Chapter 2

Literature Review

In this chapter, we briefly introduce OT and then refer to previous proposals and approaches in loanword phonology which are relevant to our analyses.

2.1 Optimality Theory

OT (Prince and Smolensky 1993, McCarthy and Prince 1995) proposes that Universal Grammar (UG) is composed of output constraints, and their different rankings among languages are reflected by the diversity in language forms. In the OT framework, language forms are generated through an input-output mechanism, where GEN(ERATOR), when applied to an input, produces an indefinite number of candidates, all of which are possible analyses of the input. When the other function we call EVAL(UATION) is applied to these output candidates, it produces the optimal output. The grammatical organization is schematically represented below:
EVAL comprises a series of universal, and violable constraints. In terms of ‘universality’, it is proposed that all constraints are part of UG. Namely, all constraints exist in natural languages. OT does not, however, assert that every constraint has an equal effect in all languages, but the constraints are language-specifically ranked. A constraint that is top-ranked in a language, which means it is never violated, can be totally ineffective in another language. ‘Parallelism’ of EVAL is also essential in OT: all the constraints pertaining to a certain type of structure evaluate and interact within a single hierarchy, with no intermediate level involved.

By ‘ineffective’, OT proposes that a constraint has the property of ‘violability’. Although a constraint can be violated, the violation must be kept to a minimum (a principle discussed later). Constraint rankings differ from language to language. Thus if a constraint is ranked on the top in language $\alpha$, a violating candidate cannot be the
optimal output. But when the constraint is ranked rather lower in language $\beta$, the violating candidate may still be a potential output. In this interpretation, the ineffectiveness of this constraint in language $\beta$ does not indicate that there is no such a constraint in the grammar of language $\beta$.

The criterion of evaluation is the principle of ‘Minimal Violation’. The candidate with no violation at all is the definite winner. When none of the candidates completely satisfies all the constraints, however, the optimal output is the one violating the lower-ranked constraint(s), or the one which incurs the least violations of a constraint (or constraints of the same rank). The evaluation is demonstrated by a ‘tableau’: a tableau lists two or more output candidates vertically in random order in the leftmost column, and constraints horizontally in a descending ranking from left to right on the top row, as represented below:

(2) Sample of a Tableau
Ranking: $C_1 >> C_2, C_3 >> C_4$

<table>
<thead>
<tr>
<th>input</th>
<th>$C_1$</th>
<th>$C_2$</th>
<th>$C_3$</th>
<th>$C_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. candidate 1</td>
<td>*!</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. candidate 2</td>
<td></td>
<td>*!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. candidate 3</td>
<td></td>
<td></td>
<td>*!</td>
<td></td>
</tr>
<tr>
<td>d. candidate 4</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>e. …</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In (2), candidate (1) violates the undominated $C_1$ and is ruled out. A violation is
indicated by an asterisk ‘∗’, the fatal violation is indicated by an exclamation mark ‘!’, and the shading in the cells indicates that the violation content is no longer relevant.

Candidate (2) and (3) are eliminated too for their violations of C2 and C3 in the second level. The separating dotted line stands for irrelevant ranking between C2 and C3.

Candidate (4) is selected as the optimal output, as marked by the index finger ‘<sup>−</sup>’), despite its insignificant violation of the low-ranked C4.

Based on the essential concepts, OT has two ‘families’ of constraints. A ‘markedness’ constraint requires that the output structure observe certain criterion of structural well-formedness. The phonological structures of languages are numerous and complex—a language may bear a large structural difference from another. Despite this, in sound systems, there are comparatively more ‘unmarked’ segments, segment combinations, or prosodic structures which are universally favored over others. Front unrounded vowels, for example, are more unmarked than front rounded vowels, short vowels are more unmarked than long vowels, CV is more unmarked than CVC, and so forth.

Specifically, the requirements of markedness constraints follow the form of either negative (3a, b) or positive (3c, d) statements, as illustrated below:

(3) Examples of Markedness Constraints
a. Vowels must not be nasal.

b. Syllables must not have codas.

c. Sonorants must be voiced.

d. Syllables must have onsets.

