MANDARIN LOANWORD PHONOLOGY AND OPTIMALITY THEORY: EVIDENCE FROM TRANSLITERATED AMERICAN STATE NAMES AND TYPHOON NAMES

Hugo Ling-yu Guo

Graduate Program in Linguistics
National Chengchi University, Taipei, Taiwan, ROC
g5555009@grad.cc.nccu.edu.tw

ABSTRACT

This paper aims to examine how the consonant clusters and illicit codas are modified in Mandarin loanwords transliterated from English, and to argue that no rules need to be involved and that a purely constraint-based approach—within the framework of Optimality Theory—can explain the data.

The data collected from transliterated American state names and state names display the onset-coda inconsistency in Mandarin loanwords. All the onset consonant clusters in the data are faithfully parsed into Mandarin syllables, with inserting vowels to shun consonant cluster. The coda clusters and illicit codas are generally parsed. However, the coda liquids may be parsed in some cases but unparsed in others. The preference of insertion in Mandarin loanwords can be explained by the interaction between the constraints--*COMPLEX, CODACON, MAX-IO, and DEP-IO. The distinctive behavior of coda liquids can be accounted for by the effect of MINWD.

1. INTRODUCTION

Mandarin transliterated loanwords borrowed from English are prevalent, such as Maryland [mɛ́.ɹɪ.lənd] → *Mǎlǐlán* [ma.li.lan][1]. Since Mandarin and English have different syllable structures—the latter allows onset and coda clusters and all almost consonants to be codas while the former does not, the syllable structures of the loanwords must be modified or converted while they are borrowed into Mandarin.

Three repair strategies—featural change, epenthesis, and deletion—are generally found to operate on these Mandarin loanwords. A question arises: What is the motivation for these strategies? In the previous rule-based studies (Yin, 1984; Chang, 1996: 10-17), several arbitrary rules of deletion and epenthesis are proposed under the condition of context without phonological motivation. These rule-based analyses, as will be presented in section 3.3, also fail to show the connection of related facts.

Different from the rule-based perspective, this study will investigate how the consonant clusters and illicit codas are dealt with in transliterated American state names and typhoon names in Mandarin in a constraint-based perspective.[2] It will be argued that the repair strategies on the syllable structures are triggered by high-ranked well-formedness constraints. These constraints interact, in order to increase the well-formedness of the surface forms of the loanwords. It will also be argued that the epenthesis and the disyllabicity effect found in these loanwords result from universal constraints that are present in all grammars, but are masked by the effects of higher-ranked constraints in Mandarin Chinese. The questions addressed are: What constraints are involved in generating the candidate set in Mandarin loanwords? How do these constraints interact to leave one candidate as the optimal output to surface? In what ways is this constraint-based analysis better than the rule-based one in explaining this issue?

---

[1] The Pinyin romanization is used in this study. The pronunciation is transcribed in IPA. The symbol "." indicates syllabification.

[2] The term "illicit coda" in this study means the coda in (C)VC. Take, Beth [bɛθ], for example. The [θ] is not allowed to be a coda in Mandarin and is termed as "illicit coda". The coda cluster, such as [bz] in Babs [bæbz], is categorized as a consonant cluster but not an illicit coda in this study.

In the next section, the sound inventories and syllable structures of Mandarin and English will be briefly presented and compared. Section 3 presents the data collected in this study and reviews the previous rule-based studies. Section 4 introduces the theoretical background of this study—Optimality Theory (henceforth, OT), and illuminates how constraints interact to generate the surface forms of Mandarin loanwords, followed by conclusion in section 5.

2. THE SOUND INVENTORY AND SYLLABLE STRUCTURE OF MANDARIN AND ENGLISH

The Mandarin and English vowels are tabulated in tables 1 and 2 while the Mandarin and English consonants are tabulated in tables 3 and 4 respectively.[3]

Table 1. Mandarin Vowels

	Front	Central	Back
High	i ɪ y		u ʅ
Mid	e ɛ	ɚ	o ɔ ɤ
Low	a		

Table 2. English Vowels

	Front	Central	Back
High	i ɪ		u ʊ
Mid	e ɛ	ə ɚ ʌ ɝ	o ɔ
Low	æ		ɑ

Table 3. Mandarin Consonants[4]

		Bilabial	Labio-dental	Alveolar	Retroflex	Alveopalatal	Velar
Stop	-asp	p		t			k
	+asp	ph		th			kh
Affricate	-asp			ts	tʂ	tɕ	
	+asp			tsh	tʂh	tɕh	
Nasal		m		n			ŋ
Glide						j	w
Lateral				l			
Fricative	-vd		f	s	ʂ	ɕ	x
	+vd				ʐ		

(where asp=aspiration, vd=voiced)

Table 4. English Consonants

		Bilabial	Labio-dental	Inter-dental	Alveolar	Palatal	Velar	Glottal
Stop	-vd	p			t		k	
	+vd	b			d		g	
Fricative	-vd		f	θ	s	ʃ		h
	+vd		v	ð	z	ʒ		
Affricate	-vd					tʃ		
	+vd					dʒ		
Nasal		m			n		ŋ	
Liquid	lateral				l			
	retroflex				r			
Glide		w				j		

The syllable structure of Mandarin is (C)(G)V(N) (where G = glide, N = nasal).[5] Nucleus is

[3] This study, basically following Ho (1996: 40) and Tse (1992: 80), adopts IPA symbols as the transcription system for mandarin loanwords and their equivalents in American English. For details, see appendix 1.

