Asymmetries in Chengdu and Northern Mandarin Phonotactics:
Variation and Learning

Sijia Zhang & Anne-Michelle Tessier

University of British Columbia

1   Introduction

Decades of research within constraint-based phonological grammars has formally modeled many aspects of first language (L1) learning — often by comparing simulated data to observed stages of child speech (e.g., Levelt et al., 1999; Curtin and Zuraw, 2002; Hayes, 2004; Jarosz, 2010; Becker and Tessier, 2011; Jesney, 2016). However, comparable work in the second language (L2) domain of phonological acquisition remains relatively limited (though c.f. Broselow, 2004; Zhang & Tessier, 2024)[1]. This paper reports interim results from one such larger project, which investigates in particular how L1 grammars that permit multiple surface optima (i.e., grammatical variation) might serve as the basis for transfer to an L2 grammar built of the same constraints, and what consequences that inherent L1 variation might bring about in L2 learning.

This study examines phonotactic variability in sequences of low vowel + nasal coda (+ following consonant) (loVN.C) in two varieties of Mandarin, Northern Mandarin (NM) and Chengdu Mandarin (CM), as compared with North American English. We first report the phonotactic variation in the two Mandarin varieties, and then model L1 NM and CM grammars and their comparison with L2 English. We conducted a production study to collect nasometry data on the relative rates and asymmetries of phonological repairs to low vowel + nasal codas. We adopt a Maxent framework (Goldwater and Johnson, 2003 *et seq*) of weighted constraints, and a Full Transfer approach to L2 acquisition (as in van Leussen and Escudero, 2015; see also Schwartz and Sprouse, 1996). We establish the necessary weighting conditions for Maxent grammars that produce multiple possible optima across contexts, with different probability distributions for each language. We discuss the learning effects of adopting a MaxEnt framework to capture that variability, and its consequences and challenges when applying the grammar to novel inputs.

This paper is structured as follows. Section 2 compares the phonotactic restrictions of loVN and loVN.C sequences in the two Mandarin varieties vs. English. Section 3 presents the methodology for the production study, including data coding and nasality measurement. Section 4 reports the results for the production study, focusing on the variation in the proportions of nasal deletion and nasal place change/assimilation produced by L1 NM and L1 CM speakers, as predicted by logistic mixed-effects regression models. Section 5 presents a MaxEnt account of L2 grammatical learning of these sequences in different contexts, with the L1 end-state grammar as the initial L2 state. The proposed grammars successfully model the output variabilities and asymmetries, albeit with some exceptions. Section 6 discusses some remaining issues and implications.

2   The phonotactics of loVN (+C) in Mandarin vs. English

Both English and Northern Mandarin (including Mandarin spoken in Beijing and adjacent regions) have three nasal consonants /m, n, ŋ/ and two low vowels which contrast for [+/–back] in their surface inventories. Standard North American English uses [æ, ɑ] while Mandarin uses [a, ɑ] (Duanmu, 2007). However,

[1] Much of the L2 literature has focused on more phonetic and perceptual aspects of learning, e.g. Escudero & Boersma (2004); van Leussen & Escudero (2015); Zhou & Hamann (2024).

Northern Mandarin (at least) has two crucial phonological restrictions on lowV+ nasalC that do not apply in English: (1) only [n] and [ŋ] are allowed in coda position, but not [m] (Duanmu, 2007); (2) low vowels must agree for [+/–back] with a following coronal or dorsal nasal coda (Rhyme Harmony in Duanmu, 2007; Luo et al., 2020). Thus, English has 6 possible loVN combinations while Mandarin has only two, as shown in Table 1.

In addition, the surface realizations of the legal Mandarin loVN sequences /an/ and /ɑŋ/ may in fact be quite variable. In Northern Mandarin, these outputs can surface either with a nasal coda consonant, or with coda deletion along with nasalization and possible compensatory lengthening of the preceding vowel (see Chen, 2000; Duanmu, 2007 inter alia; Luo et al., 2020). For example, Chen (2000) conducted acoustic measurements on the production of Beijing Mandarin coda nasals /n, ŋ/ preceded by high, mid, and low vowels (e.g., shan-ao [ʂan1.aʊ4] *'hollow of the hill'*). It was found that more than half of the word tokens containing the loVN rhyme were produced without an oral closure for the nasal coda at normal speech rates, but that this oral closure deletion was less frequently observed when the nasal followed a high or mid vowel (Chen, 2000, p. 53-54).

Chengdu Mandarin is the Mandarin variety spoken in Chengdu, Sichuan province in China. In Chengdu Mandarin, while the nasal coda in loVN also experiences substantial amount of deletion, the [+back] sequence more often preserves its coda, whereas the [-back] loVN sequence is especially prone to deletion (Liao et al., 2022).

