Chinese-Speaking EFL Learners’ Performances of Processing English Consonant Clusters

Fang-Chi Chang
National Chiayi University

Abstract
Locke (1983), studying world languages, found that some have word-initial clusters, some word-final clusters, and others consonant clusters in both word-initial and word-final positions. Considering negative transfer, linguists would claim native speakers of tongues without consonant clusters can have difficulty in phonologically manipulating target items with consonant clusters. Chinese differs from English in syllable structure. Predictably, such syllable complexity in English causes problems for Chinese learners of English. This paper presents errors of processing consonant clusters observed in 93 Chinese-speaking EFL (English as a Foreign Language) junior college freshmen. The findings suggest two sources contributing to errors of processing English consonant clusters made by these learners: first language influence and developmental factors.

Key words: consonant cluster, error, language transfer, developmental factor

INTRODUCTION
Locke’s (1983) study of 104 world languages tabulated 39% with word-initial clusters, 13% final, and 48% with both. Considering negative transfer predicted by the Contrastive Analysis Hypothesis (Lado, 1957), linguists would claim that native speakers of languages without consonant clusters may have difficulty phonologically manipulating target items that contain consonant clusters. One common error reported is that learners tend to resyllabify the target word embedding the cluster in question and make the syllable structure an acceptable one in the first language. Chinese differs from English in their syllable structure: the former permits only one consonant in the onset, the latter as many as three in the onset (e.g., street) and four in the coda (e.g., sixths). As predicted by the contrastive analysis hypothesis, such syllable complexity in English will cause problems for Chinese learners as a result of native language interference.
Studies of consonant cluster performances were conducted on learners of English as the native language (e.g., Bruck, & Treiman, 1990; Hindson, & Byrne, 1997; Lleo & Prinz, 1996; McLeod, Doorn, & Reed, 2001; Treiman, 1991) and a foreign language (Lin, 2001a, 2001b). Studying native English speaking children’s consonant cluster performances, researchers found the youngsters had trouble saying words embedding consonant clusters and categorized several frequent errors. Most common errors observed in speech sounds are cluster reduction, cluster simplification, epenthesis, and coalescence. Lin (2001a, 2001b), examining Chinese EFL learners’ productions, reported errors of these types.

Researchers also found consonant clusters difficult for children to analyze for spelling (Bourassa & Treiman, 2001; Bruck & Treiman, 1990; Stage & Wagner, 1992; Treiman, 1991). Treiman (1991) reported that even though children knew that \( p \) and \( r \) represent /p/ and /r/ and could pronounce /pr/, they could not always represent initial /pr/ with \( p \) and \( r \). Among the identified young learners’ spelling errors of consonant clusters, the most frequent is omitting one element from the cluster. Another error is deleting one element from an initial cluster and spelling the yielded onset in orthographically inaccurate but phonetically plausible ways. Reviewing results of some studies of language processing tasks, Treiman (1989) found that consonant clusters were troublesome for adults as well.

**Errors of Pronouncing Consonant Clusters**

Cluster reduction. This is defined as “deletion of one or more consonants from a target cluster so that only a single consonant occurs at syllable margins” (Grunwell, 1987, p.217). For example, blue is pronounced as [bu].

Cluster Simplification. The error occurs when two elements of a cluster are produced with one or both being produced in a manner not matching the target phoneme (Grunwell, 1987). For example, green is pronounced as [gwin]; bread is pronounced as [bwed].

Epenthesis. This type means insertion of some vowel (frequently a schwa) between cluster elements (Dyson & Paden, 1983), such as drive (/draɪv/) pronounced as [dəraɪv].

Coalescence. Coalescence occurs when the yielded pronunciation contains a new consonant composed of features from the original consonants—e.g., swim pronounced as [fim] because the [+fricative] feature of /s/ co-occurs with the [+labial] feature of /w/, resulting in a labial fricative, [f] (Dyson & Paden, 1983).

