Chapter II

Pronunciation Difficulty and Learner Production Strategy

In order to evaluate the particular problems which various learners have with pronouncing complex codas and to rank them in an accuracy hierarchy, it is necessary to consult previous findings in the published literature.

2.1 Cross Linguistic Explanations for Pronunciation Difficulty

For reasons noted in sections 1.1.2 and 1.1.3, TESOL professionals, particularly the dominant CLT- oriented ones, have not recently accorded much importance to the pronunciation of the **–ed** and **–s** inflections. Hence, one must turn to the applied linguistics literature for the description and diagnosis of the problem, if not for its remedy. Here the pronunciation of the **–ed** and **–s** inflections has been addressed in greater length and in greatest detail, for it has been examined in a perhaps more systematic and logical way than the more psychological approach of language pedagogy. The main virtue of this linguistic approach is that it views languages comparatively, and takes account of the L1 when making observations about the learning of the L2. It does not assume that all learners of a language will face identical problems, but will encounter problems that are rather predictable, based on their own native languages. Generally, early schools of thought contended that learner difficulties stemmed from the rules of the native L1 transferring to the learner

L2, while later hypotheses have also noticed developmental patterns that are more universal. Subsequent to the development of phonological theory by Trubetskoy (1939) and others dating from the 1930's, the problems of inter-language speech have been studied and various explanations for the deviations from L2 speech (hypotheses) have been put forth by applied linguists.

2.1.1 Contrastive Analysis (CA)

Leather & James (1996) relate that, "It is evident that the role of L1 in L2 speech acquisition has formed a major, if not *the* major focus of attention almost as long as L2 speech has been studied". Phonological transfer theories date from the 1940's. One such theory, *Contrastive Analysis* (CA) describes transfers involving *sound substitution*, *allophonic processes*, *under-differentiation*, *over-differentiation*, *reinterpretation of distinctions*, *phonotactic interference*, and *prosodic interference*. Weinreich (1953) called these seven types of transfer, *interferences*, and they may apply to speakers of most, if not all foreign languages. Lado (1957: 11) claimed that, "We have ample evidence that when learning a foreign language we tend to transfer our entire language system in the process". L2 pronunciation errors were usually explained by comparing the phonemes and their allophonic distributions within the L1 and L2. Lado (1957) asked which contrasts would create the maximum phonological difficulty, and posited that the greatest difficulty would occur with the learner needing to assign two or more allophones in the L1 to different phonemes in the L2. His example involved Spanish learners of English speaking the sounds [d] and [ð], which are allophones of /d/ in Spanish, but which are phonemes (contrast) in English.

Analogous English allophones which would be difficult for English NS to use correctly when speaking Thai would thus be the unaspirated stops and affricate [p], [t], [k], and [tʃ], as represented by the Thai phonemes, ป, ต, ก, and จ[7] respectively.

Native English speakers often do not devoice จ and ก when speaking Thai, despite the fact that these unaspirated consonants are respectively allophones of the voiceless /tʃ/ and /k/, e.g., in "exchange" /ekstʃeɪndʒ/ and "sky" /skaɪ/. (By the same token, it may also be rare to find native Thai speakers who will readily voice /g/ and /dʒ/). On the other hand, are these frequent errors grave? Probably not, as /dʒ/ and /g/ do not exist in Thai, and so are taken for the unaspirated จ and ก. The Spanish English learners in Lado's example did commit gravity errors because [d] and [ð] are different phonemes in English—which when used incorrectly could be mistaken for different words. The case of the Thai consonant phonemes, ป /p/ and ต /t/ is similar, and can create both common and grave errors for English NS when speaking Thai. Thus, for example, such a speaker might pronounce the Thai word ตาย /taaj/ (die) variably as [tʰaɪ], [taɪ] or [daɪ], which in turn might be interpreted by native Thais as six different words, namely: [tʰaaj] ทาย (predict) or [tʰaj] ไทย (Thai), [taaj] ตาย (die) or [taj] ไต (kidney), and [daaj] ดาย (often used in compounds) or [daj] ใด (any, which). Here the ambiguity is compounded because of the contrastive vowel length which exists in Thai but not in English.

