Article

Language and Linguistics
16(6) 769–789

DOI: 10.1177/1606822X15602600
lin.sagepub.com

OT Factorial Typology and the Typology of Mandarin Mid Vowel Assimilation*

Yen-Hwei Lin

Michigan State University

Optimality Theory (OT) posits a universal set of constraints, and the factorial typology of the constraints, i.e. through all logically possible permutations of constraints to generate all possible constraint rankings, is predicted to produce possible grammars and exclude impossible ones. This paper examines the factorial typology of the markedness constraints relevant to Mandarin mid vowel assimilation, and checks the OT factorial typology against the typology of mid vowel assimilation. I demonstrate that there is a reasonable match between the two typologies, suggest that the few mismatches can be accounted for on principled grounds, and discuss the issue of uneven frequency distribution among different typological patterns.

Key words: factorial typology, Mandarin dialects, Optimality Theory, vowel assimilation

1. Introduction

In Optimality Theory (OT) (Prince & Smolensky 2004), all languages share the same set of universal constraints, and languages differ only in the rankings of the constraints. OT is inherently typological in that if the constraints are formulated correctly, all logically possible permutations of the constraints, i.e. the factorial typology, are expected to produce possible languages and exclude impossible ones. Consider, for example, the three constraints in (1).

(1) Constraints for obstruent voicing
- a. *VOICED-CODA — Coda obstruents are voiceless.
- b. IDENT-IO(VOICE) — The value of the feature [voice] must be identical between the input and output corresponding segments.
- c. *[+VOI, –SON] — No voiced obstruents.

The factorial typology given in (2) shows that the three constraints generate six possible rankings or grammars. The rankings in (2a–b) represent languages that have a contrast between voiced and voiceless obstruents in all positions since the faithfulness constraint is ranked highest. When *[+VOI, –SON] is ranked highest, as in (2c–d), there are no voiced obstruents in any positions.

* An earlier version of this paper was presented at the Thirteenth International Symposium on Chinese Languages and Linguistics (IsCLL-13) in 2012 and some portions of the paper were also discussed in the Third Theoretical Phonology Conference (TPC-3) in 2011. I thank the participants of both conferences and the *Language and Linguistics* anonymous reviewers for their comments and suggestions.

The ranking in (2e) also rules out voiced obstruents because of the ranking of *[+voi, –son] >> Ident-IO(voice). Positional neutralization of obstruent voicing, i.e. contrastive voicing at the onset but only voiceless obstruents in the coda, as found in German and many other languages, is produced by the ranking in (2f). This factorial typology correctly produces three types of attested languages: those with full voicing contrast in all positions, those without any voiced obstruents, and those with voicing neutralization in coda position.

(2) The factorial typology of the three constraints in (1)

a.	Ident-IO(voice)	>>	*Voiced-Coda	>>	*[+voi, –son]
b.	Ident-IO(voice)	>>	*[+voi, –son]	>>	*Voiced-Coda
c.	*[+voi, –son]	>>	*Voiced-Coda	>>	Ident-IO(voice)
d.	*[+voi, –son]	>>	Ident-IO(voice)	>>	*Voiced-Coda
e.	*Voiced-Coda	>>	*[+voi, –son]	>>	Ident-IO(voice)
f.	*Voiced-Coda	>>	Ident-IO(voice)	>>	*[+voi, –son]

In addition, the factorial typology correctly excludes impossible/unattested languages such as those that have contrastive voicing at onset but only voiced obstruents in coda or those that have voiceless obstruents only in onset but both voiced and voiceless obstruents in coda.

Drawing on the data and analysis in Lin (2002), this paper examines in detail the factorial typology of the markedness constraints relevant to Mandarin mid vowel assimilation, checks the OT factorial typology against the typology of mid vowel assimilation, and discusses cases of mismatches and the issue of asymmetrical frequency distribution of the cross-dialectal patterns.

In what follows, §2 introduces the data and analysis of Standard Mandarin mid vowel assimilation and the typology of mid vowel assimilation patterns across Mandarin dialects. The factorial typology of the relevant markedness constraints is presented in §3. I demonstrate that there is a reasonable match between the two typologies, and account for the mismatches on principled grounds. Since some assimilation patterns occur more frequently than others, in §4 I discuss potential approaches to quantitatively modeling the distribution patterns and whether the uneven distribution issue is relevant to phonological theorizing. General discussion and concluding remarks are given in the final section.

2. Mandarin mid vowel assimilation and its typology

I assume five vowel phonemes in Standard Mandarin (SM), as in (3a–c), with the mid and low vowels unspecified for [back] and [round], and three derived glides, as exemplified in (3d). All the Mandarin dialects examined in this paper can be analyzed as having the same vowel system. The maximal syllable for SM as well as the Mandarin dialects examined in this paper is CGVX (C = consonant, G = glide, V = vowel, X = C or G). I assume that the pre-nuclear glide is part of the onset (see Bao 1990, 1996; Duanmu 2000, 2007; Lin 2007), and the post-nuclear glide is within the rime and occupies the coda position.[1]

[1] How glides are structurally affiliated does not affect the general analysis, except that the technical definitions of the relevant constraints would need to be modified. See Lin (2014b) for an overview of various analyses of Chinese syllable structure and references therein.

(3) SM vowels/glides (see Duanmu 2000, 2007; Lin 1989, 2002, 2007, 2014a; Wiese 1997)

a. High vowel phonemes: i y u

b. Mid vowel phoneme: ə

c. Low vowel phoneme: a

d. Glides [j, ɥ, w] are derived from the corresponding high vowels, e.g. /iə/ → [je], /yə/ → [ɥe], /əu/ → [ow].

The examples in (4) show that the mid vowel phoneme undergoes fronting/backing and/or rounding assimilation triggered by a pre-nuclear or post-nuclear glide. The example in (4c) illustrates the lack of bidirectional assimilation even though both pre-nuclear and post-nuclear glides trigger assimilation, as shown in (4a–b). In addition, partial assimilation, i.e. the lack of rounding assimilation, occurs in (4d) when the glide is front rounded.