**Errors of Spelling Consonant Clusters**

Cluster reduction. Most common error in spelling initial clusters is representing a first consonant of a cluster in a legal manner and omitting the second (Bruck & Treiman, 1990).
In spelling final two-element consonant clusters with nasals (/n/, /m/) or liquids (/r/, /l/) as the first element (i.e. /m, n, l, r/ + C), spellers often omit nasals and liquids (Treiman, Zukowski, & Richmond-Welty, 1995). For example, children tend to spell \textit{went} as \textit{wet} and \textit{belt} as \textit{bet}. Read (1975) explained that while being coarticulated, /n/ in a cluster is very short; moreover, the vowel itself is nasalized. Children seem to perceive /n/ as one of the qualities of the vowel that precedes it rather than as an individual phoneme itself (Bourassa & Treiman, 2001), and therefore spell \textit{went} as \textit{wet} with the vowel carrying /n/ in its vowel quality.

Phonetically possible spelling. In representing the first consonant of a cluster, spellers tend to spell words in orthographically inaccurate but phonetically plausible ways (Treiman & Bourassa, 2000). For example, \textit{trap} may be spelt as \textit{chap} because CH closely resembles the sound of the initial blend \textit{tr}. Treiman (1985) explained that this CH spelling reflects slow release of /t/ in this context and in 1989 posited that though cluster \textit{tr} contains two phonemes, it functions in some ways as a single unit. Thus, even children who can articulate clusters may have trouble accessing individual phonemes and representing them in spelling.

Factors Causing Errors

Analyzing the deleted element in consonant clusters, Ohala (1999) tested a hypothesis of consonant cluster reduction based on the Sonority Theory (Clements, 1990) and proved that children’s cluster reductions are sonority-driven. This theory arranges sounds into a Sonority Hierarchy wherein vowels are the most sonorous elements in a syllable and stops the least (Figure 1). This Sonority Hypothesis states that children reduce clusters in such a way that yielded syllables exhibit an optimal sonority contour. That is to say, “syllable initial clusters reduce to whichever consonant in the cluster creates maximal sonority rise, and syllable final clusters reduce to whichever consonant in the cluster creates a minimal sonority descent” (Ohala, 1999, p.402). It means deleting the less sonorous element in an onset and deleting the more sonorous element in a coda. In Table 1, Ohala presents hypothetical outcomes of cluster reductions based on the Sonority Hypothesis to illustrate these predictions. Ohala’s evaluation further supports the hypothesis and the predictions.

<table>
<thead>
<tr>
<th>Stops</th>
<th>Fricatives</th>
<th>Nasals</th>
<th>Liquids</th>
<th>Glides</th>
<th>Vowels</th>
</tr>
</thead>
<tbody>
<tr>
<td>[t,d]</td>
<td>[s,f]</td>
<td>[m,n]</td>
<td>[l,r]</td>
<td>[w,y]</td>
<td>[a,i]</td>
</tr>
</tbody>
</table>

\textbf{Figure 1. The Sonority Hierarchy}
(adopted from Ohala, 1999, p.400)
METHOD

Subjects

Data from which these 93 EFL learners’ consonant cluster processing was observed were collected in a phonological processing study conducted in 2000. That study investigated 93 five-year junior college freshmen’s phonological processing performances. These subjects, English majors at the college, were junior high school graduates at age fifteen to seventeen. Seventy-five of them had studied English for three years in junior high schools, and the rest of them had learned English for four to seven years at the time they participated in that study.

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Examples</th>
<th>Predicted reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Onsets</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. Stop-liquid</td>
<td>pl-</td>
<td>stop (p)</td>
</tr>
<tr>
<td>b. Stop-glide</td>
<td>tw-</td>
<td>stop (t)</td>
</tr>
<tr>
<td>c. Fricative-liquid</td>
<td>fr-</td>
<td>fricative (f)</td>
</tr>
<tr>
<td>d. Fricative-glide</td>
<td>sw-</td>
<td>fricative (s)</td>
</tr>
<tr>
<td>e. Fricative-nasal</td>
<td>sn-</td>
<td>fricative (s)</td>
</tr>
<tr>
<td>Codas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f. Liquid-stop</td>
<td>-lp</td>
<td>liquid (l)</td>
</tr>
<tr>
<td>g. Liquid-fricative</td>
<td>-rf</td>
<td>liquid (r)</td>
</tr>
<tr>
<td>h. Liquid-nasal</td>
<td>-rn</td>
<td>liquid (r)</td>
</tr>
<tr>
<td>i. Nasal-stop</td>
<td>-mp</td>
<td>nasal (m)</td>
</tr>
<tr>
<td>j. Nasal-fricative</td>
<td>-ns</td>
<td>nasal (n)</td>
</tr>
<tr>
<td>k. Fricative-stop</td>
<td>-st</td>
<td>fricative (s)</td>
</tr>
</tbody>
</table>