Hammerly (1982) evaluated English speaking learners of Spanish and found that their greatest difficulty was in suppressing automatic L1 allophonic processes; their second greatest difficulty was producing L1 allophones with a different distribution than found in the L1. Hence, allophones which would be difficult for Thai learners of English to use correctly would be the set of voiced stops which neutralize (devoice) in final position, บ /b/, ด /d/. The English consonant phonemes, which when transliterated into Thai, are pronounced as [t]: ส /s/, ช /tʃʰ/, จ /tʃ/; or [n]: ล /l/ post-vocalically in Thai, particularly in loanword adaptations, are a special case, as has been pointed out in section 1.7. As the English phonemes /s/, /l/, and /tʃ/ often occur post vocalically, another kind of automatic process analogous to an allophonic process often can occur. In an evaluation of Thai ESL students, Hancin-Bhatt (2000) found that they gradually re-ranked their 'faithfulness' constraints to accommodate postvocalic fricatives, e.g., /s/ and liquids, e.g., /l/. English loanwords have increased the original set of 12 beginning consonant clusters to perhaps as many as 18. New clusters: /br/, /bl/, /fr/, /fl/, /st/, and /dr/[8]. Final /f/ and /s/ are becoming more common. A "safe" is commonly referred to as a ตู้เซฟ [tûu seéf] or [tûu seép˺].

The case of postvocalic /f/ (ฟ, ฝ) is interesting, as it did not exist in writing prior to the introduction of loanwords from English, and so there were no allophonic or orthographic-phonological rules assigning a different sound after the vowel, as there are with the consonants discussed above. The case of post vocalic /s/ shows that its frequent occurrence in English (hence, frequent adoption in loanwords) is perhaps overwhelming the Thai orthographic-phonological rule (see Page 5).

The strong form of CA attempted to explain the difficulties that L1 learners have in mastering L2 phonology in terms of the phonemic and allophonic differences between the L1 and L2 systems. The strong form of CA has been generally abandoned since the 1970's when Kohler (1971) and others showed that a comparison of L1 and L2 phonologies could not account for all errors in L2 speech. Fricatives like /θ/ and /ð/ and voiced final *obstruents* (ending /b/, /v/, /ð/, /d/, /z/, /ʒ/, /dʒ/, /g/) are acquired late not only by L2 learners, but also by L1 learners (Bernhardt & Stemberger, 1998). In addition to languages having differing priorities or constraints, there are different degrees of difficulty or *markedness* of the phonemes themselves.

2.1.2 The Markedness Differential Hypothesis (MDH)

Eckman (1977:315-330) defines typological markedness logically as follows. "A phenomenon X is more marked than a phenomenon Y if its presence in a language implies the presence of phenomenon Y, but the presence of Y does not imply the presence of X". Typological Markedness is usually described in terms of relations to other languages. Examples of marked features would be voiced final obstruents, since their presence automatically implies the existence of voiceless final obstruents, but not vice-versa, and 3-consonant ending clusters, since their presence automatically implies the existence of 2-consonant cluster finals, but not vice-versa. Eckman attempted to improve on the strong form of CA by taking a different approach. Instead of transfer processes mechanically leading to error reflective of the L1 phonology, his Markedness Differential Hypothesis (MDH) tries to predict exactly which phonological areas will cause L2 learners the most trouble: Those areas of the target language which differ from the native language *and* are more marked than the native

language will be difficult; the more marked these areas are, the more difficult they will be to learn; and those areas of the target language which are different from the native language, but are not more marked than the native language will not be difficult.

Hence, the dental fricatives, being universally marked phonemes, will be difficult for Thai (and other) learners of English. In final position they are even more marked, and so they will be yet more difficult. The ending voiced interdental fricative, /ð/ is the most marked English segment of all (since voiced final obstruents are also marked), and when combined with another consonant to form a consonant cluster, as in "bathed" /beɪðd/, would be a challenge for speakers of L1's that do not have voiced final obstruents, dental fricatives, and consonant clusters. Thais would have even more difficulty with ending /ðd/ or /ðz/ finals after the English diphthongs, /aɪ/, /aʊ/, and /ɔɪ/ in words such as *mouths* /maʊðz/, or *writhed* /raɪðd/, since no consonant can follow these diphthongs within the Thai syllable. This is because the Thai equivalents of /aɪ/, /aʊ/, and /ɔɪ/ are usually analysed in Thai as /aj/, /aw/, and /ɔj/, i.e., as already containing an ending consonant. *Writhed* violates Thai phonology three ways, viz., by point of articulation (inter-dental fricative), final voicing, and triple consonant cluster, namely, /rajðd/. (This phonotactic constraint disallowing consonants following a diphthong is examined in the present study).