(4) SM mid vowel assimilation (Tones are omitted for ease of presentation.)

a.	/piə/	[pje]		'don't'
	/kuə/	[kwo]		'wok'
b.	/pəi/	[pej]		'cup'
	/kəu/	[kow]		'dog'
c.	/uəi/	[wej]	*[wøj]	'tail'
	/tiəu/	[tjow]	*[tjøw]	'to throw'
d.	/yə/	[ɥe]	*[ɥø]	'moon'

Partial assimilation in (4d) is usually attributed to the prohibition of *[ø] (Duanmu 2007:52–54, 66; Lin 2002, 2007:156; Ma 2003:146). For the failure of double assimilation in (4c), previous studies either stipulate rule ordering for regressive assimilation to take precedence over progressive assimilation (Cheng 1973; Lin 1989) or proposes a syllable-based account. One syllable-based approach ranks the constraint responsible for rime internal assimilation higher than the constraint for assimilation across the onset-nucleus boundary or within the syllable domain (e.g. Lin 2007:157; Ma 2003:142–144; Wu 1994:72–73). The other approach denies the existence of regressive assimilation and maintains that progressive assimilation only applies in open syllables when the nuclear vowel is phonetically long, e.g. [jeː], [woː] (Duanmu 2000, 2007; Wang 1993). In this paper, I assume a syllable-based account with both regressive and progressive assimilation, in which the domains of regressive and progressive assimilation are syllable rime and an open syllable respectively, but do not require a ranking between the two assimilation constraints (see Lin 2002).

The data in (4) can be analyzed in OT as presented in (5)–(8).[2] The constraints in (5a–b) require assimilation of [back] and [round] from the glide to the vowel within the syllable rime and from

[2] See Lin (2002), Duanmu (2007:52–54), Ma (2003:139–146) for similarities and differences in the analyses and technical details.

the pre-nuclear glide to the vowel in an open syllable,[3] respectively. The universal ranking hierarchies in (5c) claim that [ø] is universally more marked than the other three mid vowels and schwa is the least marked. With the ranking in (5d) for SM, the tableaux in (6)–(8) show that the correct mid vowels are selected accordingly.

(5) Constraints and constraint rankings

a. Rime-Harmony: The vowel is assimilated in [back]/[round] to the glide within the rime.

b. GV-Harmony: The vowel is assimilated in [back]/[round] to the pre-nuclear glide in an open syllable.[4]

c. Segmental markedness constraints and universal ranking hierarchies

(i) *ø >> *e >> *ə (ii) *ø >> *o >> *ə

d. Constraint ranking for SM

*ø >> Rime-Harmony, GV-Harmony >> *o >> *e >> *ə

I assume that the high vowels/glides [i/j], [u/w], and [y/ɥ] are specified as [–back], [+back, +round], and [–back, +round], respectively, and these features are organized as the vowel place features (see Clements & Hume 1995). Based on Feature Class Theory (Padgett 2002), the harmony constraints target the feature class of vowel place but assess the violation of each individual feature. For example, when a schwa is not assimilated to [w], two violation marks are incurred because the vowel is not assimilated in [+back] and [+round]; on the other hand, an unassimilated schwa next to [j] incurs only one violation for the lack of [–back]. This mode of constraint assessment is gradient rather than categorical (Padgett 2002). In (6), the candidate selection is straightforward.

(6) Mid vowel assimilation in syllables with one glide

/əi/	*ø	Rime-Harmony	GV-Harmony	*o	*e	*ə
əj		*!				*
→ ej					*	

/əu/	*ø	Rime-Harmony	GV-Harmony	*o	*e	*ə
əw		*!*				*
→ ow				*		

[3] If one assumes that both elements of a diphthong in the rime are dominated by the nucleus node in Chinese, thereby such a syllable is considered an open syllable, then the domain of the progressive assimilation would be a simplex open syllable to exclude diphthongs.

[4] Since I assume that the pre-nuclear glide is part of the onset, GV-Harmony has to apply within the domain of an open syllable.

/iə/	*ø	Rime-Harmony	GV-Harmony	*o	*e	*ə
jə			*!			*
→ je					*	

/uə/	*ø	Rime-Harmony	GV-Harmony	*o	*e	*ə
wə			*!*			*
→ wo				*		

In (7), the candidate with both progressive and regressive assimilation, i.e. candidate (7c), is ruled out by the markedness constraint against [ø] and/or Rime-Harmony. GV-Harmony is irrelevant since this is not an open syllable, and the two harmony constraints can remain unranked. The crucial role of *ø becomes evident in (8).

(7) Mid vowel assimilation in a syllable with two glides

		/iəu/	*ø	Rime-Harmony	GV-Harmony	*o	*e	*ə
	a.	jəw		*!*				*
	b.	jew		*!*			*	
	c.	jøw	*!	*				
→	d.	jow				*		

(8) Mid vowel assimilation in a syllable with the front rounded glide

		/yə/	*ø	Rime-Harmony	GV-Harmony	*o	*e	*ə
	a.	ɥə			**!			*
	b.	ɥø	*!					
	c.	ɥo			*	*!		
→	d.	ɥe			*		*	

For an open syllable with the front rounded glide, as shown in (8), the completely assimilated candidate (8b) is excluded because of the ban on [ø], and the totally unassimilated candidate (8a) loses to the partially assimilated candidates (8c–d), indicating that when total assimilation is unattainable, partial assimilation is better than a complete lack of assimilation. The selection between the two partially assimilated candidates (8c–d) is determined by the ranking of *o higher than *e.

The typology of Mandarin mid vowel assimilation in Table 1 is taken from Lin's (2002:311–312) survey of 160 Mandarin dialects.[5]

---

[5] For the data sources of the representative dialects, see Lin (2002:343, note 15). For the collection and interpretation of the dialects, see the explanation and references cited in Lin (2002:311–312). Specifically, different descriptions of the same dialect were treated as sub-dialects, and by doing so the proposed analysis has to account for all possible variation patterns reported in the literature.