[4] The symbols "C+h" represent not a complex consonant but the aspiration of the C.

[5] The V may be a monophthong or a diphthong in Mandarin.

obligatory while onset, prenuclear glide, and coda are optional. The coda position only allows [n] or [ŋ] in Mandarin, such as [khàn] "see" and [fə̄ŋ] "wind". Syllable types include V, GV, VN, GVN, CV, CGV, CGVN. Except CG-combination, there is no consonant cluster in Mandarin syllables.

The syllable structure of English is (C)(C)(C)V(C)(C)(C). Nucleus is obligatory while onset, and coda are optional. The onset and coda position both maximally allow triconsonantal clusters. When there is only one coda consonant, English allows all consonants but /h/ in repertoire to be the coda.

Two major differences are found between Mandarin and English: first, while Mandarin allows only CG-clusters on onset position within syllables, English allows bi-or tri-consonantal cluster, either on onset or coda position. Second, as for the (C)VC syllable type, Mandarin allows only [n] or [ŋ] to occupy to coda position, whereas English permits all consonants except /h/ to do so. Since the syllable structure of Mandarin is more restricted than that of English, when English words are borrowed into Mandarin, the modification in syllable structure become necessary. The focal concern of this study will fall on how the CC- and CCC-clusters and the illicit codas are repaired or modified while the American state names and typhoon names are transliterated into Mandarin, such as Florida [flɔ́.rə.də], Vermont [vɚ.mánt], Babs [bæbz], and Beth [bɛθ].[6]

3. CORPUS OF DATA AND LITERATURE REVIEW

3.1 Data Collection and Transcription

The data in this study comprise of transliterated American state names and typhoon names (see Appendix 2 and 3). These transliterated names are adopted mainly because they are more unified than other proper names. The unification results form the reason that the transliteration of American state names is based on the textbook of Senior High School Geography and that of typhoon names is available from the website of Central Weather Bureau in Taiwan.

The pronunciation of these Mandarin loanwords and their equivalents in American English are transcribed in IPA symbols (for details, see Appendix 1). Examples are given in (5).

(5) a. Alabama	[æ.lə.bǽ.mə]	→	ālābāmā	[a.la.pa.ma]
b. Maryland	[mɛ́.rɪ.lənd]	→	mǎlǐlán	[ma.li.lan]
c. Abel	[é.bəl]	→	àibó	[aɪ.pɔ]
d. Paul	[pɔl]	→	bǎolúo	[paʊ.lwɔ]

Since the tone of each Mandarin syllable is not the major concern, it will be ignored in the transcription. After the pronunciation is transcribed, the syllabification of each word ensues, followed by syllable-structure comparison between these Mandarin loanwords and their equivalents in English.

3.2 Data Analysis

The goal of this study is to investigate how the onset/coda clusters and illicit codas are dealt with in Mandarin loanwords. In the database, no tri-consonantal clusters are found whereas the di-consonantal clusters and illicit codas are prevalent. As for the di-consonantal clusters in English, Mandarin mostly adopts two strategies to modify them—insert vowels to syllabify the consonant(s) or delete consonant(s), as (6) illustrates.

(6) a. Florida	[flɔ́.rə.də]	→	fóluóľidá	[fɔ.lwɔ.li.ta]
b. Babs	[bæbz]	→	bābìsī	[pa.pi.sɪ]
c. Maryland	[mɛ́.rɪ.lənd]	→	mǎlǐlán	[ma.li.lan]
d. Bart	[bart]	→	bātè	[pa.thɤ]

Examples (6a) and (6b) show that the onset and coda clusters in English are parsed faithfully into Mandarin, with the insertion of vowels to satisfy the syllable structure constraint. Examples (6c) and (6d) reveal that deletion may also be one strategy to deal with the consonant cluster.

[6] The featural changes of segments, such as [k]→[tɕ] in [kæ.lə.fɔ̃r.njə]→[tɕja.li.fu.ni.ja], are not the focal concern and will not be discussed detailedly.

The illicit codas also trigger the repair strategies of epenthesis or deletion, demonstrated in (7).

(7) a. Michigan [mí.ʃɪ.gən] → mìxīgēn [mi.ɕi.kən]

b. Texas [ték.səs] → dékèsàsī [tɤ.khɤ.sa.sɪ]

c. Pennsylvania [pɛn.səl.vé.njə] → bīnxīfǎníyǎ [pin.ɕi.fa.ni.ja]

The [n] in (7a) is a licit coda and thus parsed faithfully into Mandarin without conversion. Examples (7b) and (7c) represent the prohibited codas in Mandarin. Different strategies are employed: (7b) inserts vowels after the illegal codas [k] and [s] to syllabify them whereas (7c) deletes the banned coda [l] when this state name is transliterated into Mandarin.

Furthermore, the coda may also undergo featural changes to conform to Mandarin syllable structure, such as Hampshire [hǽmp.ʃɪr]→ *hǎnpǔxià* [xan.phu.ɕja]. The strategy is quite limited to nasal and liquid codas, especially [m], [l], and [r]. In most cases, epenthesis and deletion are still preferred.