Mandarin		English	
/loVN/			
[an] ~ [ã] 山 shan1 *'mountain'*	*[aŋ]	[æn] *'ban'*	[æŋ] *'fang'*
[ɑŋ] ~ [ɑ̃] 忘 wang4 *'forget'*	*[ɑn]	[ɑn] *'dawn'*	[ɑŋ] *'wrong'*
/loVN/ + /C/		Monomorphemic	Compound
[an.k] ~ [ã.k] ~ [aŋ.k] 翻开 fan1.kai1 *'open up'*		[æŋ.k] *'anchor'*	[æŋ.k], ?[æn.k], *[æ̃.k] *'pancake'*
[ɑŋ.t] ~ [ɑ̃.t] ~ [ɑn.t] 商讨 ʂɑŋ1.tao3 *'discuss'*		[ɑn.t] *'context'*	[ɑn.t], ?[ɑŋ.t], *[ɑ̃.t] *'long term'*

Table 1. Comparison of Mandarin and English loVN.(C) with examples.

We extend the pattern to the loVN.CV context, which in Mandarin occurs only at a morpheme boundary between a word-final nasal coda and a following word-initial C (e.g. /fan1/ + /kai1/ 'open up'). English, though, requires nasal place assimilation to the following stop C onset within a morpheme boundary (e.g., [æŋkɚ] 'anchor') (Kiparsky, 1985). This NC assimilation process is much more variable across morphemic boundaries (e.g., gree[n b]ox ~ gree[m b]ox), depending on factors like speech rate (Coetzee & Pater, 2011). In Mandarin, NC assimilation is arguably more variable and applies to a lesser degree than English, partially due to the fact that VN.C only occurs across words. Luo (2020), for example, reported partial/incomplete assimilation for /ŋ/.

Our particular focus in this paper are nasal-coda + stop-onset inputs with different places of articulation (mismatched NC) – i.e. /an+kV/ and /ɑŋ+tV/, and how these sequences are produced by the two varieties of Mandarin speakers. The L2 learners must resolve the competing pressures between applying nasal place assimilation vs. conforming to the L1 phonotactics (Rhyme Harmony) when producing English mismatched NC sequences. Previous findings (e.g., Chen, 2000; Luo, 2020) suggest three attested output candidates: (a) coda place assimilation, which would violate Rhyme Harmony (RH) for loVN; (b) faithful sequence, as a repair for loVN [+/–back] agreement; (c) deletion, as another repair strategy to avoid the violation of RH. We will compare the variability in the proportions of the three output candidates across L2 learners from the two Mandarin L1 varieties, by examining different types of sequences (e.g., Mandarin-legal vs. illegal VN).

3   Methods: a production study

3.1   *Participants*   Eleven L1 Northern Mandarin (NM) speakers and eight L1 Chengdu Mandarin (CM) speakers participated in the study. All L1 Northern Mandarin speakers self-reported that they spoke a

Northern variety of Mandarin (e.g., Beijing Mandarin) as their first language, and did not speak any other Chinese language. All L1 Chengdu Mandarin speakers learned to speak Chengdu Hua before the age of 5, in addition to Mandarin.[2] All L1 Mandarin speakers learned a North American variety of English as a second language, with the mean onset age of acquisition of 6.8 years old. They arrived in Canada at the mean age of 20.1 (range = 16-26), and had been living in Canada for 2.5 years on average. They were all living in Vancouver at the time of the study. Ten L1 North American English speakers took part in the study as the control group. All participants were compensated with $20 CAD cash for their participation.

3.2   *Procedure*   L1 Mandarin-speaking participants completed both Mandarin and English word production parts, and L1 English speakers only completed the English part. In each trial, participants saw the written text of the target word on the screen. They were instructed to click the "next" button on the screen to listen to the recorded audio of the target word. Three carrier sentences were presented after the end of the audio, and they were asked to read and complete the sentences using the target word they saw and heard. For the Mandarin part, participants listened to the audio stimuli recorded in the corresponding Mandarin varieties that they spoke. All instructions and prompts in the Mandarin part were presented in simplified Chinese characters, while the English part was presented in English only. There were a few practice trials before they started a different block or section during the production experiment.[3] All trials were fully randomized within a block, including 16 filler trials in the English part (loVC or loVC.C sequences). There were 96 test trials in total for the Mandarin part (96 words * 3 repetitions = 288 tokens), and 77 test trials in total for the English part (77 words * 3 repetitions = 231 tokens).

The study was implemented in JsPsych. The session was conducted at the Child Phonology Lab at the University of British Columbia. Participants sat in front of a laptop which presented the jsPsych experiment in a recording room. The nasometry software (PENTAX Medical, Model 6450) was installed and run on a Windows 7 Dell PC next to the laptop, and the nasometer headset was connected to the PC. After calibration, the experimenter placed the nasometer headset on the participant at the beginning of the study. The separator plate rested against the philtrum between the nose and the mouth to isolate nasal and oral airflow. The two microphones positioned on the two sides of the plates measured speech signals from the nasal and oral cavities. All audio stimuli were played through a speaker connected to the laptop.