While Thai phonology contains both voiced and voiceless plosives, the voiced plosives neutralize (devoice) at the end of syllables, so it follows that voiced complex codas would probably be harder than voiceless. Of six English fricatives (/v/, /z/, /θ/, /ð/, /ʃ/, /ʒ/) which have no equivalents in Thai, four are voiced, and both of the Thai

consonants which do have somewhat close English equivalents (ก for /g/, and จ for /dʒ/) are voiceless, unlike their English equivalents. Thus, the Thai phonemic inventory has fewer voiced consonants than the English—as is also shown in Section 1.7 under the definition of the term **Voiced**. Since voiced consonants and their allophones in codas are more marked than their voiceless counterparts, the MDH and the SCH also would predict more difficulty with voiced complex codas.

By common consent, /ð/ is the hardest and most marked English phoneme in final position, even for NS. The voiced /ʒ/ would probably be quite difficult for Thais, as they lack even its voiceless counterpart /ʃ/. A provisional ranking of the marked English phonemes from hardest to easiest might be: /ð/, /θ/, /dʒ/, /ʒ/, /z/, /v/, /tʃ/, and /ʃ/. There is some question of whether the affricate pair, /dʒ/ and /tʃ/ are more marked than their final sibilant-fricative constituents, /ʒ/ and /ʃ/. Thai actually contains a close equivalent (ฉ) to the voiceless affricate, /tʃ/, so it may well come last. However, the Thai orthographic-phonological rules would prevent its realisation as [tʃ] after vowels. The position of /tʃ/ is probably comparable to /s/: its frequent postvocalic occurrence in English may eventually lead to its realization as [tʃ] with more advanced speakers, or as Thais adopt more English loanwords containing postvocalic /tʃ/. But it is safe to say that postvocalic /s/ is more widely spoken than postvocalic /tʃ/ both in English and Thai.

There is no doubt that triple consonant clusters constituting a complex coda are more marked than double clusters, and that quadruple consonant clusters are yet more marked. Only a few languages exceed the complexity of English complex codas—

among them Norwegian, with as many as five consonants, and Georgian, with six or more.The negative correlation in English between voicing and complexity, pointed out in section 1.3.5, may indicate that voicing and complexity are both marked 'competitors' which seldom inhabit the same extremes. Thus, we find no voiced quadruple-clusters and a minority of voiced tripletons. The MDH would predict a descending mean accuracy order of doubletons, tripletons, to quadrupletons. Their relative frequencies would also indicate that learners would have been exposed to fewer higher-order complex codas.

The MDH does not predict how L2 learners will alter or simplify marked structures, and in this sense is not an advance on CA, which did attempt to so predict (the 7 transfer interferences). The MDH also tends to predict (wrongly) that similar structures will cause no difficulty when they are marked. Altenberg & Vago (1983) and Eckman (1984) found that Hungarian and Farsi speaking learners of English respectively tended to devoice final obstruents, despite the presence in both languages of final voiced obstruents. This seems to indicate that the process of voicing final obstruents is developmental in the inter-language (learner L2), itself. Eckman attempted to rectify the MDH with his Structural Conformity Hypothesis (SCH) (Eckman 1991:24), which states that inter-languages (learner L2's) conform to the same universal rules as do the various L1's, and hence, more marked structures are more difficult, and will involve more errors. Eckman (1991), Carlisle (1997, 1998) and Eckman & Iverson (1994) considered the case of consonant clusters in onsets or codas, where the L2 (English) allowed both a greater number of clusters, as well as more marked clusters, than did the L1. All three studies evaluated L1 speakers of languages that do not permit consonant cluster finals (Spanish, Cantonese, Korean, and Japanese). The studies tended to indicate that more marked structures were generally

acquired after the less marked structures. Mandarin speakers tended to devoice final English obstruents even though their L1 does not *have* any final obstruents. Again, a universal developmental process seems to be at work.