A question arises here: are the strategies of epenthesis and deletion predictable? As the database reveals, all the onset clusters in English are faithfully parsed into Mandarin syllables, with vowel inserted to shun consonant clusters. No deletion is found. The coda clusters and illicit codas, however, display inconsistent phenomena—the strategies of insertion, deletion, and featural changes are all used. Then, what is the factor causing this discrepancy between onset and coda behavior? Before this question is addressed, the behavior of illicit codas and coda clusters must be inspected. Tables 5 and 6 reveal the proportion of parsing and deletion of each class in illicit codas.[7] The behavior of liquids, quite crucial to the present study, will be tabulated individually.

(8) **Table 5. The Modification of Stop, Nasal, and Fricative Codas**

		Stop		Nasal		Fricative	
		N	%	N	%	N	%
State	Parse	3	60	13	100	5	100
	Deletion	2	40	0	0	0	0
	Total	5	100	13	100	5	100
Typhoon	Parse	10	90.9	18	100	7	100
	Deletion	1	9.1	0	0	0	0
	Total	11	100	18	100	7	100

(N=number, %=percentage)

Table 6. The Modification of Liquid Codas

		Monosyllabic		Polysyllabic	
		N	%	N	%
State	Parse	0	0	0	0
	Deletion	0	0	7	100
	Total	0	0	7	100
Typhoon	Parse	5	71.4	5	27.5
	Deletion	2	28.6	8	62.5
	Total	7	100	13	100

Table 5 displays that all the illicit nasal and fricative codas of the syllables within American state names and typhoon names are parsed faithfully into Mandarin by inserting vowels or changing the segmental features. The illegal stop codas tend to be parsed into syllable, however, with few exceptions, such as David [dé.vɪd]→*dàwèi* [ta.we] and Mexico [mék.sɪ.kɔ]→ *mòxīgē* [mɔ.ɕi.kɤ]. These exceptions may not

[7]The term "parse" in this study means that the segment in the English names has its correspondent in Mandarin. Two strategies are available to parse the segments—featural changes or epenthesis, such as Dale [del]→ dàiér [taɪ.ɚ] and Cam [kæm]→ kǎimǔ [khaɪmu]

result from phonological factors but other governing conventions.[8] Generally speaking, if the syllables of these names end with a stop, nasal, or fricative coda, the coda will be parsed faithfully while transliterated, undergoing vowel insertion or featural changes of segments.

The liquid coda displays a different phenomena, as table 6 reveals. As the database shows, within a polysyllabic English word, the liquid coda tends to be deleted while transliterated, such as Pennsylvania [pɛn.səl̠.vé.njə]→ *bīxīfǎnīyǎ* [pin.ɕi.fa.nja]. Contrastively, within a monosyllabic English word, the liquid coda tends to be parsed, such as Dale [del]→*dàiér* [taɪ.ɚ]. The two unparsed liquid codas within monosyllabic words are the [r]'s in Bart [bart]→ *bātè* [pa.thɣ] and Mort [mɔrt]→*mótè* [mɔthɣ]. It is now postulated that the liquid coda tends to be deleted. Its being parsed within monosyllabic words is because a disyllabic word is preferred in Mandarin (Broselow et al, 1997). However, when the monosyllabic words have codas other than a liquid, these codas rather than the liquid will be parsed to meet the requirement of disyllabicity.

Why do the liquid codas tend to be unparsed? According to Fay and Culter (1977), the liquids, acoustically, have vowel-like formants and cannot stand out from the surrounding vowels saliently. Harris (1994: 230) indicates that while the onset [r] is a rough variant of /r/, the coda [r] is a smooth variant. The coda [r] often displays non-rhotic pattern—coda [r] tends to be deleted before a consonant or a pause. Yip (1993: 268) also suggests that the liquids are less salient than other consonants and that "this lack of salience renders them relatively vulnerable to deletion." These statements imply that the liquid coda is smooth, vowel-like, and not salient. Because of this lack of salience, the liquid coda is liable to be deleted, or is not to be perceived by hearers significantly. The notion of salience can explain why the liquid onset is parsed faithfully while the liquid coda tends to be deleted in most cases.

The strategies of epenthesis and deletion are also employed to deal with the coda clusters. As for the typhoon names, the coda clusters—ks, -bz, -tz, -ŋk, -rt—are found.[9] Except the [r] in rt-cluster, all the consonant clusters are parsed into syllables while transliterated into Mandarin. The effect of salience also plays a dominant role in determining which segment should be parsed or unparsed. As for the state names, the coda clusters—n(d), -(ts), -(r)k,-nt, -mp—are available (Parenthesis marks the unparsed segments).[10] The notion of salience can also be referred to illuminate the deletion of [d] and [r] in these Mandarin loanwords. The deletion of -ts in Massachusetts [mæ.sə.tʃʊ́.sɛts]→*másàzhūsài* [ma.sa.tʂu.saɪ] deviates from the present prediction that the deletion of a segment is due to its lack of salience, since -ts is endowed with high-intensity noise. This deviation may be explained by Chang's (1996: 18) finding that Mandarin, if possible, tends not to reinterpret the English words with more than four syllables. Thus, the effect of salience is still stable in predicting which coda segment should be parsed or unparsed without the interfering non-phonological factors.