3.3   *Materials*   Mandarin stimuli included 16 monosyllabic (monomorphemic) words with loVN sequences, 40 disyllabic words with loVN.C sequences, and 40 nonce disyllabic words. All real words were selected from the corpus DOWLS (Neergaard, et al., 2021). The loVN monosyllabic words consisted of 8 words with /an/ (e.g., /ʂan1/ 'mountain') and 8 words with /ɑŋ/ (/wɑŋ4/ 'forget'). For the loVN.C words, the following Cs included stops (/t, tʰ, k, kʰ/), fricatives (/s, ɕ, x/), affricates (/ tsʰ, ts, tɕ, tɕʰ/) and glides (/j, ɥ, w/). The place of the N could either match or mismatch the place of the following C. For example, /ʂɑŋ1.xai4/ 'harm' is an example of matched place, whereas /fan1.kʰai1/ 'open up' is an example of mismatched place (see Table 2 for details).

Word type	Monosyllabic	Real disyllabic	
Context	VN	VN.C: matched NC	VN.C: mismatched NC
/an/	/ʂan1/ 山 *'mountain'*	/tan1.tu2/ 单独 *'alone'*	/fan1.kʰai1/ 翻开 *'open up'*
/ɑŋ/	/wɑŋ4/ 忘 *'forget'*	/tʰɑŋ2.kʊo3/ 糖果 *'candy'*	/ʂɑŋ1.tao3/ 商讨 *'discuss'*

Table 2. Mandarin stimuli type and examples.

77 English target words were included in the production experiment, including 16 monomorphemic words with VN sequences (e.g., b[æn]) and 21 with VNC sequences (e.g., c[ɑn]tact), 16 real compounds with VNC (e.g., p[æn] [k]ake), and 24 nonce compounds with VNC (e.g., dawn yield). All real words were selected from the corpora IPhOD (Vaden et al., 2009). A Canadian English speaker (female, aged 29) recorded all English stimuli. For the loVN.C words, the following Cs included stops (/t, d, k, g/), fricatives (/s, z, ʃ/), affricates (/tʃ, dʒ/) and glides (/j, w/). As presented in Table 3, the monomorphemic VNC words

---

[2] Chengdu Hua is a local Chinese language spoken in Chengdu, Sichuan province in China.
[3] There was a nonce word block in the production study but we only report data on the real words in this paper.

had matched place between N and C (e.g., t[æŋg]le), wheras the compounds always had mismatched place (e.g., l[ɑŋ] [t]erm). The VN context included Mandarin-legal and -illegal sequences.

Type	monomorphemic		compound
Context	VN	VN.C: matched NC	VN.C: mismatched NC
legal: /æn/, /ɑŋ/	b**an,** wr**ong**	**hand**y, d**onk**ey	p**an c**ake, l**ong t**erm
illegal: /ɑn/ /æŋ/	pr**awn,** b**ang**	c**ont**act, **anch**or	l**awn c**are, h**ang d**og

Table 3. English stimuli type and examples.

3.4 *Data coding and nasality measurement*   The criteria for determining whether the nasal coda was deleted or not were based on combining evidence of nasalance measurements and subjective coding by three phonetically trained coders.

The degree of nasality of the nasal stop and the preceding vowel was measured through nasalance scores. Nasalance score was calculated as the relative proportion of energy (amplitude) coming from the nose vs. the mouth: $A_{nasal}/(A_{oral} + A_{nasal})$. A Praat script was run to measure the average root-mean-squared (RMS) amplitude from both the oral and nasal channels for the vowel and nasal intervals respectively.

The subjective coding was based on acoustic evidence from the waveform and the spectrogram in Praat. In general, nasal consonant production was indicated by increased amplitude in the nasal channel and decreased amplitude in the oral channel. When a nasal coda was coded as present, the oral channel showed zero or near-zero amplitude while the nasal channel showed clear energy, indicating a separation of the two channels as vowel energy declined and disappeared. This pattern was further supported by weakened vowel formants on the spectrogram. In contrast, when a nasal coda was coded as absent, changes in the oral and nasal channels were more gradual, with oral energy persisting alongside nasal energy throughout the nasal interval.

A nasal coda was considered deleted if it was coded as "absent" based on acoustic cues or if its nasalance score was below 80 (Shi et al., 2011). Conversely, a nasal coda was considered fully produced if it was coded as "present" and had a nasalance score above 80. These two criteria were largely consistent.

Regarding the coding of the place of the nasal coda (alveolar vs. velar), each token was independently coded by two coders based on perceptual judgements and acoustic cues observed at the VN boundary on the spectrogram such as velar pinch. Inter-coder agreement was high; cases of disagreement were resolved by a third coder.