On the other hand, a ‘faithfulness’ constraint regulates the similarity between the output and input; that is, it requires that the properties of the input be preserved in the output. From a communicative viewpoint, the ‘faithful’ relation between the input and output protects the contrast of lexicon, making a language have different sets of lexical items to express different meanings. A few principal members belong to the faithfulness family. First, MAXIMALITY (MAX) requires that every element in the input have a correspondent in the output, prohibiting any ‘deletion’ of elements. Secondly, DEPENDENCY (DEP) requires that every element in the output must have a correspondent in the input, disallowing any ‘insertion’ of elements. Still another crucial faithfulness constraint is IDENTITY (IDENT), which requires correspondent elements have identical properties or values, with no ‘alternation’ between the input and the output. A tableau showing the violations of faithfulness constraints is represented below:
(4) Violations of Faithfulness Constraints

<table>
<thead>
<tr>
<th>Input</th>
<th>MAX</th>
<th>DEP</th>
<th>IDENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. AB</td>
<td>*(C)!</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. ABCD</td>
<td></td>
<td>*(D)</td>
<td></td>
</tr>
<tr>
<td>c. ABE</td>
<td></td>
<td></td>
<td>*(C→E)</td>
</tr>
</tbody>
</table>

In this tableau-format, candidate (a), owing to the deletion of element C, violates MAX; candidate (b) violates DEP for the epenthesis of element D; candidate (c) incurs a violation of IDENT since the element C in the input is changed to E.

A recent development in OT suggests that an Optimality-theoretic grammar should not be restricted to only two families of constraints, namely markedness and faithfulness constraints, but should admit anti-faithfulness constraints to the constraint inventory. Evidence shown in morpho-phonological processes that serve to strengthen the opposition between two morphological classes gives rise to Alderete’s (1999a, 2001) Transderivational Anti-faithfulness (TAF) theory. The essence of TAF theory is described in (5) below:

(5) Anti-Faithfulness (Alderete 1999a)

For every faithfulness constraint $F$, there is a corresponding anti-faithfulness constraint $\neg F$ that is satisfied in a string S iff S has at least one violation of $F$.

There is an anti-faithfulness constraint for every faithfulness constraint. An
anti-faithfulness constraint triggers alternations by requiring a violation of a related faithfulness constraint. The TAF theory thus successfully accounts for the phonological alternations between two morphological categories that enhance the coding properties of them. The anti-faithfulness hypothesis is largely based on while in the meantime antithetical to Benua’s (1997) Transderivational Faithfulness Theory, in which he aims to formalize the observation that morphologically related words require a certain phonological similarity in the surface forms.

Optimality Theory contends that surface forms in languages are the optimal outputs chosen from all possible output candidates generated by GEN, through the parallel evaluation of ranked constraints. That is, any repairing rule is no longer needed to account for phonological phenomena. Constraints are violable, and the candidate(s) with the least or the most trivial violations wins as the optimal output. Language variety is reflected by different rankings of a set of constraints in every language. Compared to the traditional Generative Phonology, where every language representation is formed through an idiosyncratic step-by-step derivation under a fixed phonological rule order, OT in Generative Grammar leads to a revolutionary shift in phonological theories.

2.2 The Role of Perception in Loanword Adaptation
As loanword adaptations reveal various phonological processes (e.g. vowel epenthesis and consonant deletion), questions as to what mechanism governs the processes and how perception plays a role in it have been the central issues in loanword phonology. In what follows, we review the major proposals regarding these issues in previous studies.

### 2.2.1 Perception and Production as Separate Levels

One approach is to take both perception and production into account but each is governed by a different grammar in a separate level (Silverman 1992, Yip 1993, Kenstowicz 2001). In such a proposal the input to the production level is the output of the perception level.

Silverman (1992) investigates English loanwords in Cantonese and suggests that loanword adaptation involves perception and production, which are governed by two different grammars. A crucial point is that the loanword input does not enter \( L_T \) with the phonological representation of \( L_S \), but instead the \( L_T \) speaker hears a sequence of non-linguistic signals. His multi-scansion model consists of two separate, ordered levels, namely Perceptual Level (Scansion One) and Operative Level (Scansion Two). In the Perceptual Level, constrained by the native segment and prosodic inventory, the borrower perceives the input as a string of non-linguistic signals, which is parsed into
segment-sized chunks. For example, a Cantonese speaker cannot tell apart English voiced and unaspirated voiceless obstruents and perceive both as voiceless, since Cantonese phonological system does not provide access to the contrast. In the Operative Level, various phonological operations are triggered by the LT’s phonotactic constraints and the ‘raw’ material (the output of Scansion One) is adjusted to fit the native syllable structures. For example, in ‘film → [fey lirm]’, since Cantonese has fricatives in the phoneme inventory and the [f] occurs in the onset, it is completely preserved in the surface form. When [f] fills the coda position, however, it may undergo feature change in this level, as in ‘lift → [lip]’, since Cantonese does not allow a fricative to be the coda.