Table 1: Mid vowel assimilation across Mandarin dialects

Pattern	Representative dialect	Number of dialects	/yə/	/iə/	/uə/	/əi/	/əu/
A	SM	14	ɥe	je	wo	ej	ow
B	Changzhi	45	ɥe	je	wə	ej	əw
C	Jinan	12	ɥe	je	wə	ej	ow
D	Dalian	3	ɥe	je	wə	əj	əw
E	Changge	3	ɥe	je	wo	əj	əw
F	Zhangjiakou	23	ɥe	je	wo	ej	əw
G	Yantai	1	ɥø	je	wo	ej	ow
H	Jiyuan	2	ɥə	jə	wə	əj	əw
I	Jining	5	ɥə	jə	wə	ej	ow
J	Lucheng	6	ɥə	jə	wə	ej	əw
K	Qingdao	7	ɥə	je	wə	ej	ow
L	Baode	4	ɥə	je	wə	ej	əw
M	Yonghe	1	ɥə	je	wo	ej	əw
N	Yanggu	1	ɥə	je	wə	əj	ow
O	Shangqiu	6	ɥo	je	wo	ej	ow
P	Qintong	1	ɥo	je	wo	əj	əw
Q	Linqi	2	ɥo	je	wo	ej	əw
R	Tongchuan	13	ɥe/ɥo	je	wo	ej	ow
S	Huozhou	11	ɥe/ɥo	je	wo	ej	əw

As we can see, the degree to which mid vowel assimilation applies varies among Mandarin dialects. At the two extremes, there are cases where mid vowel assimilation is fully enforced (pattern G) or totally lacking (pattern H).[6] The predominant surface form for /yə/ is [ɥe], but all three other possible outputs are attested. There is only one case with the fully assimilated [ɥø] (pattern G). The totally unassimilated [ɥə] (patterns H–N) occurs more frequently than the partially assimilated output [ɥo] (patterns O–Q). Moreover, the two partially assimilated outputs coexist in some dialects (patterns R–S). For the four sequences containing [j] or [w], when the two extremes (patterns G and H) are excluded, the variation ranges from full assimilation (patterns A, O, R), total lack of regressive (patterns D, E, P) or progressive assimilation (patterns I, J), to complete lack of rounding (patterns B, D, J, L), the lack of either progressive or regressive rounding but not both (patterns C, E, F, I, K, M, N, P, Q, S), and the lack of either progressive or regressive fronting but not both (patterns D, E, I, J, N, P).

[6] The glide–vowel–glide forms are not included since they follow the same assimilation process within the rime applicable to /əi/ and /əu/.

Based on the same set of constraints in (5) for the analysis of SM mid vowel assimilation, the constraint ranking for each pattern is shown in Table 2. For simplicity, the lowest ranked constraint that is not involved in cross-dialectal re-ranking, i.e. *ə, is excluded from Table 2 and subsequent discussion.

Table 2: Constraint ranking for each pattern (see Lin 2002)

Pattern	Dialect	/yə/	Constraint ranking
A	SM	ɥe	*ø >> Rime-Harmony, GV-Harmony >> *o >> *e
B	Changzhi	ɥe	*ø >> *o >> Rime-Harmony, GV-Harmony >> *e
C	Jinan	ɥe	*ø >> Rime-Harmony >> *o >> GV-Harmony >> *e
D	Dalian	ɥe	*ø >> *o >> GV-Harmony >> *e >> Rime-Harmony
E	Changge	ɥe	*ø >> GV-Harmony >> *o >> *e >> Rime-Harmony
F	Zhangjiakou	ɥe	*ø >> GV-Harmony >> *o >> Rime-Harmony >> *e
G	Yantai	ɥø	Rime-Harmony, GV-Harmony >> *ø >> *o, *e
H	Jiyuan	ɥə	*ø >> *o, *e >> Rime-Harmony, GV-Harmony
I	Jining	ɥə	*ø >> Rime-Harmony >> *o, *e >> GV-Harmony
J	Lucheng	ɥə	*ø >> *o >> Rime-Harmony >> *e >> GV-Harmony
K	Qingdao	ɥə	*ø >> Rime-Harmony >> *o >> GV-Harmony >> *e
L	Baode	ɥə	*ø >> *o >> Rime-Harmony, GV-Harmony >> *e
M	Yonghe	ɥə	*ø >> GV-Harmony >> *o >> Rime-Harmony >> *e
N	Yanggu	ɥə	*ø >> Rime-Harmony >> *o >> GV-Harmony >> *e
O	Shangqiu	ɥo	*ø >> Rime-Harmony, GV-Harmony >> *e >> *o
P	Qintong	ɥo	*ø >> GV-Harmony >> *e >> *o >> Rime-Harmony
Q	Linqi	ɥo	*ø >> Rime-Harmony, GV-Harmony >> *e >> *o
R	Tongchuan	ɥe/ɥo	*ø >> Rime-Harmony, GV-Harmony >> *o, *e
S	Huozhou	ɥe/ɥo	*ø >> Rime-Harmony, GV-Harmony >> *o, *e

The tableaux in (9) illustrate how different constraint rankings produce five different patterns derived from /yə/. An example deriving [ɥe] under a different ranking from SM is shown in (9a). In (9b), where harmony constraints are ranked highest, full assimilation occurs, but when harmony constraints are ranked lowest, the unassimilated output is selected, as in (9c). In (9d), the alternative partially assimilated output [ɥo] is possible when *e is ranked higher than *o, which is opposite to the *o >> *e ranking in SM. Pattern R in (9e) has the same ranking as in SM except that *o and *e are unranked, resulting in variation between the two partially assimilated outputs.