4   Results

Figure 1 and Figure 2 plot the overall raw data, namely proportions of nasal coda outputs that were coded as faithful, deleted, or changed in place (i.e., alveolar to velar or vice versa). Fig.1 presents the two groups' Mandarin productions, and Fig 2. presents their L2 English productions; each split by following

segmental context (note that here, we only analyze the N.C clusters where the C was a stop).
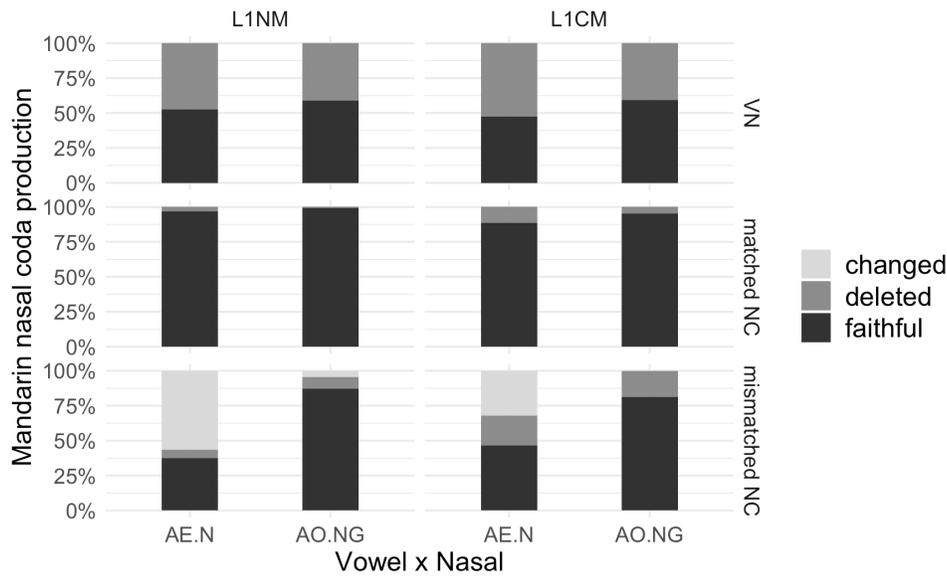
Figure 1. Mandarin nasal coda produced in simple words with VN and disyllabic compounds with VN + matched stop C and mismatched stop C by L1 NM and L1 CM speakers. %Outputs were calculated for nasal deletion, faith and nasal place change (assimilation).
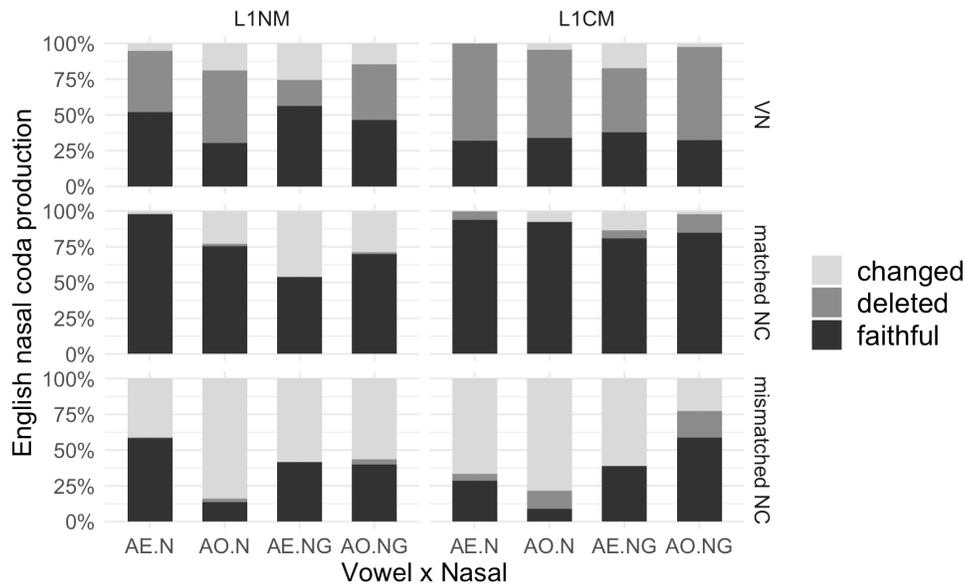
Figure 2. English nasal coda produced by L1 NM and L1 CM speakers in simple words with VN, simple words with VN + matched stop C, compounds with VN + mismatched stop C. %Outputs were calculated for nasal deletion, faith and place change.

We ran four separate logistic mixed-effects models: two predicting the proportion of deletion compared to other treatments, and another two to predict proportion of place change (repair or NC place assimilation), one each for Mandarin and English productions.[4] For further analysis and a larger set of

---

[4] We do not report results from the L1 English control group, but they produced faithful nasal coda outputs with almost no deletion, and quite little nasal place changes, as compared to the L2 speakers.

segmental and morphological contexts, see Zhang (forthcoming).