Following Silverman (1992), Yip (1993) agrees that the perception of the LT speaker is governed by the native phonotactics, but unlike Silverman, she adopts a constraint-based framework to account for the phonological processes in the Operative Level. The asymmetry of deletion and preservation of certain consonants in loanword phonology is attributed to their degrees of acoustic salience. For example, /s/ is saliently perceived from surrounding stops or vowels for its high-intensity noise and thus it tends to be preserved in the loaning process. This argument can well account for the three-way distinction that in her data, /s/ is never lost because it is salient, liquids are sometimes lost because they are not very salient, and
pre-consonantal stops are always lost because they are not detected at all.

Kenstowicz (2001) reviews Gbéto’s (1999) study of French, Portuguese, and English loanwords in Fon and singles out a few noteworthy problems which can be better accounted for by means of perception cues. Compared with Yip (1993), in which only the Operative Level is analyzed within OT framework, he sketches two different constraint rankings for both perception and production. This approach well accounts for the adaptation of word-final obstruent-stop clusters in Fon, as exemplified below:

(6) Perception Mapping

\[
\begin{array}{cccc}
/post/ & \text{Dep-V} & *\text{stop/obstruent}_# & \text{Max-C} \\
\text{post} & & *! \\
\rightarrow \text{pos} & & *
\end{array}
\]

Production Mapping

\[
\begin{array}{cccc}
/\text{pos}/ & \text{Max-C} & *\text{stop/obstruent}_# & \text{Dep-V} \\
\rightarrow \text{posu} & & * \\
\text{po} & & *!
\end{array}
\]

(Kenstowicz 2001, p. 11, simplified)
In (6), ranking Dep-V over Max-C enables the word-final stop after an obstruent to delete in the perception mapping, while the reverse ranking in the production mapping prefers vowel epenthesis to consonant deletion.

The main advantage of regarding perception and production as governed by two independent grammars is that it provides a plausible account for the deletion of the unsalient segments in the Perception Level and the preservation of the salient ones in the Operative Level. However, such a treatment is challenged by the loanword adaptations from English to Mandarin.

Firstly, from the acoustic angle, Yip (1993) relates Fay and Cutler’s (1977) evidence on speech errors to the adaptation tendency of the ‘stop-liquid’ onsets in Cantonese loanwords, as in ‘place → [pʰey.si]’, where liquids are more vulnerable to deletion in the context of ‘Cl’ or ‘Cr’. Yip’s observation of the effect of acoustic saliency on loanword phonology thus reinforces Silverman’s Perceptual Level in that a more salient segment is more likely to be detected in the Perceptual Level than a less salient one. In our loanword data, however, consonants are mostly preserved in the onset position (as simplex or complex onset), regardless of the saliency of the consonant(s). This inter-linguistic contrast in loanword adaptation naturally leads to a question: if the input is nothing but a sequence of non-linguistic superficial signal,
why do Cantonese and Mandarin speakers not equally perceive the liquid in the onset as well as the post-consonant stops in the coda, provided that neither of the two governing phonological systems allows the ‘C + liquid’ sequence to be the onset?

Secondly, the two-level grammar as formulated by Kenstowicz (2001) can only account for part of our data, for the consonant cluster may be simplified or completely preserved in Mandarin loanwords, as will be discussed in Chapter 3.

2.2.2 Perception as Integrated into Production

Steriade (2001b) sketches a proposal of P-map, which successfully accounts for several language universals that the earlier OT potentially fails to. The P-map is like a repository of knowledge, comprising hierarchies of absolute and relative human perceptibility of different sound contrasts. The proposed mechanism may thus project the relevant constraints and their mutual rankings in the OT framework. The aim of judging the degrees of ‘confusability’ or ‘similarity’ of a set of comparative pairs is to correctly point out the phonotactic repair strategy that involves the minimal change when a departure from the underlying form is inevitable.

The advantage of such an approach as integrates perception into production in loanword phonology is twofold. Firstly, the mechanism of loanword formation is simple, for all the relevant phonological processes are operated within a single
grammar. Secondly, the two-level approach, posited by Silverman (1992) and others, forces researchers to separate the phonological processes involved into two levels of operation, which is sometimes infeasible. By mapping perceptual similarity to loanword phonology, however, we are able to avoid such a redundancy. Our standpoint is thus theoretically closer to the perceptual-similarity approach in that we formulate a single constraint ranking for loanwords, in which the features that contribute more to dissimilarity, such as [+continuant] and [+nasal], tend to preserve, while a segment in a certain position that contributes less to dissimilarity, such as post-nuclear retroflexes, tends to be ignored.