(9) Sample tableaux for various surface forms for /yə/

a. Jinan (pattern C)

	/yə/	*ø	RIME-HARMONY	*o	GV-HARMONY	*e
	ɥə				**!	
	ɥø	*!				
	ɥo			*!	*	
→	ɥe				*	*

b. Yantai (pattern G)

	/yə/	RIME-HARMONY	GV-HARMONY	*ø	*o	*e
	ɥə		*!*			
→	ɥø			*		
	ɥo		*!		*	
	ɥe		*!			*

c. Jiyuan (pattern H)

	/yə/	*ø	*o	*e	RIME-HARMONY	GV-HARMONY
→	ɥə					**
	ɥø	*!				
	ɥo		*!			*
	ɥe			*!		*

d. Shangqiu (pattern O)

	/yə/	*ø	RIME-HARMONY	GV-HARMONY	*e	*o
	ɥə			**!		
	ɥø	*!				
→	ɥo			*		*
	ɥe			*	*!	

e. Tongchuan (pattern R)

	/yə/	*ø	RIME-HARMONY	GV-HARMONY	*o	*e
	ɥə			**!		
	ɥø	*!				
→	ɥo			*	*	
→	ɥe			*		*

The tableaux in (10) illustrate how pattern B, the one that occurs most frequently in the database, can be derived. The difference between this pattern and that of SM is the lack of both progressive and regressive rounding assimilation. For detailed illustration of all patterns, see Lin (2002).

(10) Changzhi (pattern B)

	/əi/	*ø	*o	Rime-Harmony	GV-Harmony	*e
	əj			*!		
→	ej					*

	/əu/	*ø	*o	Rime-Harmony	GV-Harmony	*e
→	əw			**		
	ow		*!			

	/iə/	*ø	*o	Rime-Harmony	GV-Harmony	*e
	jə				*!	
→	je					*

	/uə/	*ø	*o	Rime-Harmony	GV-Harmony	*e
→	wə				**	
	wo		*!			

	/yə/	*ø	*o	Rime-Harmony	GV-Harmony	*e
	ɥə				**!	
	ɥø	*!				
	ɥo		*!		*	
→	ɥe				*	*

Note that in Table 2, some distinct patterns in Table 1 are shown to have the same constraint ranking: patterns C, K and N, patterns B and L, patterns F and M, patterns O and Q, and patterns R and S. As will be discussed later, these cases involve complications in which some patterns assimilate only one feature, e.g. only [–back] or only [+round], or some patterns assess the harmony constraints categorically.

3. The factorial typology and matching the two typologies

In Table 3, I list the logically possible rankings of the proposed five constraints and indicate whether a particular ranking is attested in the Mandarin database. For simplicity, I assume that relevant faithfulness constraints are ranked at a position that does not interfere with the re-ranking of these constraints, so this factorial typology does not include faithfulness constraints. Note that because of the fixed universal rankings of *ø >> *o and *ø >> *e, the factorial typology does not produce any rankings in which *o or *e is ranked above *ø. Moreover, some rankings produce the same pattern; e.g. the two rankings Rime-Harmony, GV-Harmony >> *ø >> *o >> *e and Rime-Harmony, GV-Harmony >> *ø >> *e >> *o have the same effects, so they are combined into one ranking: Rime-Harmony, GV-Harmony >> *ø >> *o, *e.

Table 3: Factorial typology (see Lin 2002: Appendix)

Ranking 1	RIME-HARMONY, GV-HARMONY >> *ø >> *o, *e
Pattern	full assimilation (pattern G)
Dialect (number)	Yantai (1)
Ranking 2	*ø >> *o, *e >> RIME-HARMONY, GV-HARMONY
Pattern	no assimilation (pattern H)
Dialect (number)	Jiyuan (2)
Ranking 3	*ø >> RIME-HARMONY, GV-HARMONY >> *o, *e
Pattern	full assimilation except /yə/ as [ɥe] or [ɥo] (pattern R)
Dialect (number)	Tongchuan (13)
Ranking 4	*ø >> RIME-HARMONY, GV-HARMONY >> *o >> *e
Pattern	full assimilation except /yə/ as [ɥe] (pattern A)
Dialect (number)	Standard Mandarin (14)
Ranking 5	*ø >> RIME-HARMONY, GV-HARMONY >> *e >> *o
Pattern	full assimilation except /yə/ as [ɥo] (pattern O)
Dialect (number)	Shangqiu (6)
Ranking 6	*ø >> RIME-HARMONY >> *o, *e >> GV-HARMONY
Pattern	no progressive assimilation, unassimilated [ɥə] (pattern I)
Dialect (number)	Jining (5)
Ranking 7	*ø >> GV-HARMONY >> *o,*e >> RIME-HARMONY
Pattern	no regressive assimilation; /yə/ as [ɥe] or [ɥo]
Dialect (number)	**no example**; possible dialects similar to ranking 8
Ranking 8	*ø >> GV-HARMONY >> *o >> *e >> RIME-HARMONY
Pattern	no regressive assimilation; /yə/ as [ɥe] (pattern E)
Dialect (number)	Changge (3)
Ranking 9	*ø >> GV-HARMONY >> *e >> *o >> RIME-HARMONY
Pattern	no regressive assimilation; /yə/ as [ɥo] (pattern P)
Dialect (number)	Qintong (1)
Ranking 10	*ø >> RIME-HARMONY >> *o >> GV-HARMONY >> *e
Pattern	no progressive rounding; /yə/ as [ɥe] (pattern C) /yə/ as [ɥə] (pattern K)
Dialect (number)	Jinan (12) Qingdao (7)
Ranking 11	*ø >> RIME-HARMONY >> *e >> GV-HARMONY >> *o