2.1.3 Morpheme Acquisition Order (MAO)

Krashen (1977) reviewed studies of morpheme acquisition order among disparate groups of L1 learners of English and found that **-ed** inflections arising from regular verb past tense appear relatively late in the acquisition process—later than irregular verb forms, and especially in comparison with the appearance of the continuous **–ing** suffix, which is always a syllabic allomorph. Among **–s** inflections, plurals were acquired shortly after **–ing** inflections, while possessives and third person singulars were acquired late. This last finding would indicate that MAO is not fully dependent on phonological difficulty, but may also relate to semantic difficulty, e.g., plurals are conceptually easier and more easily memorized for learners than are third person singulars or possessives, which also appear less often. Conversely, it might also predict that **–s** codas are easier than **–ed** codas, as *some* of the former (namely, the plurals) are acquired relatively early. The existence of an entire class of /s/ and /z/ codas which are earlier acquired may indeed indicate that they are on the whole more easily pronounced than /t/ and /d/ codas resulting from the later-acquired regular past tense.

2.1.4 Feature Permutations

In an evaluation of 60 Thai Matayom 5 students' pronunciation of 6 different doubleton feature permutations, Mano-Im, R. (1999) found the following order from most to least accurate: Nasal-Stop /nt/, Nasal-Fricative /ns/, Lateral-Stop /lt/, Nasal-

Affricate /ntʃ/, Stop-Fricative /ks/, and Stop-Stop /kt/. It is noteworthy that the two easiest permutations began with a nasal /n/, and that the two hardest began with a stop /k/. Unfortunately, each permutation was represented by only a single coda. It stands to reason that there might be significant variations in the accuracy outcomes from differing places of articulation or voicing status, viz., that /ps/ might be easier or more difficult that /ks/, let alone /gz/.

Since Thai has far fewer fricatives (3) than English (9), and as all fricatives, affricates, and laterals become either voiceless stops or nasals, and do not remain unaltered post vocalically as they do in English (Appendix A), there might be substantial differences in the accuracy rates between permutations involving stops and nasals vs. fricatives, affricates, and laterals. Mano-Im's finding that the Nasal-Stop permutation, at least in its /nt/ incarnation, had the highest accuracy rate, might follow from the absence of allophonic processes in Thai for both nasals and voiceless stops. The question remains, would any accuracy hierarchy be consistent for both voiced and voiceless specimens of a given feature permutation—due to allophonic processes and the markedness of the former?

2.1.5 Natural Phonology

One theory which uses the idea of typological markedness to describe the evolution of phonological production in child L1 speech, and attempts to apply this to adult L2 speech is *Natural Phonology* (Stampe, 1979). Major (2001) describes the process of child L1 learners learning how to produce final voiced consonants in English; there are similarities to the processes of most adult L2 learners. Some studies (Levelt, C., Schiller, N. & Levelt, W, 2000; Stites, J., Demuth, K., & Kirk, C, 2004) have

shown fairly uniform developmental processes with young children's acquisition of consonant finals—from simple to more marked—in both Dutch and English, though there are apparent differences in the order in which various features (stops, fricatives, liquids, etc.) are acquired. Major (2001) cites the common progression of deletion, substitution, epenthesis, and final (correct) form for native Brazilian learners of English with tasks like pronouncing consonant finals. Zamuner & Gerken (1998) studied the effect of syllable stress on pronunciation by young children of word final position consonants. They found that English-speaking two-year-olds were more likely to delete word-final codas in unstressed syllables than in stressed syllables. Codas in monosyllabic words had more perceptual salience than those in bi-syllabic words. This would indicate that a proper setting for the evaluation of complex codas should be uniform, with only monosyllabic examples of the various complex codas chosen, if possible.