In sum, based on transliterated American state names and typhoon names, this study finds the onset-coda inconsistency in Mandarin loanword phonology. The onset clusters of these English names are faithfully parsed while they are transliterated into Mandarin. The coda cluster or illicit codas, however, are not always faithfully parsed—the salient segments are parsed into Mandarin syllables with some modification, such as fricative and nasal codas, whereas the unsalient segments tend to be unparsed, such as liquid codas. This onset-coda inconsistency results from the factor of salience—because of the lack of salience, some segments in coda positions are overlooked and unparsed. Furthermore, to avoid the inappropriate syllable structures, the strategy of epenthesis rather than deletion is preferred in Mandarin loanword phonology. This preference can be accounted for merely by constraint interaction and satisfaction. Before the constraint-based analysis is presented, the rule-based analysis will be reviewed in section 3.3 first.

---

[8] The only exception in typhoon names is David. The Mandarin transliteration of this name may follow the fixed translation, which is based on the Hebrew pronunciation of this name instead of the English pronunciation. Similarly, the unparsed coda [k] in Mexico may also be explained by the factor of fixed translation. While Mexico first came into Mandarin, it may be borrowed from Spanish, with its pronunciation to be [mɛ́xɪgɔ]. It was thus transliterated as *mòxīgē* [mɔ.ɕi.kɣ]in Mandarin and became a convention since then.

[9] See Appendix 2: 1-6.

[10] See Appendix 2: 7-11.

3.3 A Rule-based Analysis

Chang (1996) proposes a rule-based analysis to illuminate the phonological structure of transliterated English loanwords in Mandarin. Several rules of insertion and deletion are framed to account for how the English syllable-initial or -final consonant clusters are modified to comply with Mandarin syllable structures, such as inserting [i] between syllable-initial [s] and stops or nasals, [o] or [i] between Consonant + liquid, and deleting postvocalic [r] or word-final consonants (Chang, 1996: 18).

The first problem this rule-based analysis encounters is: where do the rules of epenthesis and deletion originate? In terms of Standard Theory, these rules should exist in Mandarin or be imported from English. However, native Mandarin forms provide no evidence for underlying representations with consonant clusters (henceforth C-cluster) and illicit codas, so there are no alternations providing evidence for a rule to epenthesize vowels after each consonant of C-clusters or illicit codas. These so-called rules do not originate from Mandarin and English, and thus there is room for doubt about the status of these rules.

Second, some rules cannot apply to all the members of identical context in the database. In the database of the present study, the syllable-final [l] is unparsed in most cases. In the rule-based perspective, a rule of syllable-final [l]-deletion would be proposed. This rule, however, fails to explain why the syllable-final [l] can be parsed in some other cases.

The problems mentioned above can be solved reasonably in terms of constraints within the framework of Optimality Theory, which describes a grammar as a set of universal ranked constraints. The surface forms of these transliterated loanwords are the consequence of constraint interaction and satisfaction. No rules at all are needed. Next section will briefly introduce the theoretical background of Optimality Theory and demonstrate its application to Mandarin loanword phonology.

4. OPTIMALITY THEORY AND MANDARIN LOANWORD PHONOLOGY

4.1 Theoretical Background

The basic assumption of Optimality Theory (McCarthy and Prince, 1993; Prince and Smolensky, 1993; Archangeli and Langendoen, 1997) is that the grammar of languages consists of a set of ranked violable well-formedness constraints. While the constraints are universal, the ranking of constraints is language specific.

OT mainly deals with the correspondence between input and output. While in standard theory the output is derived form an input via a context-driven rewrite rule, in OT the output is chosen from a set of candidates which are associated with an input (Hung, 1994: 2).

More specifically, given an input, the function called Gen (Generator) will generate a set of possible candidate analyses, based on the universal well-formedness constraints. Then this candidate set is submitted to another function Eval (evaluation). The function Eval, composed of a language-specific ranking of constraints, evaluates all the possible candidates generated by Gen in parallel and selects one which best satisfies or minimally violates the ranking as the optimal output. The notion of minimal violation or best satisfaction needs to be defined in terms of this ranking. Tableau 1 represents one situation of minimal violation.[11]

(9) Tableau 1

$Input_i$	Constraint A	Constraint B	Constraint C
☞Candidate 1			*
Candidate 2		*!	
Candidate 3	*!		

[11] The tableaux have the following convention: (1) The domination order of constraints is shown in left-to-right order. (2) Violation of a constraint is marked by * while satisfaction is indicated by a blank cell. (3) The sign ! signifies a fatal violation, the one that is responsible for a candidate's nonoptimality, whereas the symbol ☞ indicates the optimal output. (4) Shading emphasizes the irrelevance of the constraint to the fate of the nonoptimal candidate.