Pattern	no progressive fronting; /yə/ as [ɥo]
Dialect (number)	**no example**
Ranking 12	*ø >> GV-Harmony >> *o >> Rime-Harmony >> *e
Pattern	no regressive rounding; /yə/ as [ɥe] (pattern F) /yə/ as [ɥə] (pattern M)
Dialect (number)	Zhangjiakou (23) Yonghe (1)
Ranking 13	*ø >> GV-Harmony >> *e >> Rime-Harmony >> *o
Pattern	no regressive fronting; /yə/ as [ɥo]
Dialect (number)	**no example**
Ranking 14	*ø >> *o >> Rime-Harmony, GV-Harmony >> *e
Pattern	no rounding; /yə/ as [ɥe] (pattern B) /yə/ as [ɥə] (pattern L)
Dialect (number)	Changzhi (45) Baode (4)
Ranking 15	*ø >> *e >> Rime-Harmony, GV-Harmony >> *o
Pattern	no fronting; /yə/ as [ɥo]
Dialect (number)	**no example**
Ranking 16	*ø >> *o >> Rime-Harmony >> *e >> GV-Harmony
Pattern	no rounding; no progressive fronting; unassimilated [ɥə] (pattern J)
Dialect (number)	Lucheng (6)
Ranking 17	*ø >> *o >> GV-Harmony >> *e >> Rime-Harmony
Pattern	no rounding; no regressive fronting; /yə/ as [ɥe] (pattern D)
Dialect (number)	Dalian (3)
Ranking 18	*ø >> *e >> Rime-Harmony >> *o >> GV-Harmony
Pattern	no fronting; no progressive rounding; unassimilated [ɥə]
Dialect (number)	**no example**; possible dialects similar to ranking 16.
Ranking 19	*ø >> *e >> GV-Harmony >> *o >> Rime-Harmony
Pattern	no fronting; no regressive rounding; /yə/ as [ɥo]
Dialect (number)	**no example**
Ranking 20	GV-Harmony >> *ø >> *o, *e >> Rime-Harmony
Pattern	no regressive assimilation; fully assimilated [ɥø]
Dialect (number)	**no example**

Ranking 21	GV-Harmony >> *ø >> *o >> Rime-Harmony >> *e
Pattern	no regressive rounding; fully assimilated [ɥø]
Dialect (number)	**no example**
Ranking 22	GV-Harmony >> *ø >> *e >> Rime-Harmony >> *o
Pattern	no regressive fronting; fully assimilated [ɥø]
Dialect (number)	**no example**
Ranking 23	Rime-Harmony >> *ø >> *e >> GV-Harmony >> *o
Pattern	no progressive fronting; partially assimilated [ɥo]
Dialect (number)	**no example**
Ranking 24	Rime-Harmony >> *ø >> GV-Harmony >> *o, *e
Pattern	full assimilation except /yə/ as [ɥe] or [ɥo]
Dialect (number)	possible dialects in pattern R (see ranking 3)
Ranking 25	Rime-Harmony >> *ø >> GV-Harmony >> *o >> *e
Pattern	full assimilation except /yə/ as [ɥe]
Dialect (number)	possible dialects in pattern A (see ranking 4)
Ranking 26	Rime-Harmony >> *ø >> GV-Harmony >> *e >> *o
Pattern	full assimilation except /yə/ as [ɥo]
Dialect (number)	possible dialects in pattern O (see ranking 5)
Ranking 27	Rime-Harmony >> *ø >> *o >> *e >> GV-Harmony
Pattern	no progressive assimilation; /yə/ as [ɥə]
Dialect (number)	possible dialects in pattern I (see ranking 6)
Ranking 28	Rime-Harmony >> *ø >> *o >> GV-Harmony >> *e
Pattern	no progressive rounding; /yə/ as [ɥe] or [ɥə]
Dialect (number)	possible dialects in patterns C/K (see ranking 10)
Ranking 29	GV-Harmony >> *ø >> Rime-Harmony >> *o, *e
Pattern	full assimilation; [ɥø]
Dialect (number)	possible dialects in pattern G (see ranking 1)

Given the fixed universal rankings of *ø >> *o and *ø >> *e, those patterns predicted not to exist are indeed not found in the dataset. For example, there is no pattern in which [o] is lacking (i.e. [wə]/[əw] but *[wo]/*[ow]) and yet [ɥo] or [ɥø] is allowed to surface, and there is no pattern in which [e] is lacking (i.e. [jə]/[əj] but *[je]/*[ej]) and yet [ɥe] or [ɥø] is allowed to surface.

Note that rankings 10, 12 and 14 are applicable to more than one pattern: patterns C and K, patterns F and M, and patterns B and L, respectively. Lin's (2002) analysis suggests that the second patterns, i.e. K, M or L, with only one different output for /yə/ → [ɥə], result from a categorical assessment of GV-Harmony, as the example for pattern L shows in (11). Baode (pattern L) has exactly the same assimilation patterns as Changzi (pattern B) except that /yə/ surfaces as [ɥə] rather than [ɥe]. With categorical assessment, candidate (11a) incurs only one violation for GV-Harmony even though two features are unassimilated (see the tableaux for Changzhi pattern B given in (10)).

(11) Baode (pattern L): Categorical GV-Harmony (see Lin 2002:322)

		/yə/	*ø	*o	Rime-Harmony	GV-Harmony	*e
→	a.	ɥə				*	
	b.	ɥø	*!				
	c.	ɥo		*!		*	
	d.	ɥe				*	*!

The intuition behind this categorical assessment is that some dialects prefer not to entertain partial assimilation when full assimilation is not possible. An alternative phonetics-based explanation could be that for some dialects, an output form like [ɥə] can better maintain the perceptibility of a front rounded vowel/glide since such a vocoid is acoustically/perceptually more ambiguous and less robust than [i] and [u]. In such an account, some phonetically-based constraint(s) would be responsible for the second patterns in rankings 10, 12 and 14, and I leave the nature of such constraints for future research. Despite the complication and regardless of which analysis is adopted, these additional patterns are subsumed under the rankings predicted by the factorial typology.

The OT factorial typology in Table 3 also predicts additional patterns unattested in the dataset. Some accidental gaps are expected since the dataset is confined to a group of related dialects and does not constitute a comprehensive list of Mandarin dialects and other languages. For example, rankings 24–29 are those that Lin (2002:339) has identified to be either unattested or possible patterns similar to some of the existing ones. If we assume that [ɥ] cannot end a syllable in Mandarin to produce a form like [øɥ], due to some higher ranked constraints, then rankings 24–29 would produce the same patterns as some of the existing patterns, as indicated in Table 3. Ranking 18 is likely to be an accidental gap since it is similar to the attested ranking 16 (pattern J) but involving a different assimilatory feature, and ranking 7 is similar to the attested ranking 8 (pattern E) but exhibiting variation between [ɥe] and [ɥo].

