this tableau, the optimal output is the faithful [plet], because the only constraint it violates is the low-ranked markedness constraint *COMPLEX. The second candidate, [pet], violates MAX, which prohibits deletion, and the third candidate, [pəlet], violates DEP, which prohibits insertion. Both of these constraints are higher ranked than *COMPLEX, but the relative ranking of MAX and DEP does not seem crucial.

If, on the other hand, the output is [pet], as commonly attested in child speech via a cluster reduction process, then we will have the following:

/plet/ “plate”	*COMPLEX	DEP	MAX
(a) plet	*!		
☞ (b) pet			*
(c) pəlet		*!	

*COMPLEX >> DEP >> MAX

Here, *COMPLEX is the highest ranking constraint and thus is placed in the leftmost position. Candidate (a) violates the highest ranking and is thus eliminated from further consideration. Between the two remaining faithfulness constraints, DEP (no insertion) and MAX (no deletion), the ranking will be in that order. Candidate (c), [pəlet], violates DEP by inserting a vowel, and won’t be selected. Candidate (b) violates the lowest-ranked MAX by deleting a consonant from the input, and thus is the choice.

L2 phonology and OT

In the following, we will give examples from OT approaches to some of the observed phenomena in L2 phonology. Our first example comes from a segmental substitution of the English /θ/ as [s] or [t] in languages that lack the interdental fricative (Lombardi 2003). What is interesting here is that some languages use [s] and others utilize [t] despite the fact that all first languages have both segments. The idea advanced by Lombardi is that languages that use the substitute [s] (e.g. German, French, Japanese) do so because of native language transfer, whereas others that use the substitute [t] (e.g. Turkish, Persian, Russian) do so because of a universal markedness constraint (fricatives are more marked than stops, thus *[continuant] >> *[stop]). Also relevant is the markedness constraint *θ, which conspires against the occurrence of interdentals in inventories. Finally, the relevant faithfulness constraint for this substitution is IDENT-manner, which is defined by the manner features [stop], [continuant], and [strident]. The explanation lies in the ranking of the manner faithfulness constraint relative to the markedness constraints. We have the following tableaux for the two different substitutions. First, we look at the situation where /θ/ is replaced by [t]:

/θ/	*θ	*cont	*stop	IDENT-manner
θ	*!	*		
s		*!		
☞ t			*	*

*θ >> *cont >> *stop >> IDENT-manner

Here, the markedness constraints are higher than the IDENT-manner, and the candidate that violates the lower-ranked constraint is chosen.

Second, we look at /θ/ being replaced by [s]:

/θ/	*θ	IDENT-manner	*cont	*stop
θ	*!		*	
☞ s			*	
t		*!		*

*θ >> IDENT-manner >> *cont >> *stop

Because of the re-ranking of the faithfulness constraint (IDENT-manner), [s] violates a lower-ranking markedness constraint and is the substitute.

Our second example will be on the native language transfer effects on complex onsets. Turkish does not allow complex onsets. When Turkish speakers learn English, target complex onsets are rendered with an epenthetic vowel (e.g., <u>group</u> [grup] → [gurup], <u>speak</u> [spik] → [sipik]). The situation can be described in the following way:

/spik/	*COMPLEX	MAX	DEP	IDENT-VH
(a) spik	*!			
(b) pik		*!		
☞ (c) sipik			*	
(d) supik			*	*

The leftmost constraint, *COMPLEX, is a markedness constraint against having onset clusters. The second and third are faithfulness constraints that disallow consonant deletion and vowel insertion. The last one relates to the vowel harmony. Candidate (a) violates the highest ranking, *COMPLEX, and is eliminated from further consideration. The remaining three avoid violating *COMPLEX; however, they do this at the expense of other constraints. Candidate (b) violates MAX (no deletion) and candidates (c) and (d) violate DEP (no insertion). In Turkish, DEP is more violable than MAX, and thus is placed lower in the hierarchy. The epenthetic vowel in Turkish is chosen from the set of four high vowels, /i, y, ɯ , u/, following the vowel harmony rules that call for agreement with the other vowel, /i/, in backness (thus, /ɯ/ and /u/ are eliminated) and in rounding (thus, /y/ is eliminated). Consequently, /i/ is inserted and candidate (c) is the surviving one.