Data Collection

A study conducted in 2000 fathomed 97 subjects’ capacity for phonological manipulate English spoken and written stimuli, and the data of 93 subjects were valid. Six phonological processing tasks conducted were: syllable deletion, phoneme segmentation, phoneme deletion, phoneme recognition, pseudo-word reading, and spelling dictation. All the subjects worked individually with the researcher on the first five tasks, with the whole procedure recorded. In the first four tasks, each stimulus word was repeated until the subject said it correctly, then instruction for further processing was given. The fifth task, pseudo-word reading, was a read-aloud task, and each subject worked on decoding aloud himself/herself. The sixth, spelling dictation, required subjects to represent spoken words with letters. This task was conducted in a class setting, and hence subjects’ repeating each stimulus for confirmation was
Chinese-Speaking EFL Learners’ Performances

not required.

Since spoken stimuli were given orally by the researcher, the researcher’s pronunciation was considered as one of the factors. Before conducting the tasks, the researcher recoded her productions of all the stimulus items. The tape was given to an English native speaker by a third party for pronunciation evaluation. The evaluator, holding a doctorate in linguistics, was a university professor unacquainted with either the speaker on the tape or with the researcher. He was informed of the purpose of the evaluation; written feedback was “The pronunciation is clear and correct. The nasal sound of the speaker’s /n/ is too strong.”

In all, 127 vocal and 20 written stimuli were used, 93 responses collected for each. From each subject, 147 responses were collected—i.e., a total of 13,671 responses (147 x 93) were recorded. As to test reliability and construct validity, tests had moderate internal consistency, with $\alpha$ around .75, and were highly interrelated. That indicated that the tasks as tests were tapping a similar construct and thus lending construct validity to the concept of phonological processing. This study was not designed to investigate consonant cluster processing, but the data collected revealed valuable information on consonant cluster processing.

FINDINGS AND DISCUSSION

The commonest errors of manipulating consonant clusters observed on native English speaking children occur in their pronunciation and spelling. In the current study, some errors occurred in reading aloud written words containing consonant clusters (pronunciation), and some in spelling oral words with such clusters (spelling). Most errors of processing consonant clusters occurred while subjects processed segments of spoken words containing clusters (processing speech segments). Errors in “pronunciation” and those in “processing speech segments” did not seem to be the same type in nature, yet it was intriguing to see the similarity between these two types. Errors of processing consonant clusters observed in these EFL learners are 1) omitting a second element of an initial two-element consonant cluster or third element of an initial three-element consonant cluster, 2) treating a consonant cluster as a single unit, 3) inserting a vowel between the elements of a consonant cluster, 4) identifying a phonetically similar sound as the sound of a target initial consonant cluster, 5) representing a target initial consonant cluster with phonetic spelling, and 6) omitting nasals and liquids at the first position of a final two-element consonant cluster in spelling.

**Omitting One Element of a Cluster**

In segmenting a spoken word, instead of segmenting the initial two-element cluster into
individual phonemes, subjects gave the first element and omitted the second. In segmenting a three-element cluster, they gave the first two and omitted the third. For example, speech code of grow, blue, train, and drive respectively was segmented into /g/-/o/ (68 subjects, 73.2%), /b/-/u/ (45 subjects, 48.4%), /t/-/e/-/n/ (10 subject, 10.7%), and /d/-/aI/-/v/ (56 subjects, 60.2%). The /r/ of speech codes street and spring was omitted, and the subjects’ answers were /st/-/it/, /s/-/t/-/it/, /sp/-/t/-/n/, etc. No deletion of the first element was observed on processing both two- and three-element clusters. Deletion of second elements of three-element clusters, though rare, did exist. Eight subjects read aloud the pseudo-words spiech as /slI∫/, /slaI∫/, etc., omitting the second consonant.