As tableau 1 shows, constraint A is ranked higher than constraint B, followed by constraint C in Language X. The violation of higher-ranked constraints is fatal. Thus, though all the candidates violate only one constraint respectively, candidate 2 and 3 are eliminated since they violate higher-ranked constraints. Candidate 1 violates the lower-ranked constraint with the minimal penalty and is selected as the optimal output in Language X. This illustrates a key characteristic of OT: simple violation of a constraint is never in self fatal. While constraints are in conflict, the lower-ranked constraint can be toleratedly violated to satisfy a dominant constraint to avoid the fatal violation. Different languages would have different rankings of constraints and thus given an identical input, the optimal candidate will not be the same cross-linguistically. If Constraint C is ranked higher than Constraint B in Language Y, Candidate 2 will be the optimal surface form in Language Y. Tableau 2 displays another situation of minimal violation.

(10) Tableau 2

$Input_j$	Constraint D	Constraint E
☞Candidate 4	*	*
Candidate 5	*	**!

Candidate 4 and 5 cannot satisfy the Constraint D, and thus this constraint fails to decide the optimality of candidates even though it is ranked higher. Constraint E must be consulted. While Candidate 4 violate one time for Constraint E, Candidate 5 violates two times. Thus, while compared to Candidate 5, Candidate 4 violates constraints minimally and becomes optimal. This illuminates another properties of OT: "violation is only fatal while there are other competing constraints that pass the constraint" (McCarthy and Prince, 1993:7).

In sum, no specific rules are needed within the framework of OT since "the candidate analyses, evaluated by the constraint hierarchy, are admitted by very general considerations of structural well-formedness" (McCarthy and Price, 1993: 5). No derivational processes are proposed since the best satisfaction of the candidate set is computed in parallel. Moreover, OT analysis to language mainly focuses on the surface or output structure. The focus on surface forms or outputs makes OT well suited to the description of loanword phonology. The loan language (Mandarin) may introduce underlying representations that are not motivated by itself. But since all the underlying representations will ultimately be forced to conform to the surface constraints in the loan language, the foreign underlying forms (English) will come out looking like the surface forms of the loan language (Mandarin). By means of inspecting the surface forms or outputs, the constraints and their dominancy or ranking in a particular language can be worked out.

4.2 A Constraint-based Analysis to Mandarin Loanword Phonology

The **Basic Syllable Structure Constraints**, as proposed by Prince and Smolensky (1993: 85-8), describe the universally unmarked characteristics of syllable structures. The CV-combination is the most unmarked syllable structure.[12] Based on the CV structure, Prince and Smolensky set constraints for a preferred unmarked syllable structure universally, as (11) states.

(11) a. **ONSET**: Syllables must have onsets.
b. **NOCODA**: Syllables must not have a coda.
c. **NUC** (nucleus): Syllables must have nuclei. [13]
d. ***COMPLEX** (no complex): No consonant cluster is allowed within a syllable (where consonant cluster means [+cons][+cons] . . .)

Since Mandarin does allow codas but the codas are highly restricted to be [n] or [ŋ], the constraint NOCODA may be refined as CODACON, as (12) dictates.

(12) **CODACON**: Syllables must have no coda, except an alveolar or velar nasal.

[12] See Chomsky and Halle (1991: 408).

[13] According to McCarthy and Prince (1993: 104), "if a constraint is never violated, it is essentially part of the basic definition of a structural category and one might as well regard it as part of Gen rather than part of the Eval." NUC is a good example, and need not be ranked.

Furthermore, basic to OT is consistency, that is, there must exist a one-to-one correspondence between the input and output (McCarthy and Prince, 1995b). Any mismatch between the input and output will violate MAX-IO or DEP-IO, which are referred to as FAITHFULNESS constraint family.

(13) a. **MAX-IO (maximum-input/output, first version)**

Every segment in the input must have a correspondent in the output.

b. **DEP-IO** (dependent-input/output)

Every segment in the output must have a correspondent in the input.

DEP-IO requires no insertion since any inserted segment in the output cannot have a correspondent in the input; on the other hand, MAX-IO craves no deletion since any deleted segment in the output will lead to some segments in the input having no correspondent in the output.

As the data show, to avoid C-clusters, epenthesis instead of deletion is triggered; that is, to satisfy *COMPLEX, the constraint DEP-IO rather than MAX-IO will be violated in Mandarin. This means that *COMPLEX is ranked higher than MAX-IO, followed by DEP-IO. If DEP-IO is ranked higher than MAX-IO, deletion of segments will be prevalent, which contradicts the fact. The ranking among CODACON, MAX-IO, and DEP-IO is also crucial. To shun the illicit coda except liquid codas, the strategy of insertion is also preferred, implying that CODACON is more dominant than MAX-IO, followed by DEP-IO. The interaction among these constraints is demonstrated in tableau 3.

(14) Tableau 3 (Nebraska [nə.bræs.kə]→[ne.pu.la.sɪ.tɕja])[14]

Input: nə.bræs.kə	*COMPLEX	CODACON	MAX-IO	DEP-IO
☞a. ne.pu.la.sɪ.tɕja				**
b. ne.la.sɪ.tɕja	*!	*		
c. ne.la.tɕja			*!	
d. ne.las.tɕja		*!		

The ranking between *COMPLEX and CODACON is not crucial in determining the optimal candidate. It means that there is no interaction between these two constraints here. The ranking can be sketched in the following schema, as (15) shows.