4.1   Results of L1 Mandarin production

4.1.1   Nasal coda deletion   A logistic ME model was fit for Mandarin L1 simple word and disyllabic compound production, with independent variables of group (L1 NM vs. L1 CM), following context (null vs. matched NC vs. mismatched NC). The outcome variable was nasal coda deletion (deleted = 1, not deleted = 0). The model revealed a significant main effect of the following context (β = 4.78, p < 0.001), such that more deletion occurred in loVN in isolation compared to when a following C was added, regardless of whether the place was matched or not between the NC. There was a significant interaction effect between group and following context, such that CM speakers produced more deletion than NM speakers when there was a following C (β = -1.03, p < 0.05). Among loVN.C sequences with matched vs. mismatched NC, CM speakers deleted more nasal codas for the mismatched NC than the matched NC (β = -2.12, p < 0.01), yet NM speakers did not have such a difference.

4.1.2   Nasal place assimilation   In L1 Mandarin production, the nasal coda underwent place change only in the context of a mismatched following C, as seen from Fig. 1. We fit a model on place assimilation for mismatched NC sequences in Mandarin production. Place assimilation applied almost exclusively to the alveolar nasal rather than the velar nasal in both NM and CM (β = -3.72, p < 0.001). In addition, NM speakers produced more assimilation than CM speakers (β = -1.22, p < 0.05).

4.2   Results of L2 English production

4.2.1   Nasal coda deletion   A logistic ME model was fit for both English L2 simple word and compound production, with independent variables of group (L1 NM vs. L1 CM), following context (null vs. matched NC vs. mismatched NC), and VN legalness (whether VN follows Rhyme Harmony as in Mandarin: legal vs. illegal), and their interactions. The outcome variable was nasal coda deletion (deleted = 1, not deleted = 0). Similarly, the model suggested that there was significantly more deletion in loVN in isolation compared to when a following C was added, with either a matched or mismatched place (β = 4.51, p < 0.001). There were no effects of L1 Mandarin varieties or VN legalness.

4.2.2   Nasal place change in simple words   Moving on to L2 English production, we first report a model on how much place repair applied in simple words including loVN and loVN.C. Note that in the simple loVN.C words, the place of NC was always matched in the UR, so the change of the place must be due to the pressure of Rhyme Harmony between V and N. The model showed that NM speakers had significantly more place repair compared to CM speakers (β = -0.81, p < 0.05). Illegal loVN sequences underwent more place repair than legal sequences (β = -0.80, p < 0.05). More place repair applied to loVN.C compared to loVN in isolation (β = 0.73, p < 0.05). A significant interaction effect suggested that NM speakers applied more place repair for illegal sequences than CM speakers (β = -1.62, p < 0.05).

4.2.3   Nasal place assimilation in disyllabic words   We also fit a model on how much place assimilation applied in disyllabic words including both simple and compound words with loVN.C. Place assimilation was coded when the produced nasal coda shared the same place as the following C. A change of nasal place in the compound was considered "assimilation", whereas a faithful output in the simple loVN.C words was considered "assimilation". For the mismatched NC sequences (compounds), more assimilation applied when VN was illegal compared to legal VN for both Mandarin speakers (β = -1.57, p < 0.001). For instance, */hang/ + /dog/* is more likely to assimilate to *ha[n.d]og* compared to */fan/ + /girl/ -> fa[ŋ.g]irl*. Mismatched NC had slightly more assimilation than matched NC (β = -0.93, p < 0.05), and CM speakers had more assimilation (faithful to the UR) than NM speakers for the matched NC (β = 1.43, p < 0.001). Interestingly, there was a significant interaction effect between matched vs. mismatched NC and VN legalness (β = 2.91, p < 0.001). The effect of Rhyme Harmony on the proportion of place assimilation was greater for the matched NC than the mismatched NC.

The following table (Table 4) summarizes some of the key findings on the comparison between L1 vs. L2 production by the two varieties of speakers.

Context	Mandarin	English
VN	deletion (~0.45): NM = CM	deletion (~0.49): NM = CM
	faithful (~0.55): NM = CM	place change: NM (0.16) > CM (0.05)
VN.C: matched NC	mostly faith: NM (0.98) > CM (0.92)	faith (~0.80), with little deletion
	little deletion: CM (0.08) > NM (0.02)	place change: NM (0.27) > CM (0.06)
VN.C: mismatched NC	deletion: CM (0.20) > NM (0.07)	assimilation (~0.61): NM = CM
	assimilation: NM (0.31) > CM (0.16)	

Table 4. Summary of speaker group-level differences in the outputs of Mandarin and English.

