Mandarin Retroflex Suffixation: An OT Account

Ma Qiuwu
Department of Foreign Languages, Beijing Normal University
Foreign Languages Dept, Tianjin University of Technology
maquwu@public.tpt.tj.cn

The retroflex suffixation is one of the interesting phonological phenomena in Beijing Mandarin and has engendered much discussion. But most of the discussions are made from social, historical and experimental points of view. Recently, some of the scholars overseas have applied different theories of Generative Linguistics to analyze it and proposed different explanatory theoretic models for it. In this paper, I will first present the data of Mandarin retroflex suffixation reported by different scholars, then introduce some representative models proposed so far to account for Mandarin retroflex suffixation, and finally make a discussion on it from the OT perspective.

1. Mandarin Retroflex Suffixation

The studies on retroflex suffixation are abundant in the Chinese linguistic literature. Most of them are the (synchronic or diachronic) data descriptions (e.g., Tao & Yin 1957; Chao 1968; Xu 1960; Wang 1963; Lin 1982; Hsueh 1986; Lin & Shen 1995; Li 1986; Lu 1995) or phonetic / perceptional experiments (Wang & He 1985; Bao 1989; Li 1990; Yang 1991). These studies are often different in their data presented. For ease of exposition, I will introduce only some representative studies on it and then, after consulting and comparing several representative empirical studies, present a more general version of the retroflexed forms in Beijing Mandarin.

1.1 Some Empirical Studies

The retroflex suffix 儿 is considered to evolve from "li", "ri" and / or "er" (Chao 1968:46; Hsueh 1986:74[1]; Lin & Wang 1992:165-166) and, as Hsueh (1986:73) claims, it is now added or merged into the preceding syllable, assigning a diminutive meaning to it. "er" as a diminutive suffix probably first appears in no earlier than Ming dynasty. Xu (1999:133) points out that the dimunitive suffix "er" was originally an independent syllable and has been merged into the preceding syllable due to the long time of fast speaking. Morphologically, the retroflex suffix is usually written as an independent morpheme "er" (Lin 1982/1990:62) but, as Tao and Yin (1857:31) noted, some syllables are read with the diminutive retroflex even if the suffix "er" is not written out.

It is widely acknowledged that the suffix "er" is now the only non-syllabic morpheme or, straightforwardly speaking, a segment-affiliated feature in Mandarin. The problem is: how is it combined with or incorporated into the root or preceding syllable in the phonological process of

[1] Hsueh (1986) claims that "ri" is only the historical source of Mandarin retroflex suffix.

Mandarin retroflex suffixation?

Tao and Yin (1957) claim that the retroflex suffixation in Mandarin is simply adding the color of retroflex or rhotacization to the ending of the preceding suffixed syllable or stem. If the suffixed syllable ending is incompatible for rhotacization, it will be lost in the process of its retroflex suffixation. The ways of retroflex suffixation they proposes can be summed up as the following types:

(1) Tao & Yin (1957)
   a. Add the color of retroflex directly to the syllable ending with a, o, e, or u;
   b. Add the retroflexed mid central vowel [☯ʳ] after the singleton high vowel rime i or 𝅥;
   c. Replace the apical vowels with the retroflexed mid central vowel;
   d. Delete the post-nuclear glide from the rime ai, ei, or ui and then for the first two rimes add the color of retroflex to the remaining nuclear vowels a or e and for the last rime ui add the retroflexed mid central vowel [☯ʳ] to the end of it;
   e. For the palatal nasal rimes an, en, in and 𝅥n, first delete the nasal consonant and then, for the first two rimes, add the color of retroflex to the remaining nuclear vowels; for the last two ones, add the retroflexed mid central vowel [☯ʳ] to the remaining i and 𝅥;
   f. For the velar nasal rimes, nasalize the nuclear vowels before deleting the velar nasal ending and then retroflex the nasalized nuclear vowels.

Obviously, according to Tao and Yin (1957) there are mainly two types of retroflex suffixation: one is adding the color of retroflex to the last (remaining) vocalic segment of the syllable; and another is adding the retroflexed mid central vowel [☯ʳ] to the end of the suffixed syllable or stem.

Li (1986) claims that there are two types of retroflexion in Mandarin: additive type and fusional type. For the additive type, the retroflex is just added to the end of the rime, i.e., a retroflex liquid is usually added to the end of the vowel of the suffixed rime. Take (金)桔儿 (tɕ𝅥+☯ʳ, orange) for example. The rime of the suffixed syllable [𝅥] is incompatible with retroflexion since it is a front high vowel and thus the retroflex suffix is added to the nuclear vowel of the suffixed syllable, forming a complex rime. For the fusional type, the retroflex action starts simultaneously with the articulation of the rime, i.e., the monosyllabic retroflex suffix and the suffixed nuclear vowel are fused into a singleton retroflexed vowel. Take (眼)珠儿(tɕu+☯ʳ, eyeball) for example. The rime [u] is compatible with retroflexion and thus fused with the retroflex suffix into a retroflexed singleton rime. Li claims that Mandarin has shown a constant change from additive to fusional retroflex suffixation. And he also makes an experiment to show there exist such two types of retroflexion now in Mandarin (Li 1990).

Wang and Wang (1991:98) criticize Li's classification on the basis of two types of retroflexion and his claim that Mandarin exhibits a constant change from additive to fusional retroflex suffixation. Based on their acoustic and perceptional experiments, Wang and He (1985:27) claim that in Mandarin, the retroflex action is almost simultaneous with the articulation of the rime of a syllable. Lin and Wang (1992:169) reiterate this idea that in Mandarin the retroflex action starts simultaneously with the articulation of the rime, with the onset unaffected. They point

out that under the retroflexion the vocalic segments in the rime (not including the pre-nuclear glide) have undergone some changes and the difference in Mandarin retroflexion mainly lies in the degrees of the retroflexion of the vowels in these rimes. The retroflex suffixed forms in Beijing Mandarin that Wang and He (1985:48) report are displayed as in the following table.

(2)

开口呼			齐齿呼			合口呼			撮口呼		
例词	原韵母	儿化韵	例词	原韵母	儿化韵	例词	原韵母	儿化韵	例词	原韵母	儿化韵
丝儿	ɿ										
枝儿	ɻ		鸡儿	i					鱼儿	🖫	
碑儿	ei	☯r			i☯r	柜儿	uei	u☯r			🖫☯r
根儿	☯n		个儿	in		棍儿	u☯n		裙儿	🖫n	
歌儿	ɤ	ɤr									
			街儿	iɛ	iɛr				月儿	🖫ɛ	🖫ɛr
(婆儿)	o	or				窝儿	uo	uor			
						屋儿	u	ur			
把儿	a		牙儿	ia		花儿	ua				
牌儿	ai	ar			iar	拐儿	uai	uar			
盘儿	an		尖儿	ian		罐儿	uan		院儿	🖫an	🖫ar
刀儿	au	aur	票儿	iau	iaur						
钩儿	ou	our	球儿	iou	iour						
缸儿	aŋ	ɑ̃r	亮儿	iɐŋ	iɑ̃r	筐儿	uɑŋ	uɑ̃r			
灯儿	☯ŋ	☯r	影儿	iŋ	i☯r	瓮儿	☯ŋ	u☯r	熊儿	🖫uŋ	🖫☯r
						空儿	uŋ	(ũr)			

Note that, as Chao (1968) and many other scholars point out, the front high vowels [i, 🖫] do not take the color of retroflexion as they are incompatible with retroflexion. When the front high vowel [i][2] (precisely speaking the off-glide) is preceded by a non-high (nuclear) vowel, it will be deleted when the rime is retroflexed; when the front high vowels are in the nuclear position, an embedded lax mid central vowel will be in the target of retroflexion. To put the rimes with the velar nasal ending aside, we generally have four basic types of retroflexed rimes: [a^{r}], [u^{r}], [☯^{r}] and [ᑕ^{r}][3] (Wang & He 1985:39). These four types of retroflexed rimes exhibit that the retroflex suffix has been fused into the rimes and is realized only as a retroflex feature attached to the vowels [a], [u], [☯] and [ᑕ]. Wang and He's idea that the retroflex suffix in Beijing Mandarin is realized only as a feature attached to the suffixed rime gains support from some other scholars. Lin and Wang (1992:169) point out that the rime (not including the pre-nuclear glide) will undergo some changes under the retroflex suffixation. Specifically, when the nuclear vowel of the suffixed syllable is a non-high vowel or a back rounded high vowel, the retroflex action starts just from the

[2] Since [🖫] does not occur in coda position, the front high vowel (glide) that follows a non-high vowel is only [i].
[3] To precisely represent that the segment is retroflexed, here I will use the superscript [^{r}] to indicate that the segment with the [^{r}] superscript is retroflexed.

articulation of this nuclear vowel; when the nuclear vowel is a front high vowel [i] or [□], the retroflexed mid vowel is added to it and the front high nuclear vowel becomes a correspondent pre-nuclear glide. Wang and He (1985) and Lin and Wang (1992) show with their acoustic and perceptional experiments that the retroflex suffixation in Beijing Mandarin is not merely an addition of a retroflex suffix to the end of a preceding syllable but actually incurs the retroflexion of the whole rime of the preceding syllable. Their studies are simply the denials of the additive and fusinal classification of retroflex suffixation in Beijing Mandarin as Li (1986) advocates. Next, I will present my empirical analysis of different kinds of retroflex suffixation in Mandarin.

1.2 Reconsidering the Retroflex Suffixation

In the previous subsection, I have mentioned that Mandarin retroflex suffixation has long been considered to be the incorporation of an independent suffix (either a syllable or a retroflex segment) into the preceding syllable. In process of the incorporation, the rime of the preceding syllable will be altered to some extent that it can be naturally co-articulated in spoken Chinese. Recent acoustic and perceptional experiments show that the retroflex suffixation in Mandarin is not a simple addition of a retroflex suffix to the preceding syllable but actually incurs the retroflexion of the whole rime of the preceding syllable. Thus, some scholars (e.g. Wang 1993) assume that the retroflex suffix is lexically not a segmental entity but an entity less than a full segment, i.e. just a single suffixal feature.

Obviously, the featural concept deviates from the traditional concept that the retroflex suffix is morphologically and diachronically a de-syllabified monosyllabic segment. Such a concept seems to represent the actual changes of the rime in retroflex suffixation but, on second thought, it will eventually lead to some problems. For instance, the retroflexed forms of [sɿ] (silky, 丝) and [tʂɻ] (twig, 枝) are respectively [s☯ʳ] and [ts☯ʳ]. Here, the apical high vowels are replaced by a retroflexed mid vowel [☯ʳ]. If we assume that the retroflex suffix is not a segment but a feature, in the phonological analysis, we have to consider epenthesizing a lax mid vowel after the dental or retroflex sibilant before the feature retroflex suffixation applies. How can we deal with the feature retroflex suffixation for the above cases? Possibly there are two ways of treatment. ①Suppose that the apical vowels are lexically empty. The epenthesis of the lax mid vowel is triggered by the retroflex suffixation. A retroflex feature is added to the syllable without a vocalic segment but this feature finds no place or segment to depend on and thus triggers the vowel epenthesis. But the problem with this way of treatment is why the retroflex feature applies to a lax mid vowel instead of other mid vowels. Obviously, the feature-based retroflex suffixation finds no motivation for the epenthesis of a mid vowel. ②Suppose that the apical high vowels are lexically present. The process of the retroflex suffixation is the combination of the apical vowel deletion and the mid vowel epenthesis. That is, the retroflex feature is added to the syllable and the lexically present apical high vowels are deleted due to the incompatibility for articulation with a retroflex feature. After the apical high vowels are deleted, the additive feature finds no place or segment to depend on and triggers the mid vowel epenthesis. A question will arise as to why the retroflex suffixation triggers the deletion of the only apical high vowels instead of all the high vowels. The naturalness of such an explanation is doubtful.

Consider the following facts concerning the retroflex suffix in Beijing Mandarin: ①The retroflex suffix evolves from a monosyllabic word. ②The suffix is spelled out in most cases as a morphological entity. ③In some dialects, the retroflex suffix is still an independent monosyllabic entity. And ④even among some old Beijing natives, the suffix is still retained as an independent morpheme. All these facts lead us to the following conclusion that the retroflex suffix in Beijing Mandarin is, at least lexically, an independent entity. There is an important point that needs to be called into attention. The lexically independent monosyllabic entity does not mean that it will be necessarily realized as a segmental entity. In fact, the lexically independent retroflex suffixal entity in Beijing Mandarin is, in most cases, phonetically realized as a rime-dependent feature. That is to say, the retroflex feature is a kind of superficiality, or, technically speaking, a phonetic realization of the underlying lexical form. A phonological analysis can and must present a clear and insightful picture of the whole process of the phonological phenomena.