In addition, I suggest that some of the predicted but unattested patterns can be accounted for on phonetic, functional and/or markedness grounds. First, given the rarity of [ɥø] (only one dialect in the dataset), it is not surprising that some of the nonexistent patterns are those that produce [ɥø] (rankings 20–22). One possible explanation for disfavoring [ɥø] is the lack of perceptual distance between [ɥ] and [ø], similar to the lack of [ji] or [wu] in some languages. Second, note that [ɥo] occurs less frequently than [ɥe] and [ɥə] and many of the unattested patterns involve *e >> *o (rankings 11, 13, 15, 19, 23). Only patterns O and P (seven dialects in total), i.e. rankings 5 and 9 respectively, require *e >> *o, and the majority of the dialects either have *o >> *e or have the two constraints unranked. These generalizations seem to indicate that *e >> *o tends to be disfavored. If [o] involves two relevant features [+back] and [+round] whereas [e] involves only one feature [–back], as I have assumed, then [o] is likely to be considered by some dialects to be more marked than [e], which may then account for the tendency. Along this line, Lin (2002:334–335) suggests that *e >> *o is a universally marked ranking and languages do not adopt such a ranking unless the actual data conflict with the ranking hierarchy. Since [ɥ] is front and [o] is back and since rounding on vowels normally functions as an enhancement feature on [+back], an alternative phonetics-based explanation is that the two segments in [ɥo] are perceptually more confusable or conflicting than [ɥe] and [ɥə], and hence are more marked.

Now that I have discussed all cases of predicted but unattested patterns, let us consider the attested patterns that the factorial typology in Table 3 does not seem to predict, i.e. patterns S, Q and N. According to Lin's (2002:§3.4) analysis, each of these patterns is still subsumed under the rankings predicted by the factorial typology, as shown in Table 4; the main difference is that assimilation within the rime does not target all the vowel place features, but an individual feature.

Table 4: Patterns involving a single assimilatory feature (Lin 2002:340)

Ranking 3	*ø >> Rime-Harmony, GV-Harmony >> *o, *e
Feature involved	Rime-Harmony targets [–back]
Pattern	no regressive rounding; partially assimilated [ɥe] or [ɥo]
Dialect (number)	Huozhou (11) (pattern S)
Ranking 5	*ø >> Rime-Harmony, GV-Harmony >> *e >> *o
Feature involved	Rime-Harmony targets [–back]
Pattern	no regressive rounding; partially assimilated [ɥo]
Dialect (number)	Linqi (2) (pattern Q)
Ranking 10	*ø >> Rime-Harmony >> *o >> GV-Harmony >> *e
Feature involved	Rime-Harmony targets [+round] GV-Harmony is categorically assessed
Pattern	no regressive fronting, no progressive rounding; unassimilated [ɥə]
Dialect (number)	Yanggu (1) (pattern N)

Huozhou (pattern S) has ranking 3, which is the same as pattern R, but differs in that Rime-Harmony targets only [–back]; Linqi (pattern Q) has ranking 5, which is the same as pattern O, but differs in that Rime-Harmony targets only [–back]; and Yanggu (pattern N) which is the same as pattern K, but differs in that Rime-Harmony targets only [+round]. Since assimilation can target either a feature class or just one feature, this variation is not unexpected. The tableaux in (12) illustrate how pattern S is derived (see Lin 2002:329).

(12) Huozhou (pattern S)

Rime-Harmony: The vowel is assimilated in [–back] to the glide within the rime.

	/əi/	*ø	Rime-Harmony	GV-Harmony	*o	*e
	əj		*!			
→	ej					*

	/əu/	*ø	Rime-Harmony	GV-Harmony	*o	*e
→	əw					
	ow				*!	

	/iə/	*ø	Rime-Harmony	GV-Harmony	*o	*e
	jə			*!		
→	je					*

	/uə/	*ø	Rime-Harmony	GV-Harmony	*o	*e
	wə			*!*		
→	wo				*	

	/yə/	*ø	Rime-Harmony	GV-Harmony	*o	*e
	ɥə			**!		
	ɥø	*!				
→	ɥo			*	*	
→	ɥe			*		*

In sum, there is a reasonable match between the factorial typology and the typology of Mandarin mid vowel assimilation. Those patterns predicted by the fixed universal rankings and the factorial typology to be nonexistent are indeed unattested. The cases of mismatches can be explained by appeal to (i) universal phonetic, functional and/or markedness factors, (ii) the choice of the assimilation target and/or of the assessment mode, or (iii) the prediction that there are dialects and languages to be documented that can fill in some of the gaps. For the current dataset, every pattern can fit into one of the rankings generated by the factorial typology. After taking into consideration single feature assimilation, categorical assessment, preferred rather than absolute universal markedness ranking, and/or phonetically-grounded factors, the predications made by OT factorial typology could become more complicated. Some discussion is given in the next two sections, but further investigation is beyond the scope of this paper and hence left for future research.

4. The issue of uneven frequency distribution

One interesting issue arises regarding the uneven frequency distribution among possible patterns predicted by the factorial typology. As we have seen in our case study, some predictable patterns are either missing or less common, and among the attested patterns, some occur more frequently than others. Although some of the accidental gaps or less common patterns can be attributed to the limitation of the dataset confined to a group of closely related dialects, there still seems to be a distinction between more favorable versus less favorable patterns. The question then is how to account for the asymmetrical frequency distribution. Lin (2002:335–336) suggests a markedness approach: (i) categorical assessment is more marked when the harmony constraints target a feature class, so patterns produced by such a marked assessment are less common (e.g. patterns K, M and L in rankings 10, 12, 14 respectively), (ii) assimilation that targets a single feature rather than a feature class is a marked option, so patterns produced by single-feature assimilation are less common (patterns N, S and Q as in Table 4), and (iii) *e >> *o is a universally marked ranking, so patterns produced by such a ranking are either missing (rankings 11, 13, 15, 19, 23) or occur less frequently (patterns O and P in rankings 5 and 9 respectively). In addition, as discussed in the previous section, one can also appeal to a phonetics-based account for the rarity of [ɥø] and the low frequency of [ɥo]. These approaches are intuitively appealing, but it remains to be seen how the frequency differences can be modeled quantitatively.

A factorial typology predicts possible grammars but does not indicate tendency or frequency with regard to which rankings/grammars are more common than others. However, a closer look points to potential quantitative modeling.