(15) *COMPLEX, CODACON >> MAX-IO >> DEP-IO

What is the status of ONSET here? The penalty of violating ONSET seems low since the onsetless syllables are abundant in Mandarin native forms and loanwords. To maintain faithfulness, the constraint ONSET can be violated, tableaux 4 and 5 reveal.

(16) a. Tableau 4 (*ān* "safe")

Input: an	MAX-IO	DEP-IO	ONSET
☞a. an			*
b. jan		*!	

b. Tableau 5 (Alabama [æ.lə.bæ.mə]→[a.la.pa.ma])

Input: æ.lə.bæ.mə	*COMPLEX	CODACON	MAX-IO	DEP-IO	ONSET
☞a. a.la.pa.ma					*
b. ja.la.pa.ma				*!	

Tableaux 4 and 5 elaborate the point that the surface constraints MAX-IO, and DEP-IO are ranked higher than ONSET in Mandarin native forms and loanwords. Counterexamples, however, are still available, such as Arizona [æ.rə.zo.nə]→ yǎlìsāngnà [ja.li.saŋ.na] and Alex [æ.lɪks]→ yǎlìshī [ja.li.ʂɿ]. In these counterexamples, ONSET is more dominant than DEP-IO. But this kind of examples are relatively rare. Besides, there may be some other factors interact with phonological factors to transliterate these loanwords into Mandarin, such as semantic-ambiguity avoiding or character choosing. Since [a.li.saŋ.na] and [a.li.ʂɿ] are also acceptable in transliteration, it is reasonable to say that the ranking between DEP-IO are generally more dominant than ONSET in Mandarin native forms and loanwords but the interfering factors would cause some constraint ranking "to be in flux" in loanword phonology (Broselow et al, 1997:

[14] The solid line between constraints means the ranking is crucial while the dotted line between constraints means the ranking is not crucial.

23). And thus the alternative outputs are available. So far, the general constraint ranking can be summarized in (17).

(17) *COMPLEX, CODACON >> MAX-IO >> DEP-IO >> ONSET

The constraint interaction in (17) explains most cases about how the English consonant clusters and codas are modified or repaired in Mandarin but the behavior of liquid codas is still left unsolved. The constraint ranking above predicts the liquid coda to be parsed everywhere but the data display that the liquid codas tend to be unparsed in polysyllabic words due to its lack of salience. The study follows Yip (1993: 278) to regard that the unsalient segments would be "faintly visible, and thus may be overlooked" by MAX-IO constraint. A final statement of MAX-IO is given below, with slight modification.

(18) **MAX-IO** (final version): Every (salient) of the input has a correspondent in the output.

Under this revised MAX-IO constraint, the unsalient segments tend to be overlooked and unparsed, but would be parsed sometimes. A question arises: when should the unsalient liquids be parsed? As the data reveal, the liquid coda is parsed within monosyllabic words, such Dale, Gil, and Neil but while the liquid coda is followed by another coda, such as Mort and Bart, the liquid remains unparsed. What is the motivation for parsing the unsalient liquids? A universal constraint MINWD can answer this question.

(19) **MINWD**: A lexical word must be disyllabic minimally.

It is the effect of MINWD triggering the unsalient segments to be parsed, implying that MINWD dominates MAX-IO. Since the liquid coda will be parsed to satisfy the effect of MINWD, two strategies—featural changes or epenthesis—can be employed to repair the illicit forms. The liquid codas seem to prefer the former. Since the liquid codas are vowel-like, their distinctive features are similar to those of vowels. The parsed liquid codas tend to converted as [ɚ] in loanwords. This conversion, however, violates another constraint IDENT (F).

(20) **IDENT (F)** (feature identity)

The distinctive features of two corresponding segments must be identical.

The constraint IDENT (F) is violated to avoid violating DEP-IO, suggesting that DEP-IO is more dominant than IDENT (F), as tableau 6 shows.

(21) Tableau 6 (Dale [del]→[taɪ.ɚ], Bart [bart]→[pa.thɣ])

Input: del	CODACON	MINWD	MAX-IO	DEP-IO	IDENT (F)
☞a. taɪ.ɚ					*
b. taɪl	*!	*			
c. taɪ		*!			
d. taɪ.lɣ				*!	
Input: bart					
☞a. pa.thɣ				*	
b. path	*!	*			
c. pa.ɚ			*!		*
d. pa.ɚ.thɣ				*	*!

Two facts are available to challenge the present analysis. First, counterexamples—Paul [pɔl]→ *pǎoluó* [paulwɔ] and Yule [jul]→ *yóulǐ* [joli]—are found. These examples, however, are relatively uncommon in the database. The use of epenthesis instead of featureal may result from some other governing factors such as character choosing or convention of translation. It is therefore concluded that while only phonological factors are considered, the constraint ranking in tableau 6 is still quite valid. If other factors beside phonology has to be involved, the ranking among the constraints can be in flux in loanword phonology.

Second, the above generalization mentions that the constraint MINWD exists in Mandarin and is ranked higher than MAX-IO and DEP-IO. However, native Mandarin does allow monosyllabic lexical words, such as *kàn* "see" and *shǒu* "hand". If the constraint interaction in tableau 6 applies to *kàn*, it would be extended to as *kànè* to satisfy the effect of MINWD, which contradicts the fact. To solve this dilemma caused by the effect of MINWD, this study accepts Yip's point that the constraint FAITHFULNESS, which states "Do not alter underlying forms", may not be identified with the pair of more specific constraints—MAX-IO and DEP-IO.[15] Thus, that the word *kàn* "see" is not extended can be explained by the fact that FAITHFULNESS is ranked higher than MINWD as tableau 7 illustrates.