But the same question will be raised: how is a segmental retroflex suffix phonetically realized as a retroflex feature attached to the rime of the suffixed syllable? This is just the question that the present paper is supposed to answer. Before answering this question, let us first talk about the significance of identifying the retroflex suffix as a lexically independent monosyllabic entity here.

The significance of identifying the Mandarin retroflex suffixation as a lexically independent monosyllabic entity can be generalized as the following points:

①It can reflect the historical source of this suffix. As many scholars point out, the suffixed syllable and the retroflex suffix were originally two separate phonological and morphological entities. The retroflex suffix lost its syllabicity and merged into the suffixed syllable due to the long time of fast speaking. The treatment of the retroflex suffix as a lexically independent entity can well reflect where this suffix comes from.

②It is fully in conformity to the morphological and psychological reality of the retroflex suffixation in Beijing Mandarin. As we know, in most of the cases, the retroflex suffix is written as an independent morpheme and the speaker usually tries to spell out the stem and this retroflex suffix at one breath. Morphologically and psychologically, the stem and the suffix are conceived to converge into one syllable when they are pronounced.

③It is of great significance to the typological generalization of retroflex suffixation across different Chinese dialects. Lu (1995), F-T. Wang (1999) and many others observe that there exist different types of diminutive retroflex suffixation across different Chinese dialects. To establish the typological relations among these different dialects, we may assume that they have the same type of a lexical diminutive suffix. Their difference in the phonetic realization of retroflex diminutive suffixation lies solely or principally in the ranking of some relevant universal constraints. As I will demonstrate, the phonetic realization of the retroflex suffix as an attached feature is mainly due to the dominance of the markedness constraints over the faithfulness constraints.

From the above analysis, we conclude that the Mandarin retroflex suffix is lexically an independent monosyllabic morpheme and, in most cases, is phonetically realized as a retroflex feature attached to the rime of the suffixed syllable in Beijing Mandarin.

Now the problem is: how is the retroflex suffix phonetically realized as a feature attached to the rime of the suffix syllable or, to put it in another way, how is the retroflex suffix incorporated

into the suffixed rime? As we have mentioned above, nearly all the scholars agree that when the nuclear vowel is a front high vowel, a retroflex mid vowel is added to it. But as for the other vowels, some scholars (e.g., Wang & He 1985; Lu 1995) consider that all of them are retroflexed; but some scholars (e.g., Li 1986) disagree and claim that some of them are retroflexed and some are not and for those that are not, a retroflex suffix is added to the rime in the retroflex suffixation. As for which vowels can be retroflexed and which are not, no agreement has been reached among the Chinese scholars.

Li (1990:171) reports that his analysis of his Beijing native speaker's pronunciation shows that the fusional retroflexion applies to the rimes ending with the back rounded high vowel [u] (see the syllables listed in Type B in his paper) while additive retroflexion applies to the other rimes (listed in Type A in his paper). Wang (1997b) studies the previous acoustic results and proposes that phonologically retroflexion is compatible with the [+back] vowels and the retroflex color of the [back] vowels incompatible with retroflexion is derived at the phonetic level of representation through the interpolation rules advocated by Pierrehumbert (1980). Specifically, Wang (1997b) assumes that phonologically the back vowels, such as [u], can be retroflexed while the front vowels, such as [i] and [a], cannot. Based on the experiment made by Bao and Yang in 1986, Bao (1989:104-5) concludes that fusional and additive retroflexion co-exist in the present-day Beijing Mandarin and which vowels are retroflexed in the fusional or additive way depends not on the retroflexed vowel itself but on the speakers. Wang and He (1985) claim that the retroflex action starts almost simultaneously with the articulation of the nuclear vowels of the suffixed syllable except when the front high vowel [i] or [ㄩ] is the nucleus. In sum, these studies show that the back vowels are compatible with retroflexion while the front high vowels are not. The front non-high vowels are in controversy.

Before presenting my analysis of Mandarin retroflex suffixation from an OT perspective, let's look at some of the pre-OT analyses made so far.

2. Past Phonological Analyses

Retroflex suffixation is very common in spoken Beijing dialect. When a word or a syllable is retroflexed, it will take a diminutive meaning. The most interesting aspect of Mandarin retroflex suffixation is that the retroflex suffix is not simply an additive suffix. When a word or a syllable is retroflexed, the retroflex suffix will lose its syllabicity and the rime of the suffixed syllable will undergo some changes. Such a linguistic phenomenon has attracted much attention of linguists. Studies have been conducted within different theoretical framework and some different proposals have been advanced so far. In the rest of this section, I will look at these studies and their proposals and see what achievements they have made and what deficiencies they have in phonological description and representation of the retroflex suffixation in Beijing Mandarin.

2.1 Linear Approach

Linear approach refers to the approach in which a phonological representation is assumed to be composed of a linear string of segments. Cheng (1973) was the first to conduct an analysis of

the Mandarin segmental phonology within the linear framework of Generative Phonology. Within this framework, Cheng (1973) presents his analysis on the Beijing retroflex suffixation. He assumes that the retroflex suffix is /ɤr/ underlyingly and the mid un-rounded vowel is deleted when the Retroflex Suffixation rule is applied. The remaining retroflex liquid /r/ is added to the rime of the suffixed syllable. To get the expected result, of course, some other rules must be required to involve in. Take “绳儿” (rope) for example. The rules and their ordering in the derivational process of the required form that Cheng (1973) proposes can be illustrated as in (3).

(3)

ʂɤ☠ ɤr
↓ ①Backness rule
ʂo☠ er
↓ ②Mid Vowel Laxing rule
ʂ☯☠ ☯r
↓ ③Retroflex Suffixation rule
ʂ☯☠r
↓ ④Velar Nasalization rule
ʂ☯☠r
↓ ⑤Velar Nasal Deletion rule
ʂ☯r

The figure in (3) shows the derivational process of the retroflexed syllable [ʂ☯r]. In the first step, the Backness rule applies and the non-high vowels are assimilated in backness respectively by their immediately neighboring segments [☠] and [r]. Then, the Mid Vowel Laxing rule involves in and makes the mid vowel before a consonantal ending become schwa. Next, the Retroflex Suffixation rule deletes both the vowel and the tone of the suffix and makes the remaining retroflex liquid as part of the suffixed syllable. After the Velar Nasalization rule nasalizes the preceding nuclear vowel and the velar nasal is deleted under the effect of the Velar Nasal Deletion rule, the expected result [ʂ☯r] is derived.

Clearly, there are some problems with Cheng’s rule-based derivation of retroflex suffixation. The r-retroflex is treated as a suffix of the retroflex liquid simply added to the rime of the suffixed syllable. But, as some acoustic and perceptional experiments show, in Mandarin the retroflex action is almost simultaneous with the articulation of the rime of a syllable. This means that the Mandarin retroflex suffixation is not the simple adding of a retroflex liquid to the end of the suffixed syllable but the complex fusional process of the retroflex suffix and the suffixed syllable. Obviously, Cheng’s (1973) analysis cannot exhibit such a phonological change in Mandarin retroflex suffixation clearly.

Like the other rule-based analyses, Cheng’s analysis is also subject to the external ordering of the rules and the inevitable divergence of the focus of study from the surface forms to the assumed abstract underlying forms. It can be clearly seen in (3) that the Backness rule, which effects the backness assimilation of the non-high vowels respectively to their immediately neighboring segments, is de-effected by the following Mid Vowel Laxing rule. It goes without saying that the problem of this kind is an ultimate outcome of the serial derivational approach. Besides, like most

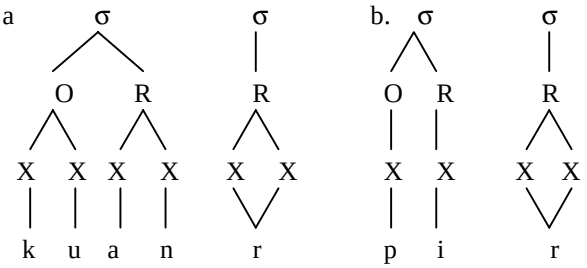
of the other rule-based theories, the rules are often formulated with excessive reference to the language specific properties and thus lack universality. And consequently, the observations of the typological variations across different languages are hard to be reflected.

2.2 Non-linear Approach

Non-linear approach to phonology starts in the 1970s. Goldsmith (1976), in studying the tones of African languages, develops the theory of Autosegmental Phonology. In this theory, phonological representations are composed of several independent tiers and tones are assumed to exist in one tier independent of the skeletal (or root) tier where a phonological change in one tier will not necessarily affect the phonological units in another tier. Autosegmental Phonology systematically develops a new way of phonological description and paves ways for the developments of other phonological theories. Clements and Keyser (1983) apply the non- linear autosegmental method to the study of syllable internal structure and advance the CV theory. Other syllabic theories advanced after the inception of Autosegmental Phonology includes X-slot theory, government phonology, mora theory, etc. Clements (1985) and Sagey (1986) adopt the non-linear approach to account for the intrinsic relations between the phonological features and develop the theory of feature geometry (see Wang 2000; Ma 2001a). Of course, these studies have great influence on study of Mandarin phonology and have touched off a number of studies on the Mandarin retroflex suffixation in the 1980s. The representative of these studies are Yin (1986), Lin (1989), Duanmu (1990), Wang (1993) and Goh (1997)[4].

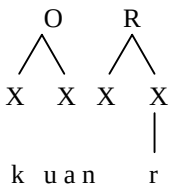
Yin (1986) proposes that the retroflexed rime in Beijing Mandarin is the result of re-syllabification of the stem and the retroflex suffix. Her autosegmental treatment of Mandarin retroflex suffixation can be illustrated as follows (see Lin 1989:101; Wang 1993:186):

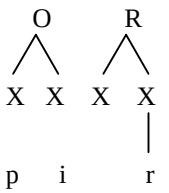
(4) a



b.

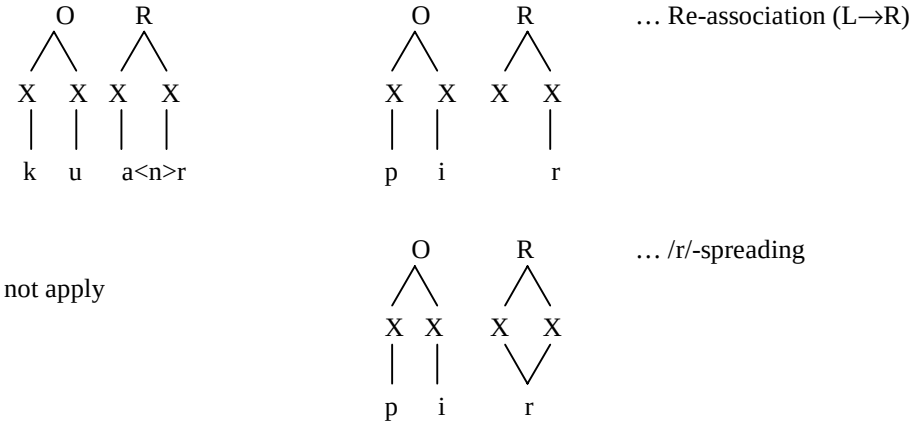
… Underlying Forms





… Syllable Merger

[4] Goh (1997) discussed Mandarin retroflex suffixation within the framework of government phonology. Since government phonology is one of the holistic theories, different from the DF-based studies made by the scholars like Yin (1986), Lin (1989), Duanmu (1990) and Wang (1993), I will make no comments on it here. Interested readers can refer to Ma (1999; 2000a, b).



In (4), the processes of retroflex suffixation for [kuan] (official, 管) and [pi] (pen, 笔) are presented. Two points need to be noted here: ①The underlying (or lexical) form is [☯^{r}] (that is, represented by /r/ associated with two timing slots dominated by R) and it has two alternative phonetic realizations [r] (when re-associated with only one timing slot) and [☯^{r}] (when re-associated with two timing slots). ②The syllabic template is not fixed. It may consist of either two syllabic constituents (O and R) or one (only R). Besides, a syllable may have either four timing slots (as in [kuan]) or two (as in [pi]). But the syllabic template for the retroflexed form is fixed. It is composed of four timing slots, two of which are dominated by O and two by R.

Within the non-linear framework, Lin (1989) also makes an analysis of the Mandarin retroflex suffixation. Lin's (1989) analysis is very similar to Yin's (1986). But hers is different in three respects: ①Lin assumes that the underlying form of the suffix is a retroflex liquid /r/. The phonetic realization [☯r] derives from the schwa epenthesis. ②She refuses that there is an invariable syllabic template for the retroflexed form (Lin 1989:101). ③Unlike Yin (1986, 1989), Lin considers that vowel epenthesis should be employed to account for the fact that schwa appears in some of the rimes like [pi] after the retroflex suffixation.