Our final example comes from final obstruent devoicing. As mentioned earlier in this chapter, this is a common process seen in the speech of many learners of English coming from a variety of languages such as German, Russian, Turkish, Dutch, and Bulgarian, to name a few. In such cases, the explanation is based on native language interference, as these languages do not allow voiced obstruents in final position. Final devoicing, however, has also been observed in learners of English whose language does not allow any obstruents (voiced or voiceless) in final position. Broselow et al. (1998) analyze such a situation in Mandarin L1 speakers learning English. While English allows both voiced and voiceless stops in final position, Mandarin lacks both in this position. When Mandarin speakers learn English, the clash created by the above-mentioned mismatch is resolved by a variety of different strategies including epenthesis (e.g. <u>bag</u> → [bægə]), deletion (e.g. <u>bag</u> → [bæ]), and final devoicing (e.g. <u>bag</u>

→ [bæk]). The last option is an unexpected one because there is no such rule in the native language. Thus, the outcome is not a result of interference, nor is it coming from the target language.

Broselow et al. analyze the situation with the following two markedness constraints:

- NO VOICED OBS CODA: syllable codas may not contain voiced obstruents;
- NO OBS CODA: syllables may not contain obstruent codas;

and the three faithfulness constraints:

- MAX (no deletion of consonants);
- DEP (no vowel insertions); and
- IDENT (VOI): an output segment should be identical in voicing to the corresponding input segment.

Initially, the constraint ranking for Mandarin, which does not allow any obstruent codas, will be: NO OBS CODA, NO VOICED OBS CODA >> MAX, DEP, IDENT (VOI).

The learners who devoice the target final stops (instead of deleting the stop, or inserting a vowel after the stop) produce an unmarked form that is not compatible with either Mandarin or English. Broselow et al. suggest that these learners have re-ranked NO OBS CODA relative to NO VOICED OBS CODA by moving the latter lowest in the hierarchy. The situation is characterized in the following tableau:

/vɪg/	NO VOICED OBS CODA	MAX (C) DEP (V)	IDENT (VOI)	NO OBS CODA
☞ (a) vɪk			*	*
(b) vɪg	*!			*
(c) vɪ		*!		
(d) vi.gə		*!		

By re-ranking the constraints in this way, Mandarin speakers who devoice target English voiced stops are in a situation comparable to German speakers who produce all English target final stops (voiced and voiceless) as voiceless.

8.7 Perception

Learners' production is partially based on how they perceive the target sounds. Literature on L2 phonology learning was heavily focused on production until two decades ago. Since then, however, we have witnessed a surge in studies on the importance of perception in shaping interlanguage productions. The relationship between perception and production is a complicated one. Questions such as "What determines the perception of foreign sounds?" and "How do things change with the experience in L2?" have been the subject of several studies. In the following, we will briefly look at three widely discussed models that offer explanations in L2 sound perception.

Native Language Magnet theory (NLM)

This theory, developed by Kuhl (1991, 1993, 2000), aims at explaining the development of speech perception from infancy to adulthood. Its main focus is on the dependence of perception on a given representation, and its consequence for production. NLM proposes that native language categories are prototypes, which are sounds "that are identified by adult speakers of a given language as ideal representatives of a given phonemic category" (Kuhl et al. 1992). Each one of these occupies a specific location in a space defined by certain phonetic properties (e.g. vowels by formant frequencies). These prototypes act as perceptual magnets that warp the perceptual space. Once these language-specific magnets are developed, infants lose the ability to discriminate sounds that they previously could, because the magnets distort perceptual space, making certain phonetic boundaries disappear; hence the perceptual reorganization from language-general to language-specific patterns of perception.

This seems to account for the facts relating to the changing abilities of small children in sound discrimination. Very young infants are capable of hearing all differences among the sounds in human languages, whereas adults display a reduced discrimination sensitivity outside their native language. Exposure to language produces a change in perceived distances in the acoustic space underlying phonetic distinctions. Infants' precocious adaptation to the native language's sound categories and adults' difficulty in discriminating non-native phonemic contrasts have been shown in several studies. Werker et al. (1981) found that English-speaking adults had difficulty discriminating two Hindi dental and retroflex stops (/t̪/ vs. /ṭ/) that Hindi-speaking adults predictably discriminated well. Yet English-learning infants at 6 to 8 months old discriminated both Hindi contrasts. Werker and Tees (1994) showed that English learners declined in their ability to discriminate between the Nthlakampx (Thompson Salish) velar and uvular ejectives (/k'i/ vs. /q'i/) as well as the above-mentioned Hindi contrasts. For both contrasts in the two foreign languages, 6- to 8-month-olds generally performed to criterion, while only about 60 percent of 8- to 10-month-olds succeeded, and very few 10- to 12-month-olds did so; thus, it was concluded that the decline was virtually complete by 10 to 12 months, except for infants learning those two languages.