One thing to be noted again is that each stimulus item was repeated until a subject said it correctly before s/he processed it. Cluster reduction did not occur while a subject pronounced the stimulus item. However, while segmenting speech codes, they omitted a certain element. Memory constraints were first used to explain this: while segmenting, they held the stimulus in mind, segmented the first phoneme, kept the remaining part, said aloud the first phoneme, segmented the second phoneme, kept the remainder, said aloud the second phoneme… This process continued till the task was done. Yet if memory constraints were to explain the failure, the last phoneme(s) of the speech code of the word in question should be the one(s) omitted. That was not the case. They completed segmenting of final phonemes while omitting one element of the cluster.

**Treating A Cluster as A Unit**

While segmenting phonemes, some subjects, instead of segmenting an initial cluster into individual phonemes, treated the whole cluster as one phoneme. For example, a speech code of train was segmented into /tr/-/e/-/n/ (30 subjects, 32.3%) and /tr/-/en/ (8 subjects, 8.6%). Segmenting the speech code of street, one subject (1.1%) responded with /str/-/it/, treating /str/ as a unit. Nine subjects giving /tr/-/en/ and /str/-/it/ actually segmented the speech codes into onset and rhyme. In the deleting task, while deleting one element from a two-element cluster, some subjects failed to do so. They either kept the entire cluster and repeated the original stimulus, or deleted the whole cluster and said aloud the yielded part of the stimulus. For example, being required to delete f from frock to yield rock, 19 subjects (20.4%) deleted the cluster fr and gave /rak/. Being required to delete n from went to yield wet, 11 subjects (11.9%) kept the cluster and gave went. Three subjects (3.2%) deleted the whole cluster nt and gave /we/. Being required to delete r from frog to yield fog, 18 subjects (19.4%) kept the cluster and gave frog as their response; this error was observed in recognizing an initial phoneme, too. Some subjects identified /tr/, /gr/ and /dr/ as the initial phoneme of tree (26 subjects, 28.0%), grass (5 subjects, 5.4%) and drink (8 subjects, 8.6%) respectively.
Inserting a Vowel in the Cluster

Epenthesis was observed in pseudo-word reading: subjects inserted some vowel between two elements of a two-element consonant cluster. Of interest here: a schwa was not their only choice. Other inserted vowels included /e/, /aI/, /i/, etc. Some relation was found between the inserted vowel sound and the vowel sound subjects gave to vowel letter(s) embedded in the stimulus item. For example, pronouncing *spiech* as /spIItʃ/, subjects inserted /I/ between /p/ and /l/ and meanwhile pronounced the digraph *ie* as /I/. Pronouncing *spiech* as /spalIʃ/, subjects inserted /aI/ and pronounced *ie* also as /aI/, an intriguing find.

Epenthesis on initial three-element was fascinating, too: some subjects syllabified initial three consonant clusters as CC.C (e.g., pronouncing *spiech* as /spIItʃ/) or C.CC (pronouncing *stroock* as /strokt/) with a vowel inserted. No similar findings were reported on the phenomenon of epenthesis in initial three-element consonant clusters. A couple of studies on processing medial consonants were found; let us see if these shed some light on epenthesis in initial three-element clusters.

Studying children’s articulation of medial consonant, Stockman & Stephenson (1981) found that children performed better at correctly reproducing medial clusters that conform to word initial clusters relative to those which do not. They posited that those children seemed able to apply knowledge of rules governing permissible sequences that begin words in the language to unclustering medial cluster elements into the coda of a previous syllable and the onset of the following syllable. On resyllabifying medial three consonant clusters in Icelandic words, Kress (1937, cited in Berg, 2001) posited that both CC.C and C.CC are possible, depending on the phonological nature of the cluster at issue. Phonotactic possibilities explain these two findings—i.e., phonotactic possibilities of word margins delimit the phonotactic possibilities of syllable margins and thereby determine syllabification.