Duanmu (1990) studies eight Chinese dialects representing eight Chinese dialect families and proposes that all the Chinese dialects share the same invariable syllabic template, which consists of three timing slots, one for the Onset and two for the Rime. With his fixed three-slot syllabic template, Duanmu presents his analysis of the retroflex suffixation in Mandarin Chinese. Duanmu (1990:53) claims that the retroflex suffixation is a process of adding a suffix /r/ to a (monosyllabic) noun and then the noun and the retroflex suffix are fused into one syllable. Specifically, if the second slot (or coda slot) of the rime is empty, the suffix /r/ will be added directly to it; if the second slot has been occupied by a post-nuclear glide or nasal consonant, it will replace the original coda that occupies the coda position. Duanmu (1990:54) points out that in the syllabification, the retroflex suffix /r/ surfaces as [+low] after [a], [+high] after [u] and [–high, –low] otherwise. He proposes to utilize feature geometry to account for the retroflex color that the nuclear vowel obtains after the suffix /r/ is added.

The three studies introduced above are very similar to each other. All of them consider that the retroflexed rimes result in the inclusion of the retroflex suffix to the stem syllabic template. The suffix /r/ is incorporated and occupies the coda position if the coda position is vacant, or replaces the segment in the coda position if this segment is already in the coda position and the replaced segment will be deleted by the stray erasure rule. There is no problem if the segment in this position is [i] or [n] since it is acoustically shown that they are lost after the retroflex suffixation. But if the segment in the coda position is [u] or [ŋ], the above way of treatment, the way of replacing the segment in the coda position with /r/, is impracticable, since the experiments show that the segment [u], or at least the feature [+round], is retained and the [ŋ] nasalizes the preceding vowel before it is replaced by the retroflex suffix [r].

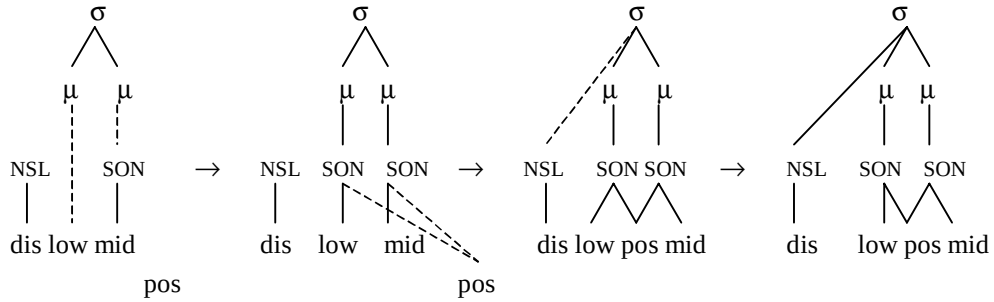
To account for these, Yin (1989) and Lin (1989) both maintain that the /u/ ending in the retroflex final, such as /aur/, is a separate segment but forms a short diphthong with a low vowel. Take /aur/ for example, they assume, following Fu's (1980) view, that /a/ and /u/ form one segment /a̯/ in the retroflex final, that is, /aur/ is assumed to be phonetically realized as /a̯r/. As for the ending /ŋ/ in the retroflexed rime, Lin (1989:106-8) assumes that the retroflexed rime with the /ŋ/ is different from the rime with /n/ in that in the former case the velar nasal is re-associated with the nuclear position after being de-linked with the coda position in the skeletal tier as the retroflex suffix is linked to it. Lin (1989:107) suggests that the re-association rule applies before the stray erasure rule comes in. The re-association rule stipulates that the segment to be re-associated must be with the feature specification [+back].

Duanmu's (1990) treatment is similar to Yin's (1989) and Lin's (1989). He proposes that the ending /u/ is first replaced by the retroflex suffix /r/ and then the feature of the /u/ is re-associated with the /r/ by means of "feature recycling" (Duanmu 1990:56-7), thus resulting in the rounded retroflex ending. As for the rime with /ŋ/, Duanmu (1990) maintains that it nasalizes the preceding nuclear vowel before it is replaced by the retroflex suffix /r/.

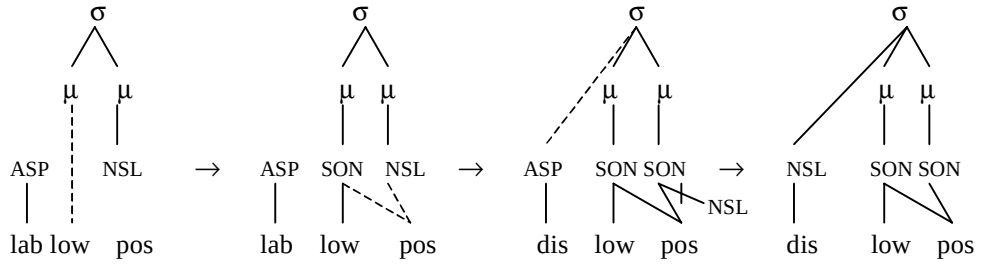
Wang (1993) criticizes that those treatments proposed above do not provide a satisfactory explanation for retaining /u/ or nasalizing the nuclear vowel in the retroflex suffixation. She argues that a satisfactory phonological explanation for the retroflex suffixation in Mandarin Chinese must have recourse to feature geometry (Clements 1985; Sagey 1986; McCarthy 1988) and under-specification theories (Archangeli 1984, 1988; Pulleyblank 1988a, 1988b) advanced in the late 1980s. The syllabic template she proposes, very similar to Duanmu's (1990), has three skeletal positions, one for the onset and two for the rime. The difference is that she adopts the moraic theory, that is, the rime of a heavy syllable in Mandarin is represented by two moras (μμ), which is linked by two roots (rt), equivalent to two segments. The retroflex suffixation is derived from the mechanism of feature incorporation. The feature retroflex suffix she proposes is [posterior], shorten as [pos] in her dissertation. [pos] is contained in the lexicon and in the process of retroflex suffixation it is compulsorily incorporated into the rime. Specifically, the feature [pos] is linked to the two roots of the rime. She points out that sometimes the incorporation of the feature [pos] in some segments is simply the addition of the feature to the roots but sometimes it may involve de-linking of certain incompatible stem features. The incompatible stem feature she claims is [coronal]. That is, the segment with the feature specification [coronal] is incompatible with the retroflex feature [pos] and will be de-linked with the root node when the retroflex feature is

compulsorily associated with the root nodes of the rime[5]. And thus, the ending /w/ or /□/ will not be replaced since they are not [coronal] segments but the ending with /y/ (in her transcription /j/) or /n/, or the rime /i/, /□/, /ɪ/ or /□/ will be replaced by the feature [pos] since they are specified with the incompatible feature [coronal]. The two phonological processes, feature addition and feature changing, in the retroflex suffixation are illustrated with the following two examples.

(5) a. feature addition: [$n^{y}a^{r}w^{r}$] (bird, 鸟儿) (Wang 1993:189)



b. feature changing: [$p^{h}a^{r}$] (dish, 盘儿) (Wang 1993:192)



(5a) and (5b) are the illustrations of how the retroflex feature [pos] is incorporated in Wang's analysis. In (5a), the feature [pos] is linked to the root node directly without triggering any change to the original structure since the features of the nodes are compatible with [pos]. Thus, it results in the addition of a single retroflex feature to the nodes. But in (5b), the linking of the feature [pos] to the nodes results in the de-linking of the feature [NSL] since these two features are incompatible with each other.

Indeed, it seems true that the retroflex suffix is a feature suffix since the suffix is, in most cases, realized as a feature attached to the rime. But, as I have pointed out, this is only the superficiality. As a matter of fact, some old Beijing natives now still treat the retroflex suffix as a segment or a monosyllable (F-T. Wang 1999:112). And therefore, it is maintained that the retroflex suffix is best to be treated lexically as a monosyllabic retroflexed mid vowel and it surfaces as a retroflex feature in accord with a set of ranked constraints.

It should be pointed out that these non-linear approaches discussed above benefit a lot from the theory of syllabic template developed in the mid 1980s. In Yin (1989) and Lin (1989), the retroflex suffix is incorporated into the stem syllable by way of the syllable merger or syllable

[5] Duanmu (1990) has proposed the similar feature geometric treatment of the retroflex suffixation. The retroflex suffix is a root dominating the place feature [cor] with the terminal feature [−ant].

contraction rules. Duanmu (1990) presents a similar analysis to Yin's and Lin's. Of course, Duanmu's analysis is principally segment- based but it is obvious that he has begun to appeal to the feature geometry feature in his analysis. Wang (1993) gives us a totally featural account for the Mandarin retroflex suffixation. It is apparent that she holds the view that retroflex suffix is not a segment but only a feature instead.

In the following sections, I will discuss in detail what constraints are required in accounting for the retroflex suffixation in Mandarin Chinese and how these constraints are ranked in deriving its optimal expected surface retroflexed rimes. It is implicitly suggested that many advantages will be reaped if we take the Mandarin retroflex suffix as a monosyllabic morpheme instead of a single feature.

3. An OT Analysis

Having reviewed and discussed the empirical studies and the past phonological analyses, I now turn to the analysis made from an OT perspective. I will first present my OT account for typological generalizations of diminutive retroflex suffixation across Chinese languages and then focus on different aspects of Mandarin retroflex suffixation.

3.1 Typological Variations and Constraint Rankings

It has been acoustically and perceptively proved that the retroflex diminutive suffix in Mandarin is phonetically realized only as a retroflexed feature attached to the rime. But from the morphological or diachronically phonological point of view, the Mandarin retroflex suffix apparently derives from the morpheme [☜ʳ]. Lu (1995) claims that there are generally three types of retroflex diminutive suffixation among different Chinese dialects. The first type is that the diminutive suffix does not lose its syllabicity and is realized as a monosyllabic morpheme when it is attached to the suffixed syllable. The second type is that the diminutive suffix loses its syllabicity but is still realized as some form of consonantal or vocalic ending of the rime. And the third type is that the diminutive suffix loses its syllabicity and is realized only as a retroflexed feature incorporated into the rime of the syllable. Beijing Mandarin is a typical example of the third type (see also F-T. Wang 1999).

Now let's take the syllable [ku] for example. The above three types of retroflex diminutive suffixation can be exemplified as follows.

(6)

Input	Type A	Type B	Type C
[ku$^{\mu\mu}$] [☯$^{\mu\mu}$]	[ku$^{\mu\mu}$.☯$^{\mu\mu}$]	[kur$^{\mu\mu}$][6]	[ku$^{\mu\mu}$]

[6] It should be pointed out that the non-syllabic lexical retroflex suffix is marked differently in the Chinese linguistic literature. Some scholars posit [r] for it (e.g. Chao 1968; Hockett 1950; Hsueh 1985; Pulleyblank 1983; Lin 1989; Duanmu 1990) and some use [☯ʳ] (Cheng 1973; R. Cheng 1966; Yin 1986). My analysis in the following subsections will implicitly show that the retroflex suffixes [r] and [☯ʳ] are non-distinctive in Mandarin Chinese, as the retroflex suffix will be erased due to some markedness constraints after triggering the assimilation process. Thus the retroflexed form for [ku] can be marked here either as [kur] or as [ku☯ ʳ].

It can be easily seen from (6) that Type A is faithful to the input both in the number of segments and in the number of moras; Type B is faithful to the input in the number of segments but not in the number of moras; and Type C is not faithful to the input either in the number of segments or in the number of moras.

In terms of the different types of diminutive suffixation listed above, I consider the following four constraints, two faithfulness constraints and two markedness constraints, are involved:

(7) a. MAX-IO: Every segment of the input has a correspondent in the output.
b. MAX-μ-IO: Every mora of the input has a correspondent in the output.
c. SUFFSTM=σ: The suffix and the suffixed stem must be one syllable.
d. ALIGN-R (Stem, Word): The right edge of the stem must be aligned with the right edge of the phonological suffixed word.

(7a) is a faithfulness constraint that penalizes the output candidate unfaithful in the number of segments. (7b) is another faithfulness constraint here and it requires that the number of moras of the input must be preserved in the output. (7c) is a markedness constraint that requires the stem and the affix be incorporated into one syllable. (7d) is an alignment constraint requiring that the right edges of the morphological stem and the phonological word must be aligned with each other.

It is assumed here that the three types of retroflex diminutive suffixation reported in different Chinese dialects can be explicitly expressed through the different rankings of the four constraints posited above.