One possibility is to calculate the number of rankings that produce the same pattern. Sometimes the same pattern can be produced by more than one ranking, and this simple calculation may suggest that such a pattern occurs more frequently cross-linguistically than another that can be produced by only one ranking. However, the Mandarin data do not seem to support this approach since among the attested patterns, the least common one, i.e. Yantai (pattern G), can be produced by six rankings, but the most common pattern, i.e. Changzhi (pattern B), can be produced by only two rankings, as shown in (13).

(13) Number of rankings that produce patterns G and B

a. Pattern G (1 dialect): rankings 1 and 29 in Table 3
   Pattern G can be produced by the following rankings
   RIME-HARMONY >> GV-HARMONY >> *ø >> *o >> *e
   RIME-HARMONY >> GV-HARMONY >> *ø >> *e >> *o
   GV-HARMONY >> RIME-HARMONY >> *ø >> *o >> *e
   GV-HARMONY >> RIME-HARMONY >> *ø >> *e >> *o
   GV-HARMONY >> *ø >> RIME-HARMONY >> *o >> *e
   GV-HARMONY >> *ø >> RIME-HARMONY >> *e >> *o

b. Pattern B (45 dialects): ranking 14 in Table 3
   Pattern B can be produced by the following rankings
   *ø >> *o >> RIME-HARMONY >> GV-HARMONY >> *e
   *ø >> *o >> GV-HARMONY >> RIME-HARMONY >> *e

More generally, the voicing contrast typology given in (2), where positional neutralization is produced by one ranking but full contrast in all positions is produced by two rankings, does not seem to support the approach either. It seems unlikely that cross-linguistically languages with full contrast in all positions are twice as many as those exhibiting positional neutralization, since it is generally believed that voicing neutralization in coda is less marked.

The second possibility is to explore typological entailments: the presence of one input–output mapping entails the presence of another input–output mapping, but not vice versa. For example, a pattern with [ɥø] entails the presence of [je] and [wo], but a pattern with [wo] does not entail the presence of [ɥø] or [ɥe]. Since such typological entailments generated by factorial typology have been shown to have consequences for quantitative patterns in sociolinguistic and grammar variation (Anttila et al. 2008), this may be a promising approach to modeling asymmetrical frequency distribution among different rankings cross-linguistically.

On the other hand, one could hold the view that the phonology component of a formal grammar has nothing to say about the different degrees of frequency or popularity of different grammars, which can be attributed to performance and functional factors (see de Lacy & Kingston 2013:344–346). Under this view, OT's factorial typology only needs to predict possible and impossible grammars.

5. Discussion and concluding remarks

In this paper, I have conducted an investigation into how well an OT factorial typology matches the typology of Mandarin mid vowel assimilation. Given the fixed rankings of *ø >> *o and *ø >> *e, the prediction that impossible patterns would not be generated is supported by the current Mandarin dataset. The prediction that the OT factorial typology produces only possible or attested patterns is less straightforward. For those mismatches between the factorial typology and the typology of Mandarin mid vowel assimilation, it has been suggested that in some cases, universal factors on phonetic, functional and/or markedness grounds are involved,[7] and in some others, it is a matter of not having a comprehensive empirical database for a full comparison. Importantly, however, every attested Mandarin pattern can be accounted for by some ranking generated by the factorial typology.

In the OT literature, it is not uncommon to find that some patterns/languages predicted by an OT factorial typology are unattested cross-linguistically, at least based on our current knowledge. For example, Pater (1999:319–320) notes that to avoid a sequence of a nasal followed by a voiceless consonant, among the repair strategies that languages have adopted—including postnasal voicing, nasal-stop fusion, gemination, nasal deletion—one strategy predicted by the relevant factorial typology is unattested, i.e. epenthesis between the nasal and the voiceless consonant. It remains unclear why this is the case, although he speculates that perhaps for an epenthetic vowel to break up a placed-assimilated nasal-consonant sequence violates geminate integrity. Steriade (2009) also points out the 'too many solutions' problem in OT; i.e. OT predicts more possible outputs and repairs than those that are actually attested in human languages. To solve the over-generalization problem, she proposes a perception-based account in which all else being equal, the optimal input–output mapping should be perceptually most similar. Some unattested patterns in Mandarin can probably be analyzed along these lines. For example, if we can show experimentally that the mapping between /yə/ and [ɥe] is perceptually more similar to that between /yə/ and [ɥø], then we can better explain the rarity of [ɥø] and the preference for [ɥe].

In addition, Kager (1999:35) argues against 'a deeply naïve assumption' of the factorial typology in which all possible rankings must be instantiated by some attested languages. He considers this assumption to be naïve, as one that expects 'all logically possible permutations of genetic material in the human genome are actually attested in individual humans.' He suggests that for the

[7] For a phonetics-based approach to low vowel assimilation/raising in Chinese dialects (e.g. /ian/ → [jen] in Mandarin), see Hsieh (2012).

typology to be meaningful, one should focus on the patterns/languages produced by permutations of broader general types of constraints. If this is the goal of OT factorial typology, then the results of this Mandarin study are in line with such an approach since all attested patterns are matched to some rankings generated by the factorial typology when the harmony constraints are considered general types of constraints, i.e. regardless of which feature is the target for assimilation and whether or not the constraints are assessed by gradient or categorically.

Broadly speaking, under the view that formal theories of phonology only need to define a set of possible grammars, but do not guarantee that every possible grammar exists, the lack or rarity of a theoretically definable possible language can be attributed to performance and functional reasons (de Lacy & Kingston 2013:344–346). The results of this Mandarin study are also compatible with this view since all attested patterns are definable possible grammars by OT's factorial typology, and the unattested patterns can be accounted for on functional grounds.

Other than the uneven frequency distribution among possible patterns discussed in §4, there are a few issues for future studies. First, the empirical bases of vowel assimilation need to be expanded and the constraints need to be refined so as to better test the matching between vowel assimilation typology and the relevant factorial typology.