2.2 Ranking Complex Codas by Difficulty

Generally, CA predicts frequent phonological problems stemming from different allophonic rules, while the MDH predicts rarer, but less surmountable problems stemming from the phonotactics, or the phonological limits of a language. There is thus more chance that an allophonic rule, such as /b/ → [p] after vowels might be overcome than that the dental fricatives will be mastered, especially in postvocalic position. To the extent that an orthographic-phonological rule, such as /s/→ [t] after vowels shares a similar habitually-induced impediment to language acquisition as the allophonic processes, CA might also indicate problems with the consonants /s/, /l/, /tʃ/, and /dʒ/ in postvocalic position, where Thai orthographic-phonological rules assign

different consonants, namely, [t] [n], [t], and [t], respectively. The MDH would indicate that the more marked phonemes— dental fricatives, voiced obstruents, and complex codas containing three or more consonants would be hardest. Previous findings in MAO indicate that /s/ and /z/ coda finals resulting from (plural) –s inflections are easier than /t/ and /d/ finals resulting from regular verb past tense –**ed** inflections. Applying all these linguistic theories, one of the easier codas should be /fs/, e.g., ***beliefs***, a doubleton (MDH) which ends in /s/ (MAO), is voiceless and does not contain a dental fricative or other marked phoneme (MDH), and for which the first postvocalic consonant (/f/) is not subject to allophonic rules (CA). This coda has a frequency of 0.1 per 1,000 (Appendix D). The most common /nd/ coda, e.g., ***rained***, which occurs 31.7 times per 1,000 words, or 317 times as often, would be hypothetically more difficult because of its voiced –**ed** status. The relative accuracy outcomes of these two codas would be a good test of the validity of the theories. The degree of learner exposure to the more common codas may count more than some linguistic features. But it is unlikely that coda frequency would entirely govern accuracy, and it is probable that accuracy outcomes would deviate from frequency in ways that could be explained by at least some linguistic theories.

2.3 Learner Repair Strategies for Complex Codas

Because most languages (section 1.3.1) favor CV syllable structures, and may at most contain CVC or CCVC syllable structures, TESOL learners often employ various repair strategies to deal with the more complex syllable structure of English. These strategies include vowel epenthesis (a short vowel inserted between two consonants), paragoge (a short vowel inserted right after the complex coda), consonant deletion, consonant substitution, or metathesis (switching the order of consonants). A

CVCC syllable undergoing epenthesis may be pronounced as CVCVC. A CVCC syllable undergoing paragoge may be pronounced as CVCCV. In both cases one syllable has been transmuted into two, while the consonantal data has been retained, if colored by the surrounding extraneous vowels. A CVCC syllable undergoing deletion of one of its final consonants becomes CVC, and potentially crucial consonant data is lost. Substitution of one or more of the final consonants by easier consonants can also mar the intelligibility of a syllable. With metathesis, a consonant's position is moved, usually from the coda to the onset: CVCC can become CCVC. This also affects intelligibility. Occasionally metathesis can proceed from the onset to the coda: CCVC can become CVCC. There is yet a third kind of metathesis occurring solely within the coda: CVCC remains CVCC with the position of the consonants reversed.

2.3.1 Deletion, Substitution, and Epenthesis

Epenthesis would seem to be the favored strategy for dealing with the phonotactic constraints of the L1, as it retains the phonemic data of the L2. Yet it is not a universal strategy among TESOL students. Hancin-Bhatt (2000) ranked the phonotactic constraints of her Thai L1 subjects and found that the highest ranked constraint was against complex codas, and the lowest ranked constraints, and thus, the most likely learner strategies were deletion and substitution. Abrahamsson (2002) in an evaluation of disparate L1 learners of Swedish, a language which also involves complex codas, found that,

> An increasing ratio of epenthesis-to-deletion is the first-order indicator of increasing L2 proficiency during early stages of acquisition, but increased target-like production becomes the first-order indicator of development at later stages. Codas that are essential for the retention of semantic information are preserved through higher accuracy rates and higher relative levels of epenthesis errors.