In sum, both groups of speakers had a substantial and equal amount of deletion in loVN and seemed to transfer this pattern into L2 English, though L1 NM speakers repaired place to satisfy VN Rhyme Harmony more often than L1 CM speakers in English. For the VN.C sequences with place-matched NC, both groups of speakers produced the faithful forms overwhelmingly and very few deletion forms. For the VN.C sequences with place-matched NC, both groups produced overwhelmingly faithful forms and very few deletions. In L1 Mandarin, NM speakers produced more faithful and fewer deletion forms than CM speakers. In L2 English, NM speakers showed more nasal place repair, paralleling their production in the VN sequences. For the mismatched NC, however, the L2 patterns cannot be fully explained by L1 transfer. In the L1, CM had more deletion while NM had more assimilation. But in the L2 production, this asymmetry did not show up as it did in the L1: both groups of speakers had around the same rates of assimilation and deletion.

5 MaxEnt Grammar and learning

In this section, we present a formal account of the grammars of NM and CM speakers in a set of MaxEnt grammars. We show how the MaxEnt-predicted proportions may (or may not) match the observed production data – that is, the proportions reported in the previous Results section – based on the proposed constraints and weighting conditions. If a pattern is significantly different among the two groups of Mandarin speakers as evidenced by the statistical models in section 4, we propose two different grammars respectively (although due to the limit of space we only present tableaux for NM and discuss the group difference for CM in the prose); otherwise we present a single grammar for both groups. To simplify the grammatical illustrations here, we only present our grammar's treatment of alveolar nasal inputs (i.e., /an/ and /ɑn/), though further analysis will need to distinguish between alveolar vs. velar inputs to capture the differentt treatment of Mandarin mismatched VN.C (see also Zhang and Tessier, 2024).

We start with the L1 grammars for the two Mandarin varieties (5.1 below) as a starting point for L2 grammatical learning of English (5.2). In what follows, we provide hand-weighted grammars, i.e. with constraint weights chosen by us (out of the many possible options), to illustrate the interconnectedness of these constraints across contexts and L1 vs. L2 languages. A full analysis which uses a MaxEnt learner (Goldwater and Johnson, 2003; Hayes and Wilson, 2008; *interalia*) to reach the closest approximation of the two groups' observed probabilities, given certain priors, is work in progress (see Zhang, forthcoming.)

First, we introduce the following Markedness constraints:

RHYME HARMONY: Assign a violation mark to every sequence of two segments, a low vowel + nasal consonant, where one is [+back] and the other [–back] [adapted from Duanmu (2007)].

*NASALV: Assign a violation mark for every nasalized vowel.

*FINALCODA: Assign a violation mark to every segment associated with the coda position of a syllable word finally.

LICENSE-CODAPLACE: Coda place must be licensed by the following consonant (adapted from Itô, 1988).

*C.C: Assign a violation mark for every coda consonant followed by an onset consonant across a syllable boundary.[5]

We will also use the following Faithfulness constraints (as well as one other novel constraint, introduced later):

IDENT-PLACE: Assign a violation mark for every pair of input and output segments in correspondence which disagree in place.

UNIFORMITY: Assign a violation for every output segment that corresponds to more than one input segment (i.e., whenever multiple input segments are fused into a single output segment).

5.1 *Mandarin grammars for loVN and VN.C sequences*   We start with the MaxEnt grammar for the loVN sequences in simple monosyllabic words for NM and CM, which are identical in this case:

(1) Mandarin monosyllabic VN:

/an/	*FINALCODA 9	UNIFORMITY 5	*NASALV 4	*H*	*prob*	*Observed*
(i). an	-1			-9	0.5	0.5
(ii). ã		-1	-1	-9	0.5	0.5

In the Mandarin monosyllabic VN, the faithful output form in (1i) violates *FINALCODA. The output form in (1ii) undergoes coda deletion and shift the nasalization feature onto the vowel, which violates both UNIFORMITY[6] and *NASALV. The constraint weights shown in (1) condition satisfy the weighting condition $w(\text{*FINALCODA}) = w(\text{UNIFORMITY} + \text{*NASALV})$, so the grammar outputs an equal proportion of deletion and faithful forms. Recall that Mandarin morphemes all satisfy Rhyme Harmony (RH), meaning that no /VN/ inputs will include */aŋ/ or */ɑn/ sequences. Thus we also note that RH must be weighted highly enough to rule out the faithful mapping of those rich base inputs. This weighting will be crucial below, starting in (3).

Next, we present the relevant constraints and their weights for Mandarin disyllabic compounds with a matched NC place in (2):

(2) Mandarin disyllabic compound VNC with matched NC:

/an + t/	*C.C 6.5	UNIFORMITY 5	*NASALV 4	*H*	*Prob*	*Observed*
(i). an.t	-1			-6.5	0.924	0.93
(ii). ã.t		-1	-1	-9	0.076	0.07

In (2), Mandarin speakers produce overwhemingly more faithful outputs and very little deletion. Thus, we have $w(\text{UNIFORMITY} + \text{*NASALV}) \gg w(\text{*C.C})$. Compared to *FINALCODA in (1), the key markedness constraint here *C.C that bans a coda consonant word-internally has a lower weight.