With these EFL subjects, influence of phonotactic possibilities in Chinese on processing English consonant clusters was observed. What was reported by Stockman and Stephenson and suggested by Kress was resyllabification on medial consonant clusters. These subjects’ unclustering initial CCC structures into either CC.C or C.CC probably means their previous English linguistic experience should be considered. Three-element clusters are more complex than two-element clusters, even to native children. Some subjects might have mastered two-element clusters but not yet three-element clusters. Furthermore, some subjects’ previous experience of encountering (seeing, saying, processing) words embedding certain kinds of cluster combination unconsciously approved the first two-element cluster and made them resyllabify CCC into CC.C, while other subjects’ previous experience approved the last two-element cluster and made them resyllabify CCC into C.CC. Unfortunately, there were not enough initial three-element stimulus clusters for investigating this assumption.
Phonetically Similar Pronunciation

Replacing a target segment with a phonetically similar segment was observed in the first five tasks. In segmenting the speech code of train (train→/t/-/r/-/e/-/n/), some subjects kept tr as a unit and pronounced it as /ts/ (7 subjects, 7.5%) and /t∫/ (32 subject, 34.4%). The dr of drive was also kept as a unit and pronounced as /ts/ (6 subjects, 6.5%) and /dз/ (40 subjects, 43.0%). Some sample answers to segmenting the speech code of train and drive are /ts/-/en/, /t∫/-/en/, /ts/-/aI/-/v/, /dз/-/aIv/, etc. In the phoneme identification task, the initial phoneme of tree was identified as /ts/ (10 subjects, 10.7%) and /t∫/ (11 subjects, 11.8%), and that of grass and drink was identified as /dз/ by 4 subjects (4.3%) and 43 subjects (46.2%) respectively. Pseudo-word dray was read aloud as /dзe/ by 3 subjects (3.2%).

Phonetic Spelling

Phonetic spelling should result from identifying phonetically similar sounds as the target sound. In the spelling dictation task, the pseudo-words /dredз/ was spelt as gage, jage, and chage; /traIdз/ was spelt as chige, jige, thige, and shige; /kred/ was spelt as cwade and kweade. Identifying phonetically similar sounds resulted in all these unconventional spellings. As reported, some phonetically similar sounds were given by some subjects; a further probe identified phonetically similar sounds and phonetic spellings as given by the same subjects.

Omitting Nasals and Liquids

While representing spoken pseudo-words with letters in a spelling dictation task, though not many, some subjects omitted nasals and liquids, just like native English-speaking children did. They spelt /fεnt/ (an expected spelling is fent) as fet (7 subjects, 7.5%), /pεld/ (peld) as ped (8 subject, 8.6%) and /fImp/ (fimp) as fip (2 subjects, 2.2%). In their study, Treiman et al. (1995) found children’s representations of CVCC syllables reflecting phonetic input. They explained when a postvocalic segment is very short (as with nasals) or lacks a consonantal articulation (in the case of /l/ and /r/) children consider the syllable to contain only three phonemes. They identify C1VC2C3 as C1VC3 and represent the syllable structure with three corresponding letters. What was observed in these EFL learners tallies with their finding.

General Discussion

Six error types are characterized in the study findings. The first, omitting the second of an initial two-element cluster (CCVC) to make a CVC syllable structure, is to make the complex syllable structure more native-like. The second error, treating a cluster as a single unit, is to turn C1C2VC into C3VC with C3 covering C1C2. The third, epenthesis, inserting a vowel between CCC to make either a CVCC structure or a CCVC structure, is still a trial to simplify the even more complicated syllable structure to make it more native like. All these reveal the interference from the Chinese language, which is negative transfer.
A fourth type error, identifying a phonetically similar sound as the target sound, reveals subjects’ awareness of phonetic features (e.g., identifying /tʃ/ as the initial cluster of /tr/). Moreover, a few subjects who identified phonetically similar sounds gave phonetic spellings. Four pronounced initial tr- and dr- as /tʃ/, identified /tʃ/ as the initial phoneme of tree, and used CH to represent the initial dr- (e.g., spell speech code of /dredʒ/ as chage) and tr- (e.g., spell speech code of /trædʒ/ as chige). Another four subjects pronounced initial tr- and dr- as /ts/, identified /ts/ as the initial phoneme of tree, and represented /tr/ as SH and TH. Spelling of ts- was not observed, however. Other phonetic spellings are G, J for /dr/, J for /tr/, KW, CW for /kr/. Difficulty in phonemic awareness seems to pose corresponding obstacles in spelling.