(8) The constraints ranked for three types of Chinese retroflex suffixation
Type A: MAX-IO, MAX-μ-IO >> SUFFSTM=σ, ALIGN-STEM-R
Type B: SUFFSTM=σ, MAX-IO >> MAX-μ-IO, ALIGN- STEM-R
Type C: SUFFSTM=σ, ALIGN- STEM-R >> MAX-μ-IO, MAX-IO

Obviously, the typological variations between the three types of retroflex diminutive suffixation are fully expressed by way of ranking the four constraints posited here. In Type A, the two faithfulness constraints are ranked over the two markedness constraints, blocking any phonological changes under the diminutive suffixation. This ranking of the constraints well predicts the first type of diminutive suffixation depicted in Lu (1995:69) in which the diminutive suffix [☯^{r}] does not lose its syllabicity in the suffixation process. In Type B, one of the markedness constraints is promoted into the dominating position in the constraint hierarchy while one of the faithfulness constraints originally in the dominant position is demoted into the dominated class. Such a transposition between the faithfulness constraint and the markedness constraint in the hierarchy indicates some changes may be undergone under the retroflex suffixation. Specifically, the diminutive suffix will lose its syllabicity in terms of the number of moras but not in terms of number of segments, as reported in Lu (1995:69-70). In the constraint hierarchy for Type C, the two markedness constraints dominate the two faithfulness constraints, triggering some structural changes in the output forms. The diminutive suffix loses its syllabicity both in the number of moras and segments, and the retroflex suffix is realized as a retroflex feature

attached to the syllable, as is shown in Beijing Mandarin.

The effects brought out by the different rankings of the constraints can be displayed below with the word [ku] as an example.

(9) Type A: MAX-IO, MAX-μ-IO >> SUFFSTM=σ, ALIGN-STEM-R

Input: $ku^{\mu\mu}$+☯$^{r\mu\mu}$	MAX-IO	MAX-μ-IO	SUFFSTM=σ	ALIGN-STEM-R
☞a. $ku^{\mu\mu}$.☯ $^{r\mu\mu}$			*	*
b. $ku^{\mu}r^{\mu}$		*!		*
c. $ku^{r\mu\mu}$	*!	*		

In the tableau above, (9c) is assessed as the most unfaithful candidate by the two dominant constraints, MAX-IO and MAX-μ-IO, and thus eliminated as a non-optimal output. (9b), though satisfying one of the two dominant faithfulness constraints (MAX-IO), is still eliminated as it fatally violates the other dominant faithfulness constraint MAX-μ-IO. (9a) turns out to be the optimal candidate since its violations are minimal compared with the other candidates.

(10) Type B: SUFFSTM=σ, MAX-IO >> MAX-μ-IO, ALIGN-STEM-R

Input: $ku^{\mu\mu}$+☯$^{r\mu\mu}$	SUFFSTM=σ	MAX-IO	ALIGN-STEM-R	MAX-μ-IO
a. $ku^{\mu\mu}$.☯ $^{r\mu\mu}$	*!		*	
☞b. $ku^{\mu}r^{\mu}$			*	*
c. $ku^{r\mu\mu}$		*!		*

In this tableau, (10a) incurs a violation of the dominant markedness constraint SUFFSTM=σ and is dropped out of the competition. (10c) is penalized as a sub-optimal output because it is not faithful to the input in the number of the segments it has. Finally, (10b) is chosen as the optimal candidate since it least violates the ranked constraints.

(11) Type C: SUFFSTM=σ, ALIGN-STEM-R >> MAX-μ-IO, MAX-IO

Input: $ku^{\mu\mu}$+☯$^{r\mu\mu}$	SUFFSTM=σ	ALIGN-STEM-R	MAX-IO	MAX-μ-IO
a. $ku^{\mu\mu}$.☯ $^{r\mu\mu}$	*!	*		
b. $ku^{\mu}r^{\mu}$		*!		*
☞c. $ku^{r\mu\mu}$			*	*

(11a) and (11b) are dropped out because they respectively fatally violate the dominant markedness constraints SUFFSTM=σ and ALIGN-STEM-R. (11c), though the most unfaithful to the input, is chosen as the optimal output as it violates the only two un-dominated faithfulness constraints.

Interestingly, the process of changing from the rime with the monosyllabic suffix to the fully retroflexed rime that has been observed in different Mandarin dialects is well predicted in the OT analysis presented above. It is assumed that this changing process is the one from the additive to the fusional retroflex suffixation. In the OT vocabulary, it is a process of faithfulness demotion and markedness promotion. The three types of retroflex diminutive suffixation reflect the three

stages of this process.

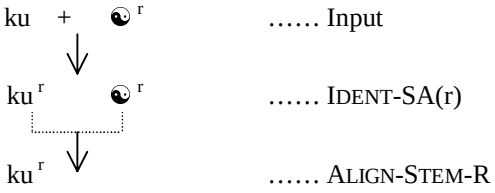
It is of great significance that the teleological or output-oriented linguistic theory can predict the different stages of historical changes of a language or typological variations of retroflex suffixation across Chinese languages. Therefore, in this sense, OT is more advantageous over the classical rule-based generative models in which extrinsic ordering of a series of rewrite rules is applied to account for those phonological changes discussed above.

As we have discussed above, in Beijing Mandarin, the diminutive suffix ☯^{r} is realized as the only feature attached to the syllable rime in the process of retroflex suffixation. Obviously, the suffixation of this kind is exactly predicted in the constraint hierarchy designed for Type C. Now we can claim that the retroflexed rime in Beijing Mandarin actually results from the structural changes triggered mainly by the dominating markedness constraints like SUFFSTM=σ and ALIGN-STEM-R I propose here. Having discussed the general type of the retroflex suffixation in Mandarin, I will turn into some specific changes in retroflex suffixation observed in Mandarin Chinese.

3.2 Opacity in Retroflex Suffixation and Sympathy Theory

In the previous subsection, I have discussed how to account for the general typological differences from an OT perspective. Careful readers will find many problems are still unsettled with the limited number of four ranked constraints. For instance, in (11) the markedness constraint ALIGN-STEM-R requires to align the right edge of the suffixed form with the right edge of the stem of the word suffixed. Of course, this constraint eliminates the candidates ended with [☯^{r}] or [r] and at the same time assesses the sub-optimal candidates like [ku] as one of the optimal ones. To make a right choice between the candidates like [ku] and [ku^{r}], some other constraints are surely needed. Let's take the [ku^{r}] for example. As is known, its input form is the stem [ku] plus the retroflex suffix [☯^{r}]. It is obvious that the retroflex feature in the rime of the syllable [u^{r}] derives from the spreading of the retroflex feature from the suffix. After the stem gains this feature from the suffix, the suffix is lost. This process can be schematically illustrated in the following figure.

(12)



(12) shows that the rime of the stem is retroflexed under the effect of the horizontal faithfulness constraint on the segments between the stem and the suffix. Thus we posit the output-output faithfulness constraint here. This OO-IDENTITY constraint is operative between the stem and the suffix and hence defined as follows:

(13) IDENT-SA(r): A segment α in the stem must be identical in retroflex features

specification to its correspondent β in the suffix.

The constraint in (13) is a faithfulness constraint that stipulates the segments in the stem must be identical to the suffix in terms of retroflex feature. Consider the S-IDENT (F) presented by Krämer (1998, 2000). The constraint proposed here in (13) is similar in nature to it and can thus be replaced by S-IDENT(r).

Please note that some scholars may express this retroflex process by positing the constraints like the retroflex harmony. Yip (1994), in accounting for the nasalization in Chaoyang dialect, one of Chinese languages, proposes two harmony constraints: Rhyme Harmony and Syllable Harmony. In doing so, she presupposes that the nasalization derives from the assignment of a feature to the rime or the syllable. Zhou (1995) adopts the same way of treatment in accounting for the Mandarin retroflex suffixation. But it is argued here that the way of treating Mandarin retroflex suffixation by the feature-rooted retroflex harmony constraints superficializes or simplifies the fact that the retroflexed rime as in Beijing Mandarin has been theoretically and empirically proven to result from the assimilation process of the independent monosyllabic retroflex suffix. The process of changing form the rime with an additive retroflex suffix to the fusional retroflexed rime can be occasionally found even in the present-day Beijing Mandarin (F-T. Wang 1999:112). Hence, the present paper claims that the feature-assigned harmony constraint does not tell us where the assigned feature comes from. The OO-IDENTITY constraint proposed here is similar to the harmony constraint in that it requires all the segments in the stem share the same feature specification with the suffixal segment but it is different from the feature-rooted harmony constraint mainly in that it can reflect the source of the feature in harmony. Another problem with the feature-rooted harmony constraint is that such a constraint is hard to be relegated into any one of constraint families proposed up to now. And, therefore, it is assumed here that this kind of constraints is transitorily employed and will be confronted with many more problems if more evidence is brought into consideration in the phonological analysis of retroflex suffixation across different Chinese dialects.

As we have pointed out, the stem and suffix featural identity constraint proposed here is a member of faithfulness constraint family and states that the segments in the stem must be identical in retroflex feature specification to the segments in the retroflex suffix. Now the problem is that the environment that triggers the retroflex assimilation in the stem is demolished in the later derivation. This is very common in Beijing Mandarin. Consider the retroflex suffixation of [kə̃ ɻ] (root, 根儿) and [kɑ̃ ɻ] (vat, 缸儿). In the former case, the consonantal coda [n] triggers the laxing of the mid vowel and then is deleted under the retroflex suffixation. And therefore the mid vowel in the retroflexed form with the lexical consonantal coda (as in [kə̃ ɻ]: root, 根儿) is different from the mid vowel in the retroflexed form without the lexical consonantal coda (as in [kɤ ɻ]: song, 歌儿). The same is true for the latter case. The nuclear vowel becomes nasalized before the lexical consonantal velar nasal is deleted under the retroflex suffixation. The above cases can be generalized as the following schema:

(14) AC ⟶ BC ⟶ B∅

A becomes B before C	C is deleted

(14) shows that A is triggered by C to change to B but the trigger C is deleted after A has changed to B. Please note that A will not become B if it does not precede C. C is obviously the trigger of the A's change. That is, some generalization G (here the OO- IDENTITY constraint given in (13) plays an active role in shaping of the surface form F (here the B). But the conditions that lead to the application of the G (here the trigger C) are not apparent or opaque in F (McCarthy 1998:1). It should be admitted that within the classical derivational generative phonology, the phonological opacity illustrated in (14) can be easily treated by establishing an intermediate representation (here BC) (Kigursky 1971, 1973). But the problem will arise for Optimality Theory, the theory in which no intermediate level of representations is allowed.

In the following subsections, I will argue that the problem that we have encountered in Mandarin retroflex suffixation can be tackled from the OT perspective, but only with reference to Sympathy Theory. To make my OT analysis of Mandarin retroflex suffixation easily accessible to the readers, I will first present a brief introduction of Sympathy Theory initiated in McCarthy (1998) in the next subsection.

3.2.1 Sympathy Theory

Sympathy Theory is a theory proposed by McCarthy (1998, 2000) for a parallel treatment of phonological opacity that some scholars claim to be a big problem for the teleological OT analysis. Phonological opacity, the term first dubbed by Kiparsky (1971, 1973), refers to the case in which a phonological rule is not surface-true or surface-apparent by the application of subsequent rules. That is, the output forms are shaped by some linguistic generalizations that are not surface-true or surface- apparent. These can be schematized as in (15).

(15) a. $\varnothing C \rightarrow BC \rightarrow B$
b. $AC \rightarrow BC \rightarrow B$[7]

In (15a), the condition C triggers the insertion of B and then C, the trigger, is deleted. It is obvious that the insertion plays an active role in the shaping of the surface form B but its function is not surface-true. In (15b), A becomes B because it appears before C and C, which triggers the change from A to B, is finally deleted. Obviously, the derivation of the surface forms in (15) depends crucially on the intermediate stages on which some linguistically significant generalizations function actively in the shaping of the output forms.

To account for the above cases within the teleological framework of OT, McCarthy (1998, 2000) systematically proposes Sympathy Theory. The crucial claim of this theory is the selection of the sympathetic candidate and the selection of the actual output take place in parallel. Specifically, one of the failed candidates is chosen as a model candidate that all the other candidates are required to resemble. The selection of this model candidate relies primarily on a

[7] (15b) is equal to the case represented in (14).

designated IO faithfulness constraint. The model candidate, generally called the sympathetic candidate, must obey the designated IO faithfulness constraint, often referred to as the sympathy selector. It is clear that usually there may be several candidates that obey the sympathy selector. Among the candidates that obey the sympathy selector, the sympathetic candidate is the most harmonic with respect to the rest of the constraints. Once the sympathetic candidate is chosen, all the other candidates are required to resemble it through the candidate-to-candidate constraint, i.e., Sympathy. Now take the often-quoted Tiberian Hebrew case for example. The process of obtaining the actual output /deše/ from the lexical form /dešʔ/ by way of parallel constraints is schematized as follows:

(16) McCarthy (1998:12)

	/dešʔ/	CODACOND	☸ MAX-V☸	⋆MAX-C_{IO}	DEP-V_{IO}
opaque	☞ a. deše			*	*
transparent	☜ b. deš		*!	*	
sympathetic	☸ c. dešeʔ	*!		√	*

Here, the sympathetic selector is ⋆MAX-C_{IO}, which selects (16c) as the sympathetic candidate. ☸MAX-V☸ is the sympathetic faithfulness constraint, which requires preservation of vowels of the ☸-candidate in the output form. (16b) is a transparent candidate but it drops out of the competition because of the fatal violation of the sympathetic faithfulness constraint ☸MAX-V☸. (16a) is selected as the optimal candidate though it is an opaque candidate.