Now we consider the grammars for VN+C with mismatched NC place, with the NM system illustrated in (3). Since the weights of *C.C, UNIFORMITY and *NASALV have all been determined already as in (2), the choice between the output candidates (3i) and (3ii) is determined by the competition between RH and IDENT-PLACE vs. LICENSE-CODAPLACE. For NM, %place assimilation (0.57) has a higher probability than %faith (0.37), so the sum of RH and IDENT-PLACE which both penalize candidate (3ii) must outweigh LICENSE-CODAPLACE. However: inspection of this tableau in (3) reveals that the deletion candidate (iii) is now more complicated to consider:

---

[5] We intend this constraint to only apply to C.C sequences at word boundaries, i.e. at the juncture between members of a compound. Any other *C.C constraints which apply to word-internal contexts must be weighted much more highly in a Mandarin grammar.

[6] We note that UNIFORMITY may well need to be a family of constraints, penalizing fusion relative to specific segmental or featural components; this particular flavour of UNIFORMITY may only be relevant to penalize the fusion of an input vowel+nasal sequence into a single nasalized vowel. See more discussion in section 5.3.

(3) NM: disyllabic compound VNC with mismatched NC

/an+k/	*C.C 6.5	LIC-CODAP 6	EXISFAITH [NASAL-C] 5.5	UNIF 5	RH 5	*NASALV 4	IDENT-PLACE 0.5	*H*	*prob*	*observed*
(i). an.k	-1	-1						*-12.5*	*0.36*	*0.37*
(ii). aŋ.k	-1				-1		-1	*-12*	*0.60*	*0.57*
(iii). ã.k			-1	-1		-1		*-14.5*	*0.04*	*0.06*

Looking now at candidate (3iii), there must be a constraint in addition to UNIFORMITY and *NASALV to further penalize this particular deletion mapping, but *not* the deletion mapping in (2ii) – in other words, it must somehow be worse to delete an input nasal from an /N+C/ cluster if the resulting surface [C] does not share that /N/'s place. For the moment we will call this constraint 'Existential Faith', relativized to Nasal Consonants (EF-NC), whose definition we will discuss in section 5.3. The gang of these three constraints – UNIFORMITY, *NasalV and EF-NC – must have a combined weight greater than w(*C.C + RH + IDENT-PLACE) and w(*C.C + LIC-CODAP), so the grammar's outputs are mostly faith and assimilation forms with a very low proportion of deletion.

To compare with CM (whose grammar we do not present in a tableau for reasons of space): this group had a lower probability for place assimilation form (0.32) compared to NM (0.57). The opposite weighting condition applies to CM as %faith (0.47) > %assimilation, i.e., w(RH + IDENT-PLACE) >> w(LIC-CODAP). CM has a higher weight of IDENT-PLACE (1) and lower weight of LIC-CODAP (5.65). Deletion has a much higher probability for CM (0.21) compared to NM (0.06), thus EF-NC has a lower weight in CM (4), with *NASALV and RH at the same weights in both groups.

5.2 *L2 English grammars for loVN and VN.C sequences*   As mentioned above: we assume that the initial learning state of the L2 is identical to the end state grammar of the L1, as presented in (1)-(3). As the L2 learner receives more and more English inputs – including those that are novel to their L1 (here, illegal VN sequences like [ɑn]) – the L2 grammar gradually develops towards the target grammar with the readjustment of the constraint weights. The production data here are from L2 learners at an advanced learning stage, which has nevertheless not reached the target grammar yet.

For the English monosyllabic words with VN that are legal in Mandarin in (4), the output proportions are slightly different from the L1 Mandarin VN (see (1)), with a bit more deletion (0.6) compared to the faithful output (0.4) in L2 English. This is not a pattern that can be fully explained by learning or transfer, as if anything we would expect *more* faithful productions, not fewer. Nevertheless, here we assume that the weights of the constraints are still mostly transferred from (1), though with a small decrease of UNIFORMITY and *NASALV, such that w(*FINALCODA) >> w(UNIFORMITY + *NASALV).

(4) English monosyllabic legal VN

/æn/	*FINALCODA 9	UNIF 4.75	*NASALV 3.75	H	*prob*	*observed*
(i). æn	-1			-9	*0.38*	0.4
(ii). æ̃		-1	-1	-8.5	*0.62*	0.6

(5) shows the grammar for NM regarding the monosyllabic Mandarin-illegal VN. The weight of RH has decreased from 5 in the L1 grammar to 0.2, so that the faithful form [ɑn] in (5i) which violates RH is often allowed (0.3). However, the largest probability mass goes to deletion outputs (0.51). IDENT-PLACE outweighs RH in NM so that more faithful (i.e. place-mismatched) VN in (5i) are produced compared to the place-repaired form in (5ii).