Since perceptual mappings differ for speakers of different languages, the perception of one's primary language is completely different from that required by other languages. For NLM, the presence of a L1 language-specific perceptual filter makes L2 learning difficult, as later learning is shaped by the initial mappings. Foreign sounds are drawn to the native prototypes as a function of their distance from them in the phonetic space. More distant foreign sounds either assimilate to another prototype if they are closer to it, or do not assimilate if there is no nearby prototype. Two foreign sounds that are the same distance in the phonetic space from a native prototype are predicted to assimilate to it equally so long as one is not closer to another prototype.

NLM makes certain testable predictions in the degree of discriminability of foreign sounds in relation to native prototypes: it is predicted that assimilations to more well-separated prototypes will be more successful than to less well-separated ones (Kingston 2003). For example, the prototypes of high rounded vowels that contrast in backness, such as /y/ vs. /u/, are farther apart than mid rounded vowels, /ö/ vs. /o/, because vowels are more dispersed higher in the vowel space. As a result, foreign vowels that assimilate to the high vowels would be predicted to be more discriminable than those that assimilate to mid vowels.

Despite these interesting claims, several shortcomings of NLM have been pointed out in the literature. For example, as noted by Mack (2003), the perceptual magnet effect does not seem able to account for the fact that some early bilinguals have two distinct VOT systems, whereas others with apparently similar dual-language experience, and hence presumably similar amounts of exposure to prototypical and non-prototypical vowels, do not. Therefore, it might be necessary to posit the existence of learner-specific prototypes. In that case, one would need to identify which learner-specific variables determine how a prototype is formed, which obviously is a momentous task.

It has also been shown that the perceptual magnet effect may not be robust across listener groups (Frieda et al. 1999), as well as that discrimination of unfamiliar phonetic contrasts can be improved even in adults through extensive natural experience, intensive laboratory training, or experimental manipulations that reduce task memory demands (Logan et al. 1991; Lively et al. 1993; Pisoni et al. 1982). Finally, there are several cases of children older than 12 months moving to a new country and acquiring native phonology.

Perceptual Assimilation Model (PAM)

This model, developed by Best (1995), aims to explain learner behavior in acquiring L2 sounds by accounting for the perception of the relationship between L1 and L2 sounds. The central premise is that listeners tend to assimilate non-native sounds to the native sounds that they perceive as most similar. In defining 'perceptual similarity', PAM draws from articulatory phonology (Browman and Goldstein 1986, 1989, 1992) in that it suggests that what listeners detect in speech is information regarding the articulatory gestures that generated the signal. Gestures are defined by the articulatory organs, constriction degree, and constriction locations. Categorizable L2 phonetic segments are perceptually assimilated to L1 phonological categories on the basis of their gestural similarity to L1 phonetic segments, unless they are uncategorizable (assimilated as an unrecognizable speech sound that gives rise to a new category) or unassimilable (heard as a non-speech sound).

PAM places emphasis on the perception (assimilation) of L2 contrasts by L2 learners, rather than on the perception of single L2 sounds. When non-native contrasting sounds are both categorizable, perceptual assimilation to the native system is predicted to show different degrees of difficulty. The non-native sounds may be phonetically similar to two different native phonemes and perceptually assimilated to separate L1 categories, which is termed 'two category assimilation' (TC). In such cases, the discrimination is expected to be excellent. If both non-native sounds are assimilated to a single L1 category, this will create a 'single category assimilation' (SC), which is predicted to be difficult to discriminate. SC assimilations are further taxonomized into those in which both foreign sounds assimilate equally to the single native category and those in which one assimilates far more than the other. In the latter case, the two foreign sounds differ in 'category goodness' (CG) with respect to the native category, and they are predicted to be discriminable to the extent that they do so. The members of such CG assimilations are still less discriminable than the members of TC assimilations, because they both assimilate to just one native category. Thus, we have a continuum that predicts listeners' success in distinguishing different foreign sounds: TC > CG > SC, with CG cases varying between TC and SC depending on whether the CG differences between the foreign sounds are larger or smaller.