It is worth noting that in Tiberian Hebrew the insertion of /e/ here is triggered by the constraint prohibiting the consonant clusters. But the final consonant is deleted due to the coda conditions imposed on it. Obviously, the final consonant is deleted after triggering the insertion of /e/ and otherwise /deš/ would be the optimal candidate. In Optimality Theory, sympathy, the candidate-to-candidate constraint, must be required so that a sympathetic candidate is chosen from all the failed candidates. With the ☸-candidate, the actual optimal output will be able to be obtained.

3.2.2 Sympathy in Mandarin Chinese

Now come back to Mandarin retroflex suffixation. In Mandarin retroflex suffixation, just as many scholars claim, the retroflex suffix is incorporated into the suffixed syllable and it is phonetically realized as a retroflex feature attached to the rime of the syllable. If the suffix is taken as a lexical monosyllabic entity, the retroflex suffixation can be expressed technically as: the suffix spreads its retroflex feature to the rime of the suffixed syllable and is then deleted. Similar cases can be found in Mandarin Chinese. For example, the velar nasal nasalizes the rime of the retroflexed syllable and is then deleted. Another example, the dental nasal creates the condition that makes the mid vowel become lax and, in the retroflex suffixation, the lax mid vowel that the deleted dental nasal has brought about is retroflexed. This might be the reason that the retroflexed forms [kə̃ˠ] (root, 根儿) and [kɤˠ] (song, 歌儿) are different.

How to account for the retroflex suffixation in terms of Sympathy Theory? It is claimed here that in Mandarin Chinese the sympathy selector is the faithfulness constraint ⋆MAX-IO, which requires preservation of all the lexical segments in the ☸-candidate. The sympathetic faithfulness constraint is IDENT-☸O(r), which requires the candidates to be faithful to the ☸-candidate in terms of the retroflex feature. This proposal can be informally characterized as in the following tableau.

(17)

	Inpt: ku+☯ r	ALIGN-Stem-R	IDENT-☸O(r)	⋆MAX-IO	IDENT-IO(r)
sympathetic	☸ a. ku r☯ r	*!		√	*
transparent	☜ b. ku		*!	*	
opaque	☞ c. ku r			*	*

Note that if the sympathetic faithfulness constraint IDENT-☸O(r) were not involved, the optimal candidate would be (17b). It is due to the fatal violation of this constraint that (17b) is eliminated.

Careful readers may find that the selection of the ☸-candidate cannot fully rely on the faithfulness constraint ⋆MAX-IO since some other candidates also satisfy it. As I have pointed out above, in addition to the sympathetic selector, some other constraints sometimes should be considered in the determination of the ☸-candidate. For instance, the candidate [ku☯ r] is one of the candidates that satisfy the sympathetic selector ⋆MAX-IO and it is chosen as the ☸-candidate because it least violates the constraint hierarchy among those candidates that satisfy the sympathetic selector ⋆MAX-IO. In this case, the OO correspondence constraint proposed in (13), rewritten here as S-IDENT (r), is obviously required.

(18)

Inpt:ku+☯ r	ALIGN-Stem-R	IDENT-☸O(r)	⋆MAX-IO	S-IDENT(r)	IDENT-IO(r)
☸ a. ku r☯ r	*!		√		*
b. ku☯ r	*!	*	√	*	
c. ku		*!	*		
☞ d. ku r			*		*

It is clear that (18a) and (18b) both satisfy the sympathy selector but the latter candidate is not chosen as the ☸-candidate obviously because it is less harmonic than (18a) with respect to the constraint set. And thus (18a) turns out to be the ☸-candidate for the Mandarin retroflex suffixation.

Up to now, we have discussed the role of sympathy in the Mandarin retroflex suffixation and proposed that the sympathy selector of Mandarin Chinese is ⋆MAX-IO, the constraint that preserves all the segments of the input and the sympathetic faithfulness constraint is the featural identity constraint, IDENT-☸O(r) for the retroflex suffixation. It should be claimed that the afore-mentioned sympathy treatment of the retroflex suffixation is also applicable to the other similar phonological opacities in Mandarin Chinese. In the following subsections, I will demonstrate how to deal with the specific cases of retroflex suffixation within the OT framework.

In the discussion, I will make analyses of different aspects of phonological processes including those afore-mentioned phonological opacities that are often encountered in the discussion of Mandarin retroflex suffixation.

3.3 The Constraints and Their Ranking

In this subsection, I will concentrate my attention on specific aspects of Mandarin retroflex suffixation from the OT perspective. What constraints most probably play an active role in the suffixation process? How do they interact in obtaining the optimal output? And in what way are these constraints ranked?

As we know, the /☯ r / (or /r/) is morphologically a suffix, which is added to the ending of the stem syllable. It is clear that the retroflexed syllable must end with the retroflexed segment. This can be expressed by a constraint in (19).

(19) R-ANCHORING-r[8]: The element [retroflex] at the right periphery of the input has a correspondent at the right periphery of the output.

This constraint stipulates that the suffixed syllable must have (at least) one retroflexed segment at its right edge. A suffixed syllable may violate it if the segment at its right edge is not retroflexed in correspondence to the retroflexed mid vowel in the input.

But on the other hand, as we know, not all the segments can be retroflexed. Chao (1968:46) proposes the "simultaneity of compatible articulation", a general principle determining whether or not the segment in the rime can be retroflexed in the retroflex suffixation. If a coda segment in the rime is incompatible with retroflexion, it will simply be dropped. Wang (1993) claims that the segment incompatible with retroflexion is specified with the coronal features. But the coronal vowel [ɿ] does not obey this constraint, that is, it is not replaced by [r] in the suffixation. Wang (1993:191) argues that the original coronal feature "dis" of the segment [ɿ] is replaced by the retroflex feature [pos] in the retroflex suffixation but the "mid" specification for its dorsal feature remains. Wang's (1993) analysis shows that what is incompatible with retroflexion is most possibly the coronal features. But her explanation for why the coronal segment [ɿ] is compatible with retroflexion is not satisfactory enough. It is claimed that it is better, at least at moment, not to stipulate that the coronal segment is incompatible with retroflexion in Mandarin Chinese.

In Section 6.1.2, I pointed out that the past empirical studies show that the back vowels are compatible with retroflexion while the front high vowels are not. The front non-high vowels are in controversy. This means that probably there is a general scale on one side of which segments are susceptible to retroflexion and on the other side of which segments are not. This scale can be roughly and provisionally embodies as: back vowels (easy to be retroflexed) to front low vowels to front mid vowels to front high vowels to consonants (hard to be retroflexed). Thus a set of

[8] (L or R) ANCHORING constraints can be served to distinguish the prefix and the suffix. McCarthy and Prince (1999) argue that this kind of constraints should subsume *Generalized Alignment*. The R-ANCHORING-r, proposed here, can be replaced by a alignment constraint, which is expressed as ALIGN-R (σ, r) (The right edge of the syllable must be aligned with the right edge of the retroflexed segment).

constraints is derived as follows:

(20) a. *r/C[9]: R-retroflexion is not compatible with a consonantal segment.
b. *r/I: R-retroflexion is not compatible with a front high vowel / glide.
c. *r/E: R-retroflexion is not compatible with a front mid vowel / glide.
d. *r/A: R-retroflexion is not compatible with a front low vowel / glide.
e. *r/U: R-retroflexion is not compatible with a back vowel / glide.

These constraints of this set, like moraic constraints, are ranked, which are shown below.

(21) *r/C >> *r/I >> *r/E >> *r/A >> *r/U

Here, the constraints on retroflexion are ranked. Like moraic constraints, some other conflicting constraints can be intervened in between. For instance, if Mandarin allows the front high vowels to be retroflexed, then the constraints triggering the retroflexion can be presumably intervened before the constraint *r/I. If Mandarin only allows the back vowel to be retroflexed, it can be assumed that the constraints triggering the retroflexion take effects before the constraint *r/U involves in. Following Wang and He (1985), I assume that except for the front nuclear high vowel, all the non-high vowels and back vowels can be retroflexed. Specifically, the constraints triggering the retroflexion must be ranked below *r/I and above *r/E, i.e., the constraints *r/C and *r/I are in the dominating position, blocking the retroflexion of the consonants and the front high vowels, while the constraints *r/E, *r/A and *r/U are lowly ranked and thus do not take effects. For ease of exposition, in the following analysis I will occasionally just use *r/I to replace the constraint set *r/C >> *r/I.

Clearly, this constraint is in conflict with some other constraints, for instance, the constraint proposed in (19). Take [pan] (signboard) for example. In the retroflex suffixation, the consonantal ending [n] is lost. The [n] deletion in the output form obviously results mainly from the effects of the two constraints proposed above. Specifically, the constraint *r/C prohibits the dental nasal to be retroflexed and the constraint R-ANCHORING-r excludes the syllable to end with a non-retroflexed segment. It is obvious that these two constraints must dominate the constraint

[9] Following Wang (1993), Duanmu (1990) and J-L.Wang (1997b), the constraint can be stated as *r/coronal, i.e., a coronal segment is incompatible with retroflexion. This constraint can effectively prevent the coronal segments from being retroflexed but at present it needs further study at least in three aspects: ①More phonetic experiments are needed to be done to see whether the coronal segments like [ʧ] or [a] are also retroflexed. If evidence shows that what Wang (1997b) and Yang (1991) claim that [a] is not retroflexed is true, *r/coronal is better than the constraint I proposed here provisionally. But there has not been enough evidence up to now to show [ʧ] or/and [a] is/are not retroflexed in the Mandarin retroflex suffixation. ②The OO identity constraint, S-IDENT(r), triggers the spreading of the retroflex feature to the segments in the stem syllable, which of course include all the consonantal segments in the onset. But as Wang & He (1985) report, the retroflexion in Mandarin starts nearly simultaneously from the beginning of the rime, including all the non-high vowels and the back high vowel [u]. It is questionable, therefore, whether the constraint, *r/coronal, can exclude all the coronal segments, including coronal consonantal segments to be retroflexed. Of course, this can be solved by further specifying the constraint, like *r/V[coronal]. ③It needs to be confirmed by more experimental researches whether the non-coronal nasal, the velar nasal [ŋ], is retained or deleted after nasalizing the nuclear vowel. If it is really deleted as many scholars including Duanmu (1990) predict, then the constraint *r/coronal must be replaced by some other constraints.

ALIGN-Stem-R, which is proposed in (7d). The conflict of these three constraints can be illustrated in the following tableau.

(22)

	*r/C	R-ANCHORING-r	ALIGN-Stem-R
a. pa^{r}n^{r}	*!		
b. pa^{r}n		*!	
☞c. pa^{r}			*

The above tableau shows that the constraints *r/C and R-ANCHORING-r must dominate the constraint ALIGN-Stem-R. And otherwise an unexpected output will be obtained.

From this analysis we know that these three constraints can be partially ranked as: *r/C, R-ANCHORING-r >> ALIGN-Stem-R. Note that this partially ranked constraint set is also applicable in obtaining the expected retroflexed forms like [tɕi☜^{r}] (chicken, 鸡儿). The interaction of these constraints in obtaining [tɕi☜^{r}] from its lexical form [tɕi+☜^{r}] can be shown in the following tableau.

(23)

Input: tɕi☜^{r}	*r/C	R-ANCHORING-r	ALIGN-Stem-R
a. tɕi^{r}	*!		
b. tɕi		*!	
☞c. tɕi☜^{r}			*

Now let's look at the retroflexion of the mid vowels. It has been recognized that the retroflexed forms for the single mid vowel rime and the rime composed of a mid vowel and a consonantal coda are different. Consider the retroflexed forms [k☯^{r}] (root, 根儿) and [kɤ^{r}] (song, 歌儿). These two surface forms derive respectively from the retroflexion of [k☯n] and [kɤ]. For the latter case, the optimal output [kɤ^{r}] is easily derived through the constraint hierarchy adopted in (17) and (18). But for the former case, things are a little more complicated since it involves the laxing of the mid vowel.

I have discussed in Section 4.2.3 that the tense mid vowel does not occur before a nasal consonant or a glide and proposed the constraint *V$_{[+mid, tense]}$C/G to account for the laxing of the mid vowel before a coda consonant or a glide. It is obvious here that the nuclear mid vowel [k☯n] is triggered by its coda consonant [n] to become schwa first and then the trigger, the coda consonant [n], is demolished under the retroflexion. Therefore, the linguistically significant generalization, the rule that brings about the laxing of the mid vowel before the nasal consonant, is non-surface-apparent. Here, I argue that the phonological opacity of this kind can also be accounted for through sympathy. We can adopt the similar way of treatment to that in (18) to obtain the expected surface form [k☯^{r}]. As we have observed above, in Mandarin Chinese the sympathy selector is the faithfulness constraint ✭MAX-IO, which requires that all the lexical segments of the input should be preserved in the ❁-candidate. The sympathetic faithfulness

constraint is IDENT-☸O(tense), which requires the candidates to be faithful to the ☸-candidate in terms of the mid vowel tenseness. The selection of the optimal candidate [k☯ʳ] (root, 根儿) is illustrated as in (24).