(22) Tableau 7 (kan "see")

Input: kan	CODACON	FAITHFULNESS	MINWD	MAX-IO	DEP-IO
☞a. kan			*		
b. kanɣ		*!			*

If FAITHFULNESS does not separate from MAX-IO and DEP-IO and rank higher than MINWD, all the Mandarin monosyllabic words have to be extended to satisfy the disyllabicity effect. The evidence from Mandarin phonology also supports Yip's viewpoint that MAX-IO and DEP-IO should be separated from FAITHFULNESS constraint family. So far, the constraint ranking in Mandarin can be sketched in the following schemata.

(23) a. CODACON, *COMPLEX, MINWD>>MAX-IO>>DEP-IO>>ONSET, IDENT(F)
b. FAITHFULNESS>>MINWD

The interaction or ranking among CODACON, *COMPLEX, and FAITHFULNESS is not crucial here. The above constraint rankings can also correctly predict how [m] in monosyllabic English words is converted while these words are borrowed into Mandarin, as tableau 8 shows.

(24) Tableau 8 (Cam [kæm]→ kǎimǔ [khaɪ.mu])

Input: kæm	CODACON	FAITH	MINWD	MAX-IO	DEP-IO	IDENT (F)
☞a. khaɪ.mu		*			*	
b. kham	*!		*	*		
c. khan		*	*!			
d. khaɪ.nu		*			*	*!

In this tableau, it is found that CODACON should dominate FAITHFULNESS, or [kham] will become the optimal output. Thus, based on Mandarin native forms and loanwords the constraint ranking in Mandarin can be modified as (25).

(25) a. CODACON >> FAITHFULNESS >> MINWD >> MAX-IO >>
DEP-IO >> ONSET, IDENT(F)
b. CODACON, *COMPLEX >> MAX-IO

It is these constraints interacting with one another to generate the optimal output in Mandarin.

5. CONCLUSION

Based on the transliterated American state names and typhoon names, this study provides a constraint-based analysis to Mandarin loanword phonology. The modification of syllable structures of the foreign words are necessary since the constraint rankings in American English and in Mandarin are different. The English forms should be filtered or repaired by the constraint rankings in Mandarin. The repair strategies—epenthesis, deletion, and featural changes—are motivated by higher-ranked well-formedness constraints in Mandarin, such as *COMPLEX and CODACON. The theoretical framework of OT also provides a reasonable explanation about why some constraints only bring out their effects in loanword phonology but fail to do so in native phonology. Furthermore, the behavior of liquid coda, which cannot be comprehended by the rule-based analysis, can also be predicted to some extent by means of constraint interaction.

However, there are still some problems in this study. First since this study is based on the written transliterated names, some other factors beside phonological ones may interfere the results, such as

[15] Yip argues that MAX-IO and DEP-IO should be separated from FAITHFULNESS. First, while FAITHFULNESS pays attention to all detectable segments, even the liquid coda, MAX-IO only cares about highly salient segments. Second, there would be a ranking paradox if one tries to combine MAX-IO and DEP-IO with FAITHFULNESS. For details, see Yip (1993: 283-7).

character choosing, semantics, and convention for translation. These interfering factors lead to lots of counterexamples that are hard to explain. Results attained from real-time transliteration are thus needed. Second, OT requires that the constraints must be universal. However, some language-specific constraints seem unavoidable, such as CODACON. Third, vowel and consonant conversion in English-Mandarin transliteration is only slightly touched. The problem about the featural changes of segments is still left unsolved. The correspondence between segments in English and in Mandarin needs more elaboration. It is therefore hoped that further refined studies will be available to solve these problems and verify the results worked out in this study.

REFERENCES

[1] Archangeli, Diana, and Terence Langendoen. 1997. ***Optimality Theory: An Overview.*** Massachusetts: Blackwell.

[2] Broselow, Ellen, Su-I Chen, and Chilin Wang. 1997. "The Emergence of the Unmarked in Second Language Phonology." Manuscript.

[3] Central Weather Bureau. www.cwb.gov.tw/typhon/

[4] Chang, Guo-bin. 1996. "The Phonological Structure of Transliterated English Loan Words in Chinese." Manuscript. National Chengchi University.

[5] Chao, Fengfu et al. Eds. 1982. ***Wen Shin's English Chinese Dictionary of Contemporary English.*** Taipei: Wenshin.

[6] Chomsky, Noam, and Morris Halle. 1991. ***The Sound Pattern of English.*** 2nd printing. Massachusetts: MIT Press.

[7] Fay, David, and Anne Culter. 1977. "Malapropisms and the Structure of the Mental Lexicon." ***Linguistic Inquiry.*** 8 (3): 505-20.

[8] Harris, John. 1994. ***English Sound Structure.*** Massachusetts: Blackwell.

[9] Ho, Da-an. 1996. ***Shengyunxue Zhong De Guannian Yu Fangfa. (The Concept and Method in Phonology.)*** Taipei: Da-an.