Support for PAM's predictions are frequently found in the literature. The American English /w/ and /j/ appear to be assimilated to the corresponding Japanese /w/ and /j/. This is a case of TC assimilation. The English /ɹ̣/ and /l/ assimilating to a single Japanese /r/ is a case of SC assimilation (Best and Strange 1992). Polka (1991) reported that English listeners tended to assimilate Farsi voiced velar versus uvular stops (/g/-/G/) as a CG contrast, and Nthlakampx velar versus uvular ejectives (/k'/ – /q'/) as an SC contrast, with a tendency toward better discrimination of the former distinction, which is in accordance with PAM's predictions.

As pointed out by Mack (2003), however, it is not clear what predictions PAM would make about the formation of two phonetic systems when simultaneous acquisition of two languages or very early acquisition of an L2 occurs, as the cases the model has been primarily applied to are those in which exposure to an L2 system occurs when an L1 system has already been well established.

In addition, the model, which is primarily concerned with the role of L1 in the perception of foreign sounds, is essentially static, and it does not include any means by which an existing L1 phonemic system might be altered by exposure to non-native segmental contrasts.

Speech Learning Model (SLM)

Flege's (1995) Speech Learning Model also treats phonological acquisition with a view of phonetic approximation and interference based on perceptual judgments. This model is concerned with 'ultimate attainment', and thus focuses on long-term bilinguals and not on beginning L2 learners. SLM is built on the ideas of categorical perception and equivalence classification in the determination of how a learner will react to and ultimately acquire sounds in an L2. The phonetic perception of an L2 sound involves a comparison of the L2 sound with all sounds in the learner's L1 system. SLM claims that the two phonetic subsystems (L1 and L2) are cognitively represented in a single phonological space and mutually influence one another. Learners relate L2 sounds to L1 positional allophones, and L2 perceptual failure occurs when the L1 phonological system filters out the distinctive features of L2 sounds. L1 and L2 phonetic segments can be related along a continuum; sounds are classified as 'new', 'similar', or 'identical' on the basis of the difference between L2 sounds and existing L1 sounds, and the model predicts how the learner will react. The different categorizations are made in terms of acoustic similarity or perceived cross-language similarity. If L2 sounds are categorized as 'similar', their assimilation to the existing L1 phonetic categories will be through a process of equivalence classification, and will be produced as the L1 sound (never as an authentic L2 sound). New categories will be formed for less similar and 'new' L2 sounds.

Applied to perception of L2 contrasts, SLM makes the following predictions: 'identical' sounds will present no problem for the learner, as all necessary knowledge is already available in the L1 (cf. PAM's TC). If two contrasting sounds of L2 are designated 'similar' and both are assimilated to the same L1 category, discrimination will be difficult (cf. PAM's SC or CG). For example, as we noted in section 8.2, the /ɑ/ – /ʌ/ contrast of English (e.g. <u>body</u> [bɑdi] vs. <u>buddy</u> [bʌdi]) creates lots of problems for speakers of several languages (e.g. Spanish, Turkish, Greek, French, Arabic, and Russian, to name a few), because these sounds are perceptually assimilated to [a] in L1s, and result in discrimination difficulties of the contrast, as well as the accented production of both English vowels. If, on the other hand, there is great dissimilarity between L2 and L1 sounds, the sound will be judged 'new' (cf. 'uncategorizable' in PAM), and it will not be assimilated to any L1 category. For example, English speakers learning French as a second language could produce French /y/ (a 'new' vowel for English speakers) more accurately than French /u/, because French /y/ is perceptually more distant from the closest English vowel than is French /u/, which has a near (but not identical) counterpart in English /u/.

SLM holds the view that there is no critical period after which the learner will be unable to acquire an L2 sound system; that is, adults can retain the capability for accurate perception of L2 contrasts. However, it is also stated that L2 development is constrained by age of learning. It is predicted that learners are more likely to have native-like perception with early age of learning (pre-puberty). The later the age of learning, the less likely a learner is to hear the differences between L1 and L2 sounds, because the learner's L1 categories will be more developed and are likely to impede the formation of new categories for L2 sounds. The model also states that L2 development is further constrained by the amount of L1 use. It predicts an inverse relationship between frequent use of L1 and attainment of native-level L2 perception. That is, L2 learners who use their L1 frequently will be less likely to have native-level L2 perception.