Second, the impressionistic descriptions of these Mandarin dialects or other Chinese dialects or any languages, on which phonological analyses are based, ideally should be confirmed by instrumental phonetic studies. In addition, to verify the accuracy of the typology of Mandarin mid vowel assimilation, it is necessary to examine more closely the contrastive nature of and the variation in the data. For example, some dialects may exhibit variation between [ow] and [əw], and the question is whether or not the assimilation pattern is phonological or phonetic.[8] In this study, I assume that all patterns are phonological or, more precisely, allophonic in the postlexical phonology, and following Lin (2002), I include all possible variations reported in the literature (see fn.5). However, since the reported variations could be different accents within the same dialect or the mid vowel is highly variable in some dialects and the transcriptions are given based on the transcribers' preferences, sociolinguistic and instrumental phonetic studies can help address the relevant issues.

Third, theoretically, there is the issue of how much predictive power should be granted to factorial typology, and in what format, as the central mechanism to generate possible languages and exclude impossible ones. Moreover, there is the question as to what language universal and/or functional factors are and to what extent they can or should be integrated to make better typological predictions.

In conclusion, as the first step to checking OT factorial typology against the typology of Mandarin mid vowel assimilation, this study shows that it is possible to maintain factorial typology as a theoretical tool for predicting possible and impossible patterns if appropriate assumptions are made and language universal and functional factors are taken into consideration. Although such an approach may, under certain views, weaken the predictive power and the theoretical status of OT factorial typology, an expansion of empirical data and a refinement of how constraints are formulated and what constraints are adopted for the purpose of constraint permutations in a factorial typology may still lead to more accurate predictions for typological patterns and help us better understand language universals.

[8] Thanks to an anonymous reviewer for raising the issue.

References

Anttila, Arto, Vivienne Fong, Štefan Beňuš, & Jennifer Nycz. 2008. Variation and opacity in Singapore English consonant clusters. *Phonology* 25.2:181–216.

Bao, Zhiming. 1990. *Fanqie* languages and reduplication. *Linguistic Inquiry* 21.3:317–350.

Bao, Zhiming. 1996. The syllable in Chinese. *Journal of Chinese Linguistics* 24.2:312–354.

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

Clements, George Nick, & Elizabeth Hume. 1995. The internal organization of speech sounds. *The Handbook of Phonological Theory* (1st edition), ed. by John A. Goldsmith, 245–306. Oxford: Blackwell.

de Lacy, Paul, & John Kingston. 2013. Synchronic explanation. *Natural Language & Linguistic Theory* 31.2:287–355.

Duanmu, San. 2000. *The Phonology of Standard Chinese* (1st edition). Oxford & New York: Oxford University Press.

Duanmu, San. 2007. *The Phonology of Standard Chinese* (2nd edition). Oxford & New York: Oxford University Press.

Hsieh, Feng-fan. 2012. Low vowel raising in Sinitic languages: assimilation, reduction, or both? *Language and Linguistics* 13.4:583–623.

Kager, René. 1999. *Optimality Theory*. Cambridge & New York: Cambridge University Press.

Lin, Yen-Hwei. 1989. *Autosegmental Treatment of Segmental Processes in Chinese Phonology*. Austin: University of Texas at Austin dissertation.

Lin, Yen-Hwei. 2002. Mid vowel assimilation across Mandarin dialects. *Journal of East Asian Linguistics* 11.4:303–347.

Lin, Yen-Hwei. 2007. *The Sounds of Chinese*. Cambridge & New York: Cambridge University Press.

Lin, Yen-Hwei. 2014a. Segmental phonology. *The Handbook of Chinese Linguistics*, ed. by C.-T. James Huang, Yen-Hui Audrey Li & Andrew Simpson, 400–421. Chichester: Wiley Blackwell.

Lin, Yen-Hwei. 2014b. Yinjie yu Hanyu yinduan yinxi [The syllable and Chinese segmental phonology]. *Dangdai Yuyanxue* [*Contemporary Linguistics*] 2014.3:328–345.

Ma, Qiuwu. 2003. *Optimality Theory and Mandarin Syllable Structure*. Tianjin: Nankai University Press.

Padgett, Jaye. 2002. Feature classes in phonology. *Language* 78.1:81–110.

Pater, Joe. 1999. Austronesian nasal substitution and other NC̥ effects. *The Prosody–Morphology Interface*, ed. by René Kager, Harry van der Hulst & Wim Zonneveld, 310–343. Cambridge & New York: Cambridge University Press.

Prince, Alan S., & Paul Smolensky. 2004. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden: Blackwell.

Steriade, Donca. 2009. The phonology of perceptibility effects: the P-map and its consequences for constraint organization. *The Nature of the Word: Studies in Honor of Paul Kiparsky*, ed. by Kristin Hanson & Sharon Inkelas, 151–179. Cambridge: MIT Press.

Wang, Jenny Zhijie. 1993. *The Geometry of Segmental Features in Beijing Mandarin*. Newark: University of Delaware dissertation.

Wiese, Richard. 1997. Underspecification and the description of Chinese vowels. *Studies in Chinese Phonology*, ed. by Jialing Wang & Norval Smith, 219–249. Berlin & New York: Mouton de Gruyter.
Wu, Yuwen. 1994. *Mandarin Segmental Phonology*. Toronto: University of Toronto dissertation.

[Received 20 October 2014; revised 18 March 2015; accepted 18 March 2015]

Department of Linguistics and Languages
619 Red Cedar Road
B-331 Wells Hall
Michigan State University
East Lansing, MI 48824
USA
liny@msu.edu

優選論階乘類型與官話方言中元音同化

林燕慧

密西根州立大學

優選理論 (OT) 採用具有語言普遍性的制約條件表達語言的共性。制約的階乘類型，也就是制約的所有可能的等級排序，則可預測可能存在的語言並排除不可能的語言。本文以漢語官話方言中元音同化現象的標誌制約所發展出的階乘類型來查驗優選論階乘類型預測的準確性。研究結果顯示階乘類型的預測大致得到印證。對於少數預測存在但是缺乏語料支持的例子，作者提出原則性的解釋，並討論不同語言類型分布不均的問題。

關鍵詞：官話方言，元音同化，優選理論，階乘類型