In Major's (1994) study of four Brazilian learners, more target-like pronunciation of consonant clusters occurred in the less formal text readings than in the more formal word lists. Weinberger's (1987) study of Chinese English learners found more epenthesis errors in the more formal task of paragraph reading than in the less formal task of story telling. Lin, Y-H (2003) reviewed some of the recent literature on the interlanguage variability of different L1 English learners, finding that the accuracy of pronounced consonant clusters was not highest in the most formal tasks, such as reading word lists. Lin Y-H (2001) evaluated 40 Taiwanese informants in different reading tasks involving beginning English clusters, and found that epenthesis was a preferred strategy for more advanced students and in formal reading tasks. Deletion was more often used by less advanced students and in less formal conversational tasks. The epenthesis/deletion ratio was as high as 18 when both the former two situations obtained, and below 1 when both the latter two situations obtained. Though epenthesis and deletion are both pronunciation errors, Abrahamsson and Lin have made a pedagogical advance in treating these errors as symptomatically different. Epenthesis acknowledges that the speaker is trying to ameliorate his/her pronunciation to accommodate the listener in the more formal, 'correct' task. The trade-off between error *rates* (epenthesis in the formal tasks) vs. error *gravity* (deletion in the less formal tasks) is left unanswered in Lin's study, but is a question which the TESOL profession should attempt to deal with. What would be the optimal balance of epenthesis and deletion which would lead to maximal intelligibility for given learners? The present study attempts to assign error-weightings based upon error gravity.

2.3.2 Extraneous High Tones

Kenstowicz M. & Suchato A. (2005) maintain that the iambic prosodic structure of Thai (strong or heavy final syllables) may discourage a repair strategy such as schwa paragoge from being used by Thai English learners. They furthermore cite a peculiarity when Thais pronounce English complex codas: they tend to associate a high tone with certain permutations—particularly voiceless Nasal-Stop—which they then pronounce without the stop. Example: the loanword, 'pump' /pʌmp/, which is pronounced /pʌ́m/. In fact, they found that voiceless codas as a whole are produced with a high tone:

> A search through the loanword corpus for monosyllabic English loans ending in an obstruent showed a strong tendency to be High if voiceless (76 examples vs. 9 with Low). If the English source word ends a voiced obstruent, then the loan was still likely to be H but less strongly so (14 High vs. 10 Low). This difference was statistically significant (p =.001).

While intonation patterns are perhaps analogous to the use of glottal stops as substitutions for a consonant which cannot be pronounced, in that they are evidence of an intention on the part of the speaker, they still do not assist the listener from a dissimilar linguistic background or the non- linguistic specialist in ferreting out *what* was intended, and are not evaluated in the present study.

2.4 Summary

Problems with complex coda production have mainly been addressed within applied linguistics, rather than within the pedagogical core of TESOL, even though it falls to the latter practitioners to effect any improvements which are possible with a

given set of learners. The cross-linguistic hypotheses cited in sections 2.1.1 to 2.1.3 seldom address complex codas specifically, but the more general areas of contrast with the L1 and the markedness of linguistic structures in the L2—of which complex codas may form a part. Nevertheless, the hypotheses have been provisionally applied in section 2.2, and one research question of the present study is: what are the correct applications in predicting areas of difficulty which a learner population will experience. Naturally, determining difficulty involves assessment of complex coda production, preferably with reference to the most common and typical errors, which have been noted in section 2.3 et seq.

Linguistic hypotheses such as Contrastive Analysis, the Markedness Differential Hypothesis, and Morpheme Acquisition Order would indicate overlapping, but slightly different areas of difficulty with Thai learners' pronunciation of English complex codas. Previous evaluations of Feature Permutations align with Contrastive Analysis insofar as those permutations which do not violate Thai allophonic rules were pronounced more accurately than those which do. Nevertheless, there is some ambiguity in the interpretation of the first two hypotheses: which consonants are more problematic, postvocalic or ultimate? Which factor among competing marked characteristics (voiced, feature, number of consonants) is of primary concern to pronunciation accuracy, and which are secondary? Would this relate to the issues of error gravity? Natural Phonology has indicated a developmental sequence from deletion to substitution to epenthesis. There also appears to be a lessening of error gravity from deletion to substitution to epenthesis, and this may have methodological implications for a comparative study such as the present one.