Although SLM states that accurate L2 segmental production cannot occur unless there is accurate perception, Flege (1995) does not claim that *all* foreign accent is perceptually motivated. For example, it is acknowledged that the typical Spanish-accented English production of [ɛskul] for school can only be accounted for with reference to phonotactic constraints.

There have been several studies in the literature that showed considerable support for SLM (Bohn and Flege 1990; Flege et al. 1994, 1997; Fox et al. 1995; Rochet 1995). A more recent study (Aoyama et al. 2004) examined the role of the L1 and the perceived phonetic (dis)similarity between L1 (Japanese) and L2 (English) sounds in the production of English /ɹ/ and /l/ by Japanese speakers. Since English /l/ is perceptually more similar to Japanese /r/ than English /ɹ/, it was hypothesized that Japanese learners of English would have greater difficulty in acquiring the L2 (English) lateral liquid than the non-lateral liquid. The study looked at L2 perception and production of English liquids by Japanese children and adults at two intervals separated by one year. The results, in general, supported the hypothesis. While Japanese children's perception of English /l–ɹ/ and /ɹ–w/ contrasts showed significant improvement after a year, the adults did not show any improvement over time. Also, the children showed greater improvement over time in the production of English /ɹ/ than English /l/. For /l/, neither the children nor the adults showed significant improvement. These findings support SLM's predictions in that they show better acquisition of the more dissimilar L2 sound (English /ɹ/) than the similar one (English /l/). Also, age-related differences in the rate of acquisition were apparent, as only the children showed significant gains over time.

There are, however, several studies whose findings are at odds with SLM's claims. For example, Zampini (1998; Zampini and Green 2001), examining VOTs of /p/ and /b/ in Spanish and English, found that students enrolled in an advanced undergraduate course in Spanish phonetics showed significant changes toward Spanish-like categories (toward more short lag for English /p/) in both production and perception, but there was very little relationship between production and perception. Sheldon and Strange (1982) found that Japanese learners of English /ɹ/ and /l/ performed better in production than in perception, a finding that is certainly at odds with SLM's claim that accurate L2 segmental production cannot occur unless there is accurate perception. De Jonge (1996) examined the production and perception of /ɹ/ and /l/ for Japanese speakers and of /b/ and /p/ for Arabic speakers. She found that while Arabic speakers mastered the contrast at early stages of proficiency, Japanese speakers did not even at high levels of proficiency. She used her results to evaluate SLM's (and PAM's) claims and found that neither of them could fully account for the data. There is also a problem with SLM's claim on 'new category'. For example, French front rounded vowel /y/, which is a 'new category', would be perceptually differentiated accurately from both back rounded and front unrounded French vowels, as well as from English vowels. However, some studies (Strange et al. 2004, 2005; Levy 2004) showed conflicting findings on perception of this vowel by American listeners.

The three models that have been looked at here (NLM, PAM, and SLM) capture the important insight that non-native contrasts are not uniformly poorly perceived. Instead, the difficulty with which a particular non-native contrast is perceived by listeners from a particular L1 background depends on the relationship between the sounds of the L1 and L2 in question. We also see that the perception–production relationship is a complicated one. Learners can have highly accurate perceptual abilities, but relatively inaccurate production ones. Alternatively, they may have more target-like production abilities than their perceptual ones. The different cues and skills used in perception and production are real challenges in understanding the learner's knowledge.

Summary

In this chapter we looked at several important variables that are influential in shaping the phonological productions of L2 learners. We saw that contrastive phonological information can accurately pinpoint several difficulties that are encountered by learners of specific languages. Beyond the simple contrastive patterns, however, lie deeper principles that can account for different degrees of difficulty related to phonemic contrasts. Target contrasts are incorporated into the interlanguage phonology progressively; learners seem to have greater facility in creating a phonemic contrast of the target language in basic (tautomorphemic) contexts than in derived (heteromorphemic) contexts. Also observed is that whenever we have two intersecting interlanguage substitutions, one of these is systematically blocked (i.e. deflected contrast), and hypercontrasts are results of overgeneralization.

Native language patterns that are in conflict with those of the target language alone are not sufficient to account for all of the learners' difficulties; markedness of the L2 structures also plays an important role in shaping the interlanguage phonology. Major's Ontogeny and Phylogeny Model, dealing with the three components of interlanguage phonology – L1, L2, and universals – has different predictions about the relative weight of these factors in the acquisition of different phenomena. Similar phenomena and marked phenomena are acquired more slowly than normal phenomena. In the earlier stages of acquisition, the patterns are basically governed by the effects of L1 for all phenomena, the effects of universals are minimal, and the gradual decrease of L1 influence is slower in similar phenomena than in others. In later stages, the influences of L1 and universals decrease more slowly in similar and marked phenomena than in normal phenomena.