(24)

Ipt: kɤn.☯ʳ	*r/C	R-A	*VCC	I-☸O(r)	I-☸O(t)	★M-IO	*VC	S-I(r)	I-IO(r)
a.kɤʳn☯ʳ			*!		*	√	*	**	*
☸b.k☯ʳn☯ʳ			*!			√		**	*
c.k☯n☯ʳ			*!	*		√		***	*
☜d. kɤʳ					*!	**		*	*
☞e. k☯ʳ						**		*	*
f. kɤ		*!		*	*	**		**	
g. k☯		*!		*		**		**	
h. k☯ʳnʳ	*!					*		*	*
i. k☯ʳn		*!				*		**	*

Look at the above tableau. The first three candidates all satisfy the sympathy selector ★MAX-IO and thus can be viewed as the potential ☸-candidate. But comparatively, (24b) incurs less violations of the constraints than (24a) and (24c) and hence is finally chosen as the ☸-candidate. (24d) is a transparent candidate because it is transparent in terms of the tenseness of the mid vowel. (24f) and (24g) both violate the higher ranked constraint that retroflex must be anchored at the right periphery of the output and besides they both violate the sympathetic faithfulness constraints here: IDENT-☸O(retroflex) and IDENT-☸O(tense) and hence drop out of the competition. (24h) and (24i) are excluded due to the fatal violations of the high-ranked constraints *r/C and R-ANCHORING-r. Finally, (24e), [k☯ ʳ], is chosen as the optimal candidate since it minimally violates or best satisfies the constraint hierarchy proposed here.

The same is true for the retroflexion of [kun]. As we know, evidence shows that the retroflexed form of [ku] (drum, 鼓) is quite different from that of [kun] (stick, 棍). The former is [kuʳ] while the latter is [kuʳ☜ʳ]. Clearly, it is easy to obtain the expected output for the retroflexed form of [ku] but it is not so easy for [kun].

It has been demonstrated in Section 4.2.3 that [kun] alternates with [ku☜n] under the effects of moraic constraints and these possible surface forms come form the corresponding lexical form [kuɤn] through the mid vowel reduction triggered by the *$V_{[+mid, +tense]}$C/G. It is obvious that the retroflexed form of [kuʳ☜ʳ] derives before moraic constraints take effects in eliding the reduced mid vowel. The process of obtaining the retroflexed form [kuʳ☜ʳ] is illustrated as below.

(25)

Ipt: kuɤn.☯^{r}	*r/C	R-A	*VCC	I-☸O(r)	I-☸O(t)	⋆M-IO	*VC	S-I(r)	I-IO(r)
a.ku^{r}ɤ^{r}n☯^{r}			*!		*	√	*	**	*
☸b.ku^{r}☯^{r}n☯^{r}			*!			√		**	*
c. ku☯n☯^{r}			*!	*		√		***	*
☜d. ku^{r}ɤ^{r}					*!	**		*	*
☞e. ku^{r}☯^{r}						**		*	*
f. kuɤ		*!		*	*	**		**	
g. ku☯		*!		*		**		**	
h. ku^{r}☯^{r}n^{r}	*!					*		*	*
i. ku^{r}☯^{r}n		*!				*		**	*

Now let's look at the relations between the moraic markedness constraints and the constraints proposed here to account for the retroflex suffixation in Mandarin Chinese. It is suggested in Ma (2001b) that one of the functions in applying the moraic markedness constraints is to identify which segments are linked to the moras and which are not. The segments linked to moras are assumed to be part of the rime; others are assumed to be of the margins or, in Mandarin, of the onset. Take [tɕi] for example. [i], not [tɕ], is linked to the moras because *MORA[CONS] is ranked higher than *MORA[HIGH]. But problems arise when the moraic constraints apply to its retroflexed form [tɕi☯^{r}]. If the [i] in the retroflexed form of [tɕi] is a high vowel, i.e., the actual retroflexed surface form is [tɕi☯^{r}], moras will be associated with [i] and [☯^{r}], as illustrated as in (26).

(26)

Input: tɕi☯^{r} + $^{\mu\mu}$	*MORA[GLIDE]	*MORA[SCHWA]	*MORA[HIGH]
☞a. tɕi$^{\mu}$☯$^{r\mu}$		*	*
b. tɕi☯$^{r\mu\mu}$		**!	

And if the [i] in the retroflexed form of [tɕi] has become a high glide, i.e., the actual retroflexed surface form is [tɕy☯^{r}], things will become a little more complicated. Only the moraic constraints cannot bring about [tɕy☯^{r}] as the expected optimal output, as illustrated as in (27).

(27)

Input: tɕi☯^{r} + $^{\mu\mu}$	*MORA[GLIDE]	*MORA[SCHWA]	*MORA[HIGH]
a. tɕy$^{\mu}$☯$^{r\mu}$	*!	*	
√ b. tɕy☯$^{r\mu\mu}$		**!	
? ☞ c. tɕi$^{\mu}$☯$^{r\mu}$		*	*

The moraic constraints in (27) selects [tɕi$^{\mu}$☯$^{r\mu}$] as the optimal candidate, which runs against what we expect.

The above analysis shows that whether [i] surfaces as a glide or as a vowel cannot be determined only by the moraic constraints. Therefore we must first make clear whether [i] phonetically has become a glide or remains as a high vowel. But up to now no agreement has been

reached on this aspect. Here I want to present the possible solutions to the two cases of [i] in its retroflexion. If [i] surfaces still as a high vowel, i.e., the actual retroflexed surface form is [tɕi☯ ʳ], the moraic constraints apply just as suggested in (26) above. If [i] surfaces as a high glide, i.e., the actual retroflexed surface form is [tɕy☯ ʳ], only the moraic constraints cannot explain how the underlying high vowel [i] becomes a high glide [y] in the retroflex suffixation. The possible solution to it is that there is another constraint ranked over the moraic constraints. The constraint can be written probably as in (28).

(28) ALIGN-L (r, μ): The left edge of the retroflexed segment of a syllable must be aligned with the mora of the left side.

The constraint proposed in (28) is an alignment constraint requiring the retroflexed segment in a syllable must be (part) of the rime. If a segment is not retroflexed in the retroflexion, it will not be relegated to the rime of the syllable. Obviously, in the above case, this constraint must be ranked over the moraic constraints. Putting this constraint into the constraint hierarchy in (27), the expected result will be obtained.

(29)

Input:tɕi☯ʳ + μμ	ALIGN-L (r, μ)	*MORA[GLIDE]	*MORA[SCHWA]	*MORA[HIGH]
a. tɕy$^{\mu}$☯$^{r\mu}$	*!	*	*	
☞b. tɕy☯$^{r\mu\mu}$			**	
c. tɕi$^{\mu}$☯$^{r\mu}$	*!		*	*

The employment of the constraint I posit in (28) effectively excludes the candidate [tɕi☯ʳ] as the optimal one and thus identifies [tɕy☯ʳ] as the only surface output for the retroflexed form of [tɕi] in Mandarin Chinese.

Another point drawn from the above analysis is that the moraic constraints must rank below the constraints concerning the retroflex suffixation. This is because the insertion of the moraic segment [☯ʳ] in the retroflexed form [tɕy☯ʳ] is triggered by those constraints relevant to the retroflexion. To put all these constraints together, the illustration in (29) will be elaborated in the following tableau.

(30) 鸡儿

Input: tɕi+☯ ʳ+μμ	*r/I	R-AN-r	AL-STEM-R	AL-L-μ	*M[y]	*M[☯]	*M[i]
a. tɕi $^{r\mu\mu}$	*!						**
b. tɕi $^{r\mu}$☯ $^{r\mu}$	*!		*			*	*
c. tɕi$^{\mu}$☯ $^{r\mu}$			*	*!		*	*
d. tɕi$^{\mu\mu}$		*!					**
☞e. tɕy☯ $^{r\mu\mu}$			*			**	
f. tɕy$^{\mu}$☯ $^{r\mu}$			*	*!	*	*	

Note that the moraic constraints are involved in obtaining the expected retroflexed output of [tɕi].

The requiring of moraic constraints in obtaining the retroflexed form can be further supported by the following example [un].

As some scholars point out, [un] appears sometimes as [u☜n] and sometimes as [un] (Lin 1989). I have discussed how to describe the rime of a syllable with schwa (as in [w☜n]) and without schwa (as in [tun]) within the OT framework and pointed out that the laxing of the lexical mid vowel [ɤ] results in the former case and the moraic constraints bring about the result that the laxed mid vowel [☜] in [u☜n] is elided (Ma 2001b:Chapter 4). I claimed there that schwa alternation in the syllables of this type can be accounted for with reference to the interaction of $^{*}V_{[+mid,\ +tense]}C$ and the moraic markedness constraints. Now the problem is how to account for the retroflexed forms of these syllables from the OT's point of view.

Consider the syllables [w☜n] and [tun]. Evidence shows that these two syllables have the same retroflexed rimes [u☯$^{\mu}$]. This shows that the laxing of the mid vowel takes place before the retroflexion and the retroflexion before the assigning of the moras to the segment(s) in the rime. But in the teleological framework of OT, all the possible candidates freely generated on the basis of the input simultaneously feed into the evaluator composed of a set of ranked constraints for assessment. The candidate that least violates the set of constraints provided by the evaluator will be selected as the optimal candidate of the input. Therefore the concept of the serial stages of derivation, i.e., the ordering of the rule applications, is totally insignificant. What is most significant is to identify which of the constraints is more active in assessing the linguistic surface forms. The more active constraint, the higher it is ranked in the constraint hierarchy.

As for the constraints related with the retroflexion of the syllable with [un] as its rime, the moraic markedness constraints must be ranked lower than the constraints relevant to the retroflexion. Otherwise the retroflexed mid vowel [☯^{r}] will not surface. To put it in another way, the constraints concerning the retroflex suffixation takes effect before the moraic markedness constraints, the constraints determining whether or not the [☯] in [un] turns up in the surface forms, come into effect.

Now the problem is the phonetic realization of [u] in the retroflexed form of [tun]. Specifically, if [u] is retroflexed, it is phonetically realized as [u]; if it is not, it is [w]. Such a theoretical prediction needs to be attested by more acoustic experiments.

If experiments attest that the retroflexed form of [tun] is [tu^{r}☯^{r}], the constraints proposed above are well enough in obtaining it. The evaluation process between the two competing candidates [tu^{r}☯^{r}] and [tw☯^{r}] can be partially illustrated as below.

(31)

	*MORA[SCHWA]	*MORA[HIGH]
☞a. tu$^{r\mu}$☯$^{r\mu}$	*	*
b. tw☯$^{r\mu\mu}$	**!	

Please note that which of these two competing candidates is more harmonic can be determined by the partially ranked constraint hierarchy employed in (24) and (25). Hence the task of determining which is the optimal candidate naturally passes down to the lower ranked moraic constraints

presented in (31). Here it has been proposed that these constraints can be approximately ranked as: *r/C >> *r/I, R-ANCHORING-r >> *VCC]$_{\sigma}$ >> ALIGN-Stem-R, IDENT-☸O(r), IDENT-☸O(t) >> ★MAX-IO, *V$_{[+mid, +tense]}$C/G, ALIGN-L(r, μ) >> *MORA[SCHWA] >> *MORA[HIGH] >> S-IDENT(r) >> IDENT-IO(r).

But if the retroflexed form of the [tun] is not [tu^{r}☯^{r}] but [tw☯^{r}], the constraints proposed above are not well enough in obtaining it and we are compelled to appeal to some other constraints. A possible solution, I suggest here now, is to define S-IDENT(r) as S-IDENT-SS(r) instead of S-IDENTμ(r) or S-IDENTσ(r). Specifically, to restrict the spreading of the suffixal retroflex feature to the segments of the stem, the syllable-to-syllable S-IDENTσ(r) or rime-to-rime S-IDENTμ(r) used here must be replaced here by the segment-to-segment S-IDENT-SS(r).

Theoretically, all the IDENT, MAX and DEP constraints are members of the correspondence constraint family. MAX and DEP constraints are different from IDENT constraints in that MAX and DEP require the faithfulness in the number of segments while IDENT requires the faithfulness in the content of the segments. These constraints are originally proposed to require the faithfulness between input and output (McCarthy & Prince 1993) but now have been extended to express the faithfulness between output and output (Benua 1995; Flemming 1995; McCarthy 1995; Burzio 1996; Kenstowizc 1996; Steriade 1996) and even between candidate and candidate (as in Sympathy, developed by McCarthy 1998, 2000). As for the IDENT constraints, the constraints requiring the faithfulness between the segments in terms of their feature(s), usually hold over the one-to-one mappings, that is, mappings between one and one segment. Pulleyblank (1997), in accounting for the consonantal place assimilation, proposes IDENTICAL CLUSTER CONSTRAINTS (often shortened as ICC). Significantly, assimilation is analyzed as an effect of correspondence constraints on more than two distinct elements within one representation. Inspired by this idea, Krämer (1998, 2000) develops SYNTAGMATIC IDENTITY constraints to account for vowel harmony in some languages.