[10] Hsu, Hui-chuan. 1994. ***Constraint-Based Phonology and Morphology: A Survey of Languages in China.*** Ph.D. Dissertation. University of California, San Diego.

[11] Hung, Henrietta J. 1994. ***The Rhythmic and Prodosic Organization of Edge Constituents.*** Ph.D. Dissertation. Brandeis University.

[12] Kenstowicz, Michael. 1994. ***Phonology in Generative Grammar.*** Massachusetts: Blackwell.

[13] McCarthy, John, and Alan S. Prince. 1993. "Prosodic Morphology I: Constraint Interaction and Satisfaction." Manuscript.

[14] —. 1995b. "Faithfulness and Reduplicative Identity." ***University of Massachusetts Occasional Papers.*** 18: 249-384.

[15] National Institute of Compilation and Translation. 1997. ***Senior High School Geography.*** Taipei: National Institute of Compilation and Translation.

[16] Paradis, Carole. 1986. "On Constraints and Repair Strategies." ***The Linguistic Review.*** 6 (1): 71-97.

[17] Prince, Alan, and Paul Smolensky. 1993. ***Optimality Theory: Constraint Interaction in Generative Grammar.*** Massachusetts: MIT Press.

[18] Singh, Rajendra. 1984. "Well-formedness Conditions and Phonological Theory." In ***Phonologica 1984.*** 273-86. Cambridge: Cambridge University Press.

[19] Tse, John Kwock-ping. 1992. ***An Introduction to Linguistics.*** Taipei: Sanmin.

[20] Wang, Chilin. 1995. ***The Acquisition of English Word-final Obstruents by Chinese Speakers.*** Ph.D. Dissertation. SUNY at Stony Brook.

[21] Yin, Yuen-mei. 1984. "A Comparative Study of the Interlanguage Phonology of Two Chinese Dialect Groups." Manuscript. University of Texas at Austin.

[22] Yip, Moira. 1993. "Cantonese Loanword Phonology and Optimality Theory." ***Journal of East Asian Linguistics.*** 2 (3): 261-91.

APPENDIX 1

This study, basically following Ho (1996: 40), adopts IPA symbols as the transcription system for Mandarin loanwords and their equivalents in American English. The correspondence between Mandarin phonetic symbols and IPA symbols are listed as follows.

ㄅ：[pə]	ㄆ：[ph]	ㄇ：[mə]	ㄈ：[fə]	ㄉ：[tə]	ㄊ：[thə]
ㄋ：[nə]	ㄌ：[lə]	ㄍ：[kə]	ㄎ：[khə]	ㄏ：[xə]	ㄐ：[tɕi]
ㄑ：[tɕhi]	ㄒ：[ɕi]	ㄓ：[tʂʅ]	ㄔ：[tʂhʅ]	ㄕ：[ʂʅ]	ㄖ：[ʐʅ]
ㄗ：[tsɪ]	ㄘ：[tshɪ]	ㄙ：[sɪ]	ㄧ：[i]	ㄨ：[u]	ㄩ：[y]
ㄚ：[a]	ㄛ：[ɔ]	ㄜ：[ɣ]	ㄝ：[ɛ]	ㄞ：[aɪ]	ㄟ：[e]
ㄠ：[aʊ]	ㄡ：[o]	ㄢ：[an]	ㄣ：[ən]	ㄤ：[aŋ]	ㄥ：[əŋ]
ㄦ：[ɚ]					

The following are some conventions for transcription in this study.

(1) The symbols [ʅ] and [ɪ] represents the vowels following the retroflex consonants as well as their corresponding unretroflex counterpart respectively.

(2) Some modification is also made for the consistency of transcription. ㄛ and ㄝ are transcribed as [ɔ] and [ɛ] respectively instead of [e] and [o]. ㄟ and ㄡ are thus transcribed as[e] and [o] respectively.

(3) While "ㄧ[i]" and "ㄨ[u]" behave as medials, they are transcribed as glides, that is, [j] and [w]. While "ㄧ[i]" and "ㄨ：[u]" behave as endings of diphthongs, they are transcribes as short vowels, that is, [ɪ] and [ʊ] respectively.

APPENDIX 2

1. Alex	[ǽ.lɪks]
亞力士	[ja.li.ʂʅ]
2. Babs	[bæbz]
芭比絲	[pa.pi.sɪ]
3. Fritz	[frɪtz]
福立茲	[fu.li.zɪ]
4. Hank	[hæŋk]
漢克	[xan.kɣ]
5. Mort	[mɔrt]
摩特	[mɔ.thɣ]
6. Rex	[rɛks]
雷克斯	[lɛɪ.khɣ.sɪ]
7. Hampshire	[hǽmp.ʃɪr]
罕普夏	[xan.phu.ɕja]
8. Maryland	[mɛ́.rɪ.lənd]
馬里蘭	[ma.li.lan]
9. Massachusetts	[mæ.sə.tʃú.sɪts]
麻薩諸塞	[ma.sa.tʂu.saɪ]
10. New York	[njú jɔ́rk]
紐約	[njo.yɛ]
11. Vermont	[vɚ.mɑ́nt]
佛蒙特	[fɔ.məŋ.thɣ]