We also looked at Optimality Theory, a model that deals with the role of markedness and language transfer effects, and their interaction. OT assumes that interlanguage grammars are natural, dynamic systems in the process of accommodating new inputs, and that L1 influence and markedness effects are merely a consequence of the system's design.

Finally, in addition to these phonological approaches to acquisition, we considered the role of perception in production and looked at perceptual models, which concentrate more on phonetic approximation and interference based on perceptual judgments.

All the above are indicative of the fact that the learning of L2 phonology is a highly structured process, and thus attempts at remediation should consider as many of these factors as possible. The capabilities of practices of remediators (language teachers, speech therapists) will definitely be enhanced by the inclusion of a greater number of linguistically based courses in their training.

Exercises

1. First, transcribe the following word-pairs, and then, with the contrastive information you had in this chapter, identify the languages whose native speakers would have problems related to these target English word-pairs.

   cheap – chip
   sieve – save
   age – edge
   bend – band
   band – bond
   fool – full
   backs – box
   look – Luke
   feast – fist
   wait – wet
   slept – slapped

2. Now, do the same for the following target pairs in contrast.

   glass – grass
   peach – beach
   pour – four
   went – vent
   feel – veal
   vowel – bowel
   dense – dens
   three – tree
   thick – sick
   those – doze
   leaf – leave
   rope – robe
   stow – stove
   curved – curbed
   math – mat
   forth – force
   soothe – sued
   clothed – closed
   sin – sing
   cart – card
   thin – chin
   lamp – ramp

sift – shift
sink – zinc
cheer – sheer
surge – search
dug – duck

3. Now, do the same for the following triplets.

   huck – hock – hawk
   panned – punned – pond
   bag – bug – bog
   bid – bead – bed
   stack – stuck – stock

4. Although contrastive phonological information is indispensable for the prediction of learners' difficulties, it is not sufficient in many cases, because for certain phenomena, constraints based on universal markedness have been shown to be influential in explaining the degree of difficulty of targets. Order the following targets in terms of difficulty (from most difficult to least difficult), and state the rationale.

   (a) single-coda consonants:
   deal, deer, deem, beat, beach

   (b) liquids:
   /l/ full, elect, lamp, fly, belt
   /ɹ̣/ green, boring, tire, room, card

   (c) /s/+ C onsets:
   slow, sticker, swing, small

   (d) aspiration:
   pig, keep, park, course, torn, tease

   (e) final voiced stops:
   lab, bid, rod, rag, rib, wig

5. Japanese lacks English target /θ/ and learners replace it with a [s] (e.g. <u>thank</u> [sæŋk]). Also, [ʃ] is an allophone of /s/ in Japanese before /i/. This results in renditions such as <u>sip</u> [ʃɪp]. While we have these two patterns (/s/ as [ʃ] before /i/, and /θ/ as [s]), Japanese speakers'

rendition of English <u>think</u> is [sɪŋk] and not [ʃɪŋk]. Does this support or counter the case made for deflected contrast in section 8.3.2? State your reasoning.

6. Transcribe the following (on "American English") from T. McArthur, *The English Languages* (Cambridge: Cambridge University Press, 1998, pp. 220–7).

   (a) The American I have heard up to the present is a tongue as distinct from English as Patagonian.

   (Rudyard Kipling, 1889)

   ....

   ....

   (b) The rich have always liked to assume the costumes of the poor. Take the American language. It is more than a million words wide, and new terms are constantly added to its infinite variety. Yet, as the decade starts, the US vocabulary seems to have shrunk to child size.

   (Stefan Kanfer, 1980)

   ....

   ....

   ....

   ....

   ....

   (c) I mean that almost everyone who touches upon American speech assumes that it is inferior to British speech. Just as the Englishman, having endured for a time the society of his equals, goes on to bask in the sunshine of aristocracy, so the American, when he has used the American language for business or for familiar intercourse, may then, for higher or more serious purposes, go on to the aristocratic or royal language of Great Britain.

   (Fred Newton Scott, 1917)

   ....

   ....

   ....

   ....

   ....

   ....

   ....

   ....