There is one point that needs to be addressed here. Our analyses show that there are two types of harmony (or assimilation) in terms of some feature(s): ①harmony between two adjacent segments and ②harmony among a cluster of segments. As for the former case, it is claimed here that S-IDENT-SS (r) can be employed to account for it. Take the retroflexed form of [xua] (flower, 花儿) for example. If only the last segment [a] is subject to be retroflexed and [u] is not, then it can be assumed that the retroflexion takes place only between two adjacent segments. Technically, S-IDENT-SS (r) is deemed to apply. But if [u] and [a] all become retroflexed, then it can be considered that the retroflexion occurs among a cluster of segments, or within a syllable domain[10]. Technically, S-IDENTσ(r) is supposed to apply. In short, Optimality Theory can well account for presumably two possible types of retroflexion. Which type of treatment should be adopted in accounting for the Mandarin retroflex suffixation all depends on the phonetic evidence provided by authentic acoustic experiments.

Now let us look at the retroflexion of apical vowels. I claimed in Ma (2001b) that apical vowels are lexically absent and demonstrated how the lexically absent apical vowels are derived

[10] The fact that consonants in the onset and nasals or front high glides in the coda are not retroflexed can be considered to be due to the incompatibility with retroflex articulation.

within the OT framework. It is obvious that retroflexion goes before the epenthesis of apical vowels. Take 丝儿 (ʂ☜^{r}, silky) for example. The retroflexion process can be illustrated as in the following tableau.

(32) 丝儿

Input: ʂ+☜^{r}	R-ANCHOR-r	ALIGN-STEM-R	MAX-IO	IDENT-IO(F)
☞a. ʂ☜^{r}		*		
b. ʂ	*!		*	
c. ʂu^{r}		*		*!

The tableau (32) shows that the lexically empty rime (or nucleus) cannot be filled in by any other retroflexed vowel (like u^{r} in 32c) than the retroflexed mid vowel [☜^{r}] because the faithfulness constraint IDENT-IO(F) penalizes the segment with the features not present in the input. Recall the constraint set I proposed in Section 4.3.2. The faithfulness constraint IDENT-IO(F) is highly ranked, or ranked above the set of markedness constraints responsible for the apical vowel epenthesis. Here, constraints triggering the epenthesis (precisely, retaining) of the retroflexed mid vowel apparently are ranked above the faithfulness constraint IDENT-IO(F). In doing so, the expected output can thus be obtained.

After discussing the retroflexion of apical vowels in Mandarin, let us come to another case of Mandarin retroflex suffixation, the retroflexed form of the rime ending with a velar nasal.

It is commonly known that in Mandarin the nasal consonantal coda [n] is deleted when the syllable is retroflexed while the velar nasal [☠] nasalizes the nuclear vowel before it is deleted in retroflex suffixation. As I have pointed out in Section 6.2, different proposals have been advanced so far to account for it. Most of the proposals (e.g., Duanmu 1990; Wang 1993; Wang 1997b) attribute the retaining of the nasal feature in the syllable ending with a velar nasal to the compatibility articulation. Specifically, the constraint *r/coronal that they proposed[11] disallows [n] but allows [☠] to surface as a coda in the retroflexed syllable. There are at least two problems with this kind of phonological explanation. ①It seems to be arbitrary to make such a generalization that [☠] is retained and [n] is deleted through the constraint *r/coronal since many studies have shown that the coronal vowels, say [a], can be retroflexed. ②Many acoustic experiments show that the segment [☠] is also deleted in retroflex suffixation. To assume that it is retained, more acoustic experiments must be provided.

I will argue here that neither of these two nasal consonants is retained in retroflex suffixation and the nasalized vowel in the retroflexed form actually results from the higher ranking of some universal constraint related with the velar nasal.

It is observed in many languages that vowels tend to be nasalized before a tautosyllabic nasal consonant. As we know, there are two nasal consonants in Mandarin, [n] and [☠]. Acoustic experiments show that phonetically vowels tend to be nasalized before these two nasal consonants.

[11] Specifically, the constraint that Wang (1993) proposes is that the retroflex feature [pos] is incompatible with the coronal feature [dis]. If the coronal feature [dis] is lexically associated with the root node, in retroflex suffixation this feature will be de-linked with the root node and be replaced by the retroflex feature [pos]. The constraint *[+retroflex, +posterior] that Wang (1997b) proposes is quite similar in nature to the constraint proposed here.

These can be expressed by the following two markedness constraints.

(33) a. *$V_{ORAL}N$ b. *V_{ORAL}☠

These two constraints rule out oral vowels that precede a tautosyllabic nasal. Consider that nasal vowels are "marked" as compared to oral vowels and many languages have only oral vowels but no nasal vowels (Maddieson 1984). This universal tendency can be expressed by a markedness constraint in (34).

(34) *V_{NASAL}

It is obvious that in Mandarin the constraint in (34) is in conflict with the constraints in (33). Evidence shows that in Mandarin these three constraints must be ranked as *V_{ORAL}☠, *$V_{ORAL}N$ >> *V_{NASAL}. Such a ranking results in nasal vowels surface before a tautosyllabic nasals. Now the problem is how these constraints take effects in obtaining the expected retroflexed forms in Mandarin.

Take [pan] and [pa☠] for example. In (22), I demonstrated how the expected retroflexed form [pa^{r}] is obtained from the lexical form [pan.☜^{r}] without considering the constraint *$V_{ORAL}N$ but the analysis clearly shows that the constraint *$V_{ORAL}N$ must be ranked below the constraints involved in obtaining the optimal retroflexed output. On the contrary, *V_{ORAL}☠ must be ranked comparatively higher since the retroflexed nuclear vowel is nasalized before the velar nasal is deleted. The retroflexion of [pa☠] can be illustrated as in the following tableau.

(35)

Ipt:pa☠+☯^{r}	*r/C	R-A	*VCC	I-☸O(r)	I-☸O(n)	★M-IO	*V☠	S-I(r)	I-IO(r)
a.pa^{r}☠☯^{r}			*!		*	√	*	**	*
☸b.pã^{r}☠☯^{r}			*!			√		**	*
c. pã☠☯^{r}			*!	*		√		***	*
☜d. pa^{r}					*!	**		*	*
☞e. pã^{r}						**		*	*
f. pa		*!		*	*	**		**	
☜g. pã				*!		**		**	
h. pã^{r}☠^{r}	*!					*		*	*
h. pã^{r}☠^{r}		*!				*		**	*

(35) shows that the vowel is nasalized before a tautosyllabic velar nasal consonant is deleted due to the dominating conflicting forces maintaining the retroflex feature at the right periphery of the suffixed syllable. It is obvious that such a non-surface- apparent phonological phenomenon can be accounted for only by means of Sympathy Theory as I have discussed above. Another point that deserves our attention is that in Mandarin the constraint *V_{ORAL}☠ is definitely ranked higher than the constraint *$V_{ORAL}N$ and thus the three markedness constraints proposed in (33) and (34) are

ranked as follows:

(36) *V_{ORAL}☠ >> *V_{ORAL}N >> *V_{NASAL}

Having discussed the major aspects of the retroflex suffixation, let us see some minor problems in obtaining the expected retroflexed forms for Mandarin.

It should be reminded that the constraints employed here to account for mandarin retroflex suffixation are not inclusive and, due to the limitation of space, some of the constraints are not fully discussed here. For instance, LINEARITY-IO[12] is obviously also active but not crucial in Mandarin retroflex suffixation. This will become clearer if we take more candidates into consideration in obtaining the optimal output presented above in (36).

(37)

Ipt: pa☠+☯^{r}	LIN-IO	*VCC	I-❁O(r)	I-❁O(n)	★M-IO	*V☠	S-I(r)	I-IO(r)
a. pa^{r}☠☯^{r}		*		*!	√	*	**	*
❁b. pã^{r}☠☯^{r}		*			√		**	*
c. pã☠☯^{r}		*	*!		√		***	*
d. pa^{r}☯^{r}☠	*!			*	√		***	*

(37) shows how the ❁-candidate is selected when a new candidate (37d) is involved. Here, (37a), (37c) and the new candidate (37d) all incurs more serious violations than the candidate (37b) and thus (37b) is chosen as the ❁-candidate for the retroflexed form of [pa☠].

It must be noted that the constraint hierarchy concerning the nasality of Mandarin vowels proposed here in (36) indicates Mandarin vowels tend to be nasalized before a tautosyllabic nasal (either before a [n] or before a [☠]), as M-C. Lin and J-L. Wang pointed out to me (personal communication). Obviously, none of the proposals made before by the Chinese scholars have made such a precise prediction in this respect. When the constraint hierarchy in (36) is incorporated into the constraint hierarchy triggering the retroflex suffixation, the retroflexed oral vowels of the syllable ending with [n] and the retroflexed nasal vowels of the syllable ending with [☠] are equally well represented. That is, the constraint *V_{ORAL}☠ is ranked higher, having already taken effects before the constraints triggering the deletion of [☠] in the retroflex suffixation, while the constraint *V_{ORAL}N is ranked lower, failing to take effects before the deletion of [n] in the retroflex suffixation.

4. Concluding Remarks

Wang (1993), within the feature geometrical framework, proposes her featural treatment of the retroflex suffixation in Beijing Mandarin. In this paper, I argued against this treatment,

[12] LINEARITY-IO, proposed by McCarthy & Prince (1995:13), Pater (1999:315) and others, states that the output reflects the precedence structure of the input, and vice versa. Specifically, this constraint is violated if the precedence structure of the input is changed in the output.

claiming that Wang's featural treatment cannot predict the typological variations of retroflex suffixation across Chinese languages. In Section 1, I pointed out that it is widely acknowledged that, morphologically and diachronically, the Mandarin retroflex suffix apparently derives from the morpheme [☜ʳ] and thus it is empirically evident and theoretically significant to treat the Mandarin retroflex suffix as a monosyllabic morpheme in the phonological analysis. Within the OT framework, I proposed that the three types of retroflex suffixation in Chinese languages reported by Lu (1995) can be easily dealt with by the following different rankings of four universal constraints as is shown in (8).

It is clear that the r-suffixation in Beijing Mandarin belongs to the third type in which the monosyllabic morpheme is incorporated into the suffixed stem due to the higher ranking of the two markedness constraints.

Then, in Section 3.2, I discussed how the rime of the suffixed stem becomes retroflexed and proposed that the rime retroflexion is triggered by the assimilation- oriented OO-correspondence faithfulness constraint IDENT-SA(r), which can be subsumed in the family constraint S-IDENT(F). I also argued that the non-surface- apparent opacity, which occurs in the retroflex suffixation, can be dealt with only with reference to the Sympathy Theory proposed by McCarthy (1998). I demonstrated and pointed out that the sympathy selector for Mandarin is ✶MAX-IO, the sympathetic faithfulness constraint is IDENT-☸O(r) and the sympathetic candidate, take the 鼓儿 ([kuʳ] drum) for example, is ☸[kuʳ☯ʳ].

Finally, in Section 3.3, I made a detailed analysis of different aspects of Mandarin Chinese within the OT framework. My analysis ranges over the following different aspects:

①In discussing the compatibility of retroflexion with the articulation of different segments, I posited a set of ranked markedness constraints *r/C >> *r/I >> *r/E >> *r/A >> *r/U in (21) and claimed that the problem of how to define compatible articulation in a class of segments can be easily solved by determining the specific place of the constraints triggering retroflexion in this constraint hierarchy. Specifically, if Mandarin allows the front high vowels to be retroflexed, then the constraints triggering the retroflexion can be presumably intervened before the constraint *r/I; if Mandarin only allows the back vowel to be retroflexed, it can be assumed that the constraints triggering the retroflexion take effects before the constraint *r/U involves in, and so on.

②In discussing the ranking of the above constraints, I showed that the constraint set *r/C >> *r/I and the constraint *r/C, R-ANCHORING-r must be ranked above the ALIGN-Stem-R. Such a ranking can block the nasal consonant or the front high vowel / glide to be the coda of a retroflexed output form.

③Then, I discussed the case in which the consonantal coda is conceived as the trigger of the mid vowel laxing. It is pointed out that the coda, which triggers the laxing of the mid vowel in the nucleus, is ultimately eliminated in the process of retroflex suffixation. I showed in detail how it is accounted for by way of sympathy selector.