Reported by Read (1975) and Treiman (1985), some beginning spellers, being aware of phonetic features, represent low-level phonetic features in their unconventional spelling. Print exposure will gradually familiarize children with the standard writing system, and they will eventually produce spelling in a conventional manner. Teenage subjects in this study had learned English for three to seven years and were mature in cognition and in the mechanisms of listening and speaking. Still, some errors they made do not seem to result from interference of their native tongue. These errors are segmenting a speech code into onset and rime rather than individual segments as required, giving phonetically similar sounds for the target sounds, representing consonant clusters with phonetically similar spellings, and omitting nasals and liquids. These errors are reported in native English-speaking children as being influenced by some developmental factors. These subjects’ nontransfer errors thus indicate that at that time the subjects might be in the middle of some developmental stages to reach mastery of English consonant clusters.

Is the cluster reduction observed in this study sonority-driven as Ohala attested? Table 2 presents the consonant clusters embedded in the stimuli, the predicted reductions based on the Sonority Hypothesis, and the element being deleted by these EFL subjects. Reduction at

<table>
<thead>
<tr>
<th>Table 2. Element to be Deleted Versus Element Deleted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulus Clusters</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Onsets</td>
</tr>
<tr>
<td>a. Stop-liquid</td>
</tr>
<tr>
<td>b. Stop-glide</td>
</tr>
<tr>
<td>c. Stop-liquid</td>
</tr>
<tr>
<td>d. Stop-liquid</td>
</tr>
<tr>
<td>e. Fricative-stop-liquid</td>
</tr>
<tr>
<td>f. Fricative-liquid</td>
</tr>
<tr>
<td>g. Stop-liquid</td>
</tr>
</tbody>
</table>
onsets is not sonority-driven, while reduction of codas is. The second element of two-element clusters and third of three-element clusters are always the ones deleted, even though there are a few cases of omitting a second element of three-element clusters. Ohala did not predict reduction on three-element clusters. No reports on deletion of three-element clusters can be referred to. Eight subjects deleted the second element /p/ in reading loud pseudo-word *spliech*. The stop /p/ is least sonorous in the Sonority Hierarchy scale. Deleting /p/ seems reasonable but also unexplainable.

Results of Bruck and Treiman’s study (1990) confirmed that deleting an entire cluster is relatively natural for children, but when they must drop a single phoneme from a cluster, it is easier to delete only the second rather than the first phoneme. Locke (1983) posited that if there is a glide or liquid present in a cluster, it typically will be the second member, and it is usually the one deleted by children. In Ohala’s study (1999), she found that children’s cluster reduction was sonority driven when processing stimulus clusters which were similar to native clusters. However, when they encountered unfamiliar clusters, they applied different rules of reduction not in accordance with predictions of Sonority Hypothesis. Bruck and Treiman’s, Locke’s, and Ohala’s notions probably contribute to explaining these EFL subjects’ unpredicted reduction.

**CONCLUSION**

This study was not especially on consonant clusters; neither did it examine the subjects’ pronunciation of consonant clusters. Errors of processing consonant clusters occurred when subjects were manipulating phonemes of consonant clusters in phonological processing tasks. Though most errors were made by only a few subjects, some were made by the majority. The findings suggest two sources contributing to the errors of processing English consonant clusters by Chinese-speaking EFL learners: native language influence and developmental factors. The findings support the claim of the native language interference proposed by the Contrastive Analysis Hypothesis and that of developmental factors proposed by researchers.
Hecht and Mulford (1982) stated that “acquisition of a given language by second language learners closely parallels its acquisition by first language learners” (p.313). Moreover, the finding that onset deletion is not sonority-driven hints other factors besides native language interference and developmental factors involved in foreign language learning. These teenage EFL learners’ diverse linguistic experience, cognition style, strategies accumulated, etc. may be interwoven in affecting their unconscious application of the rules in reduction.

REFERENCES