2.3 Native phonology and Loanword Phonology

Silverman (1992) contends that loanword phonology should operate independently from native phonological system, and the non-native forms are realized in accordance with native constraints. This point of view is later opposed by Yip (1993), where she argues that there is nothing special about loanword phonology. The reason that a phonological process may be found in the loanword domain but not in the native system is the lack of inputs (segments) for which such a process is needed in the native vocabulary. For example, /s/ is not found in the coda position in the native lexicon, and hence the process of vowel epenthesis will never occur.
Yip (2002) provides specific evidence and revises her previous assumption (Yip 1993) in her later work. She suggests that the cross-linguistic preference for excess consonants to be retained should not be attributed to MAX >> DEP in the L_T phonology. Rather, there should be a faithfulness constraint specific to loanword adaptation, termed ‘MIMIC’, that requires the L_T speaker’s ‘mimicking’ of salient segment(s) or suprasegmental qualities in the L_S input. The relative perceptual saliency among segments is judged by two acoustic cues—internal cues, which reside within the duration of the segment itself, and contextual cues, provided by the surrounding context. For example, fricatives are the segment with strong internal cues, while the saliency of unsyllabifiable stops depends heavily on good contextual cues, such as adjacency to a vowel or liquid. Under this view, loanword phonology is the native phonology plus an extra MIMIC family, enforcing faithfulness to the percept. No matter the native phonology is MAX >> DEP or DEP >> MAX, that either MIMIC or MAX ranks high triggers retention, which well accounts for the fact that universally retention is more common than deletion.

Since all the loanwords in Mandarin conform to the native phonotactics, and like native lexicon, bisyllabiciticy is also preferred in English loanwords in Mandarin, our standpoint is closer to Yip (2002) that loanword phonology is native phonology plus other constraints, which are the MIMIC constraint family in her analysis. In our study,
however, something else makes the native and the loanword phonology different.

The high-ranking MIMIC-SALIENT is like the way we outrank MAX-IO([+del rel]) and MAX-IO([+nasal]) over common faithfulness constraints like MAX, DEP, and IDENT, though it is criticized by Yip (2002) that it fails to make any predictions about which segments are cross-linguistically prone to deletion in which environments. However, we respond more to Steriade’s (2002) proposal that the universally relative perceptibility of sound contrasts is ‘mapped’ to the constraint ranking in the OT framework. In other words, we choose to keep the constraint types in the theoretic framework simple.

What makes them different is an anti-faithfulness constraint requiring that a postnuclear retroflex in the input should not have any correspondent in the output, namely ¬MAX-IO(r/V>). Though limited to the [ɔ + r] combination, as in ‘[ɔ35] (ㄈ, ‘son’), ‘[ɔ21] (ㄈ, ‘ear’), ‘[ɔ53] (二, ‘two’), etc.,[_Vr] is nonetheless a possible syllable in Mandarin. Unexpectedly, English [ɔ] or [ɔ̃] do not enter Mandarin as [ɔ], and like other English postnuclear retroflexes, such as that in [_ar], the postnuclear retroflex in both [ɔ] or [ɔ̃] is mostly deleted and surfaces as nothing. This phonological process does not operate in native phonology and is specific to loanword phonology. Owing to this distinction, we contend that Mandarin loanword phonology is native phonology plus the ¬MAX-IO(r/V>) constraint.
2.4 The Corpus Approach

Several works on loanword phonology do not base their analyses on an adequate database (Silverman 1992, Yip 1993, Guo 2001, Shih 2004), and hence the conclusions may be untenable when more counterexamples against their analyses are found. Silverman (1992) and Yip (1993) do not mention the number of their data, while those appearing in their papers are no more than 100 in total. Guo’s (2001) data are limited to the Mandarin transliterations of American state names and typhoon names, which are in total 138, as provided in his appendices. Shih’s (2004) data are primarily drawn from the common English proper names listed in the appendix of an English dictionary as well as his personal collection, which are no more than 200.

On the other hand, in Miao’s (2005) investigation into Mandarin loanword adaptations from three Indo-European languages, namely English, Germany, and Italian, she builds a sizeable corpus, containing 2423 words, with 1177 English loans, which provides more sufficient grounds and makes the arguments persuasive. Comparatively speaking, our corpus of English loanwords in Mandarin (947 words) is a little smaller, since we sift out loans with more than three syllables for the reason that we attempt to investigate the effect of disyllabic tendency in Mandarin, and the transliterations of names that are not commonly used. We believe a sufficient database
benefits the analyses in empirical authenticity.