④Next, I discussed the role of moraic constraints in making the lexical high vowel become a correspondent glide when the syllable is retroflex suffixed. Two possible solutions were provided: if, for instance, [i] in [tɕi] surfaces still as a high vowel in its retroflexed form, i.e., the actual retroflexed surface form is [tɕi☯ʳ], no other constraints are needed; if [i] surfaces as a high glide, i.e., the actual retroflexed surface form is [tɕy☯ʳ], only the moraic constraints cannot explain how

the underlying high vowel [i] becomes a high glide [y] in the retroflex suffixation. The other constraint assumed to be involved in the latter case is ALIGN-L (r, μ) proposed in (28).

⑤Besides these two possible surface forms, another possible surface forms are also discussed. For instance, if the retroflexed form of the [tun] is not [tu^{r}☯^{r}] but [tw☯^{r}], a possible solution for it, I suggested, is to define S-IDENT(r) as the segment-to-segment S-IDENT-SS(r) instead of the rime-to-rime S-IDENTμ(r) or the syllable-to-syllable S-IDENTσ(r).

⑥The retroflexion of apical vowels is also discussed. I claimed that the retroflexion of apical vowels can be easily dealt with within the OT framework. Specifically, it can be accounted for through ranking the constraints that trigger the retroflex suffixation over the constraints that trigger the epenthesis of an apical vowel.

⑦As for the retroflex suffixation of the syllable ending with a velar nasal, I proposed that the constraint *V_{ORAL}☠ triggering the vowel nasalization is more active or ranked higher in Beijing Mandarin while the constraint *V_{ORAL}N triggering the vowel nasalization is less active or ranked lower.

So far, I have discussed all different aspects of Mandarin retroflex suffixation. Using the constraint hierarchy proposed above, we can give a rather satisfactory and convincing formal account for different cases in the retroflex suffixation in Mandarin Chinese.

Notes: The present paper is part of my Ph. D. dissertation submitted in May 2001 to the Graduate School of Beijing Normal University. Here, I would like to express my sincere thanks to Professor *Zhou Liuxi*, Professor *Wang Jialing*, Professor *Lin Tao*, Professor *Lin Maocan*, Professor *Lin Shunwu*, Professor *Zhou Tongchun* and many others for their invaluable comments and instructive guidance.

References

Archangeli, Diana. 1984. *Underspecification in Yawelmani Phonology and Morphology*. Ph.D. dissertation. MIT. Published in 1988. New York and London: Garland.

Archangeli, Diana. 1988. Aspects of underspecification theory. *Phonology* 5:183-207.

Bao, Huai-Qiao. (鲍怀翘) 1989. 元音. 载吴宗济、林茂灿主编《实验语音学概要》。73-111. 北京：高等教育出版社。

Benua, Laura. 1995. Identity effects in morphological truncation. To appear in J. Beckman, S. Urbanczyk & L. Walsh, (eds.) *University of Massachusetts Occasional Papers in Linguistics* 18: *Papers in Optimality Theor*y, Graduate Linguistic Student Association, Amherst, MA. (also ROA-74).

Burzio, Luigi. 1996. Surface constraints versus underlying representation. In Jacques Durand & Bernard Laks, (eds.) *Current Trends in Phonology: Models and Methods*. 123-142. Manchester: European Studies Research Institute, University of Salford.

Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of California Press.《汉语口语语法》，赵元任著、吕叔湘译，北京：商务印书馆，1979 年。

Cheng, Chin-Chuan. 1973. *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.

Cheng, Robert L. 1966. Mandarin phonological structure. *Journal of Linguistics* 2: 135-158.

Clements, George N. 1985. The geometry of phonological features. *Phonology* 2: 225-252.

Clements, George N. & S. Jay Keyser. 1983. *CV Phonology: A Generative Theory of the Syllable*. Cambridge (Mass.): MIT Press.

Duanmu, San. 1990. *A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages*. Ph.D. dissertation, MIT.

Flemming, Edward. 1995. *Auditory Representations in Phonology*. Ph.D. dissertation. University of California, Los Angeles.

Fu, Yi-Chin. 1980. The /-r/ suffixation and the phonological structure of Mandarin features. In P-J. Li et al. (eds.) *Paper in Honor of Professor Lin Yu-K'eng on Her Seventieth Birthday*. Taipei: Wenshin Publishing Co.

Goh, Yeng-Seng. 1997. *The Segmental Phonology of Beijing Mandarin*. Taipei: The Crane Publishing Co. LTD.

Goldsmith, John. 1976. *Autosegmental Phonology*. Ph.D. dissertation, MIT.

Hockett, Charles F. 1950. Peiping morphophonemics. *Language* 26:63-85.“北京话形态音素学”，孟琮译，吕叔湘校，《国外语言学》1980 年第 5 期 21-26 页、第 6 期 25-34 页。

Hsueh, F-S. 1985.《国语音系解析》(An Anatomy of the Pekingese Sound System). Taipei: Student Book Co., Ltd.

Hsueh, F-S. (薛凤生) 1986.《北京音系解析》。北京：北京语言学院出版社。

Kenstowicz, Michael. 1996. Base-Identity and Uniform Exponence: Alternatives to Cyclicity. In Jacques Durand & Bernard Laks, (eds.) *Current Trends in Phonology: Models and Methods*. 363-393. Manchester: European Studies Research Institute, University of Salford. ROA-103.

Kiparsky, Paul. 1971. Historical linguistics. In W. Dingwall, (ed.) *A Survey of Linguistic Scienc*e. College Park: University of Maryland Press.

Kiparsky, Paul. 1973. Phonological representations. In Osamu Fujimura, (ed.) *Three Dimensions of Linguistic Theor*y. 1–136. Tokyo: Taikusha.

Krämer, Martin. 1998. A correspondence approach to vowel harmony and disharmony. *SFB 282 Working Paper 107*. Heinrich-Heine-Universität Düsseldorf. ROA-293.

Krämer, Martin. 2000. Yucatec Maya vowel alternations: harmony as Syntagmatic Identity. ROA-423.

Li，Si-Jing. (李思敬) 1986.《汉语“儿”[㘝]音史研究》。北京：商务印书馆。

Li，Si-Jing. (李思敬) 1990. 汉语普通话儿化音两种构音方式的语音实验.《王力先生纪念论文集》，130-145. 北京：商务印书馆。

Lin, Tao. (林焘) 1982. 北京话儿化韵的个人读音差异问题.《语文研究》第 2 辑。又载于林焘著，《语音探索集稿》。1990:61-70. 北京：北京语言学院出版社。

Lin, Tao & Jiong Shen. (林焘、沈炯) 1995. 北京话儿化韵的语音分歧.《中国语文》第 3 期第 170-179 页。又载于中国语文编辑部编，《庆祝中国社会科学院语言研究所建所 45 周年：学术论文集》，1997:259-269. 北京：商务印书馆。

Lin, Tao & Li-Jia Wang. (林焘、王理嘉) 1992. 《语音学教程》。北京：北京大学出版社。

Lin, Yen-Hwei. 1989. *Autosegmental Treatment of Segmental Process in Chinese Phonology*. Ph.D. dissertation, The University of Texas as Austin.

Lu, Yun-Zhong. (鲁允中) 1995.《普通话的轻声和儿化》，北京：商务印书馆。

Ma, Qiu-Wu. (马秋武) 1999. 管辖音系学与汉语普通话音节可能性组合的音系分析。载吕士楠等主编《现代语音学论文集》。173-178. 北京：金城出版社。

Ma, Qiu-Wu. (马秋武) 2000a. 管辖音系学：一种以制约为基础的音系学理论. 《解放军外国语学院学报》第 1 期第 15~20 页。

Ma, Qiu-Wu. (马秋武) 2000b. 管辖音系学的基本理论及最新发展. 《当代语言学》第 4 期第 218~226 页。

Ma, Qiu-Wu. (马秋武) 2001a. 后 SPE 音系学理论的发展取向. 《外国语》第 3 期第 15~22 页。

Ma, Qiu-Wu. 2001b. Interaction of Constraints on Mandarin Syllable Structure. Ph.D. dissertation, Beijing: Beijing Normal University.

Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge: Cambridge University Press.

McCarthy, John. 1988. Feature geometry and dependency: a review. *Phonetica* 43: 84-108.

McCarthy, John. 1995. Faithfulness in prosodic morphology and phonology: Rotuman Revisited. ROA-110

McCarthy, John. 1998. Sympathy and phonological opacity. ROA-252.

McCarthy, John. 2000. Sympathy, cumulativity, and the Duke-of-York gambit. ROA-315.

McCarthy, John & Alan Prince. 1993. *Prosodic Morphology 1: Constraint Interaction and Satisfaction*. Report No. RuCCS-TR-3. New Brunswick (NJ): Rutgers University Center for Cognitive Science.

McCarthy, John & Alan Prince. 1995. Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey & Suzanne Urbanczyk, (eds.) *Papers in Optimality Theory*. The University of Massachusetts Occasional Papers in Linguistics 18: 249-384.

McCarthy, John & Alan Prince. 1999. Faithfulness and identity in prosodic morphology. In R. Kager, Harry van der Hulst & Wim Zonneveld, (eds.) *The Prosody-Morphology Interface*. 1999:218-309. Cambridge: Cam- bridge University Press.

Pater, Joe. 1996. Austronesian Nasal Substitution and Other NC Effects. In René Kager, Harry van der Hulst & Wim Zonneveld, (eds.) *The Prosody- Morphology Interface*. 1999:310-345. Cambridge: Cambridge University Press. ROA-160.

Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. dissertation. MIT.

Pulleyblank, Douglas. 1988a. Underspecification, the feature hierarchy and Tiv vowels. *Phonology* 5:299-326.

Pulleyblank, Douglas. 1988b. Vocalic underspecification in Yoruba. *Linguistic Inquiry* 19:233-270.

Pulleyblank, Douglas. 1997. Optimality Theory and features. In Diana Archangeli & D. Terence Langendoen, (eds.) *Optimality Theory: An Overview*. 59-101. Oxford: Blackwell.

Pulleyblank, Edwin G. 1983. Vowelless Chinese? An application to the three tired theory of syllable structure to Pekingese. Paper presented for t*he XVI International Conference on Sino-Tibetan Languages and Linguistics at University of Wahington*.

Sagey, Elizabeth C. 1986. *The Representation of Features and Relations in Non-linear Phonology.* Ph.D. dissertation. MIT.

Steriade, Donca. 1996. Paradigm uniformity and the phonetics-phonology boundary. Papers presented at the 5th Conference in Laboratory Phonology. Illinois.

Tao, Y-P. & Y-X. Yin. (陶荫培、尹润芗) 1957. 略谈“儿化”. 《语文学习》10 月号 31-32 页。

Wang, Fu-Shi. (王辅世) 1963. 北京话韵母的几个问题. 《中国语文》2 月号。

Wang, Fu-Tang. (王福堂) 1999. 《汉语方言语音的演变和层次》。北京：语文出版社。

Wang, Jia-Ling. (王嘉龄) 1997b. 从音系学到语音学. 《外语教学与研究》第 4 期第 1-3 页。

Wang, Jia-Ling. (王嘉龄) 2000. 音系学百年回顾. 《外语教学与研究》第 1 期第 8-14 页。

Wang, L-J. & H-D. Wang. (王理嘉、王海丹) 1991. 儿化韵研究中的几个问题——与李思敬先生商榷.《中国语文》第 2 期第 96-103 页。

Wang, L-J. & N-J. He. (王理嘉、贺宁基) 1985. 北京话儿化韵的听辨实验和声学分析. 载林焘、王理嘉主编《北京语音实验录》北京：北京大学出版社。

Wang, Zhi-Jie. 1993. *The Geometry of Segmental Features in Beijing Mandarin*. Ph.D. dissertation, University of Delaware.

Xu, Shi-Rong. (徐世荣) 1960. 儿化韵的变化规律. 《语文学习》1 月号。

Xu, Shi-Rong. (徐世荣) 1999. 《普通话语音常识》。北京：语文出版社。

Yang, Shun-An. 1988. Preliminary synthesis of neutral-tone syllable in Standard Chinese. *Phonetic Laboratory Annual Report of Phonetic Research* 1988. Beijing Phonetic Laboratory, Institute of Linguistics, Chinese Academy of Social Sciences.

Yang, Shun-An. (杨顺安) 1991. 普通话儿化音节规则合成的初步研究. 《中国语文》第 2 期。

Yin, Yun-Mei. 1986. *An Autosegmental Approach to Retroflex Suffixation and Reduplication in Chinese*. Taipei: The Crane Publishing Co.

Yin, Yun-Mei. 1989. *Phonological Aspects of Word Formation in Mandarin Chinese*. Ph.D. dissertation, The University of Texas at Austin.

Yip, Moira. 1994. Morpheme-level features: Chaoyang syllable structure and nasalization. ROA-081.

Zhou, Jin-Ying. 1995. *An Optimality Theoretic Approach to R-retroflexion in Beijing Mandarin*. MA thesis. Tianjin: Tianjin Normal University.