(5) NM: English monosyllabic illegal VN

/ɑn/	*FINALCODA 9	UNIF 4.75	*NASALV 3.75	IDENT-PLACE 0.5	RHYME HARM 0.2	*H*	*prob*	*observed*
(i). ɑn	-1				-1	*-9.2*	*0.27*	*0.30*
(ii). ɑŋ	-1			-1		*-9.5*	*0.20*	*0.19*
(iii). ɑ̃		-1	-1			*-8.5*	*0.53*	*0.51*

Again comparing (5) for NM to the CM group: the CM grammar has a higher weight of IDENT-PLACE (1.5) and the same weight of RH, which leads to a smaller proportion of place change (0.05) compared to NM (0.19) in the output. Note that the group difference in the place change here mirrors the L1 Mandarin production in VNC with mismatched NC. Namely, NM speakers change nasal place more often compared to CM speakers, with a lower weight of IDENT-PLACE in both L1 and L2 grammars.

In our current grammatical analysis, we will skip the grammar for English simple disyllabic words, with matched NC (e.g., *anchor*). We recognize the result in 4.2.2 and Fig. 2 ("matched NC") suggest a small proportion of place changes i.e. /ɑŋ.k/ → [ɑn.k] in NM. Since this input has matched place between both VN and NC, and this winner has neither, it would appear that a harmonically-bounded candidate is winning much more frequently than any MaxEnt grammar would allow. We therefore hypothesize that learners must sometimes be using the wrong stored input for these forms, i.e. /ɑn.k/, and so set them aside.

Now we move on to the disyllabic compounds with legal VN and mismatched NC (e.g., *pan cake*) as shown in (6). According to section 4's results, the NM and CM group difference disappears (e.g., in place change), with both groups producing outputs of mostly place assimilation (0.53) and a decrease in deletion (0.03). UNIFORMITY, *NASALV, IDENT-PLACE and RH all have their weights already determined for this L2 grammar in the monosyllabic words illustrated in (4) and (5). Compared to the L1 grammar, the weights of LIC-CODAP and ENCF have dropped dramatically, from 6 and 5.5. to 0.9 and 1 respectively.

(6) NM: English disyllabic compound with legal VN and mismatched NC

/æn + k/	*C.C 6.5	UNIF 4.75	*NASV 3.75	EF-NC 1	LIC-CODAP 0.9	IDENT-PLACE 0.5	RH 0.2	*H*	*prob*	*observed*
(i). æn.k	-1				-1			*-7.4*	*0.43*	*0.44*
(ii). æŋ.k	-1					-1	-1	*-7.2*	*0.52*	*0.53*
(iii). æ̃.k		-1	-1	-1				*-9.5*	*0.05*	*0.03*

If we translate these same constraint weights from (6) to the context of illegal VN + mismatched NC (e.g., *lawn care*) as in (7), the difference between the faithful and assimilation output probabilities increases compared to the legal VN context as expected. The grammar prefers the assimilation form much more over faith, as the former option satisfies both RH and LIC-CODAP. However, this asymmetry predicted by the Maxent grammar (faith: 0.34; assimilation: 0.61) is not to the same extent as the observed values (faith: 0.12; assimilation: 0.81). The same problem occurs in a CM grammar, i.e., we also cannot not predict the exact same great amount of asymmetry if we translate the grammar for legal VN to the illegal context in compounds with the current set of constraints. We leave for future work the task of investigating what further constraints might best amplify the difference between these two contexts.

(7) NM: English disyllabic compound with illegal VN and mismatched NC

/ɑn + k/	*C.C 6.5	UNIF 4.75	*NASV 3.75	EF-NC 1	LIC-CODAP 0.9	IDENT-PLACE 0.5	RH 0.2	*H*	*prob*	*observed*
(i). ɑn.k	-1				-1		-1	*-7.6*	***0.34***	***0.12***
(ii). ɑŋ.k	-1					-1		*-7*	***0.61***	***0.81***
(iii). ɑ̃.k		-1	-1	-1				*-9.5*	***0.05***	***0.07***

5.3 *What is 'Existential Faith' and its relevance to nasal consonants?*   In our MaxEnt analysis of VN+C sequences, we find that the observed rates of nasal deletion across contexts require an additional penalty for deletion in N+C place-mismatched contexts (compare candidates (2ii) vs. (3iii)). Our tentative suggestion as to what constraint imposes this penalty is a brand of 'existential' faithfulness, with regard to the set of features that an input nasal is specified for (for a sort of precedent, see Struijke, 2000). More specifically, we provide a definition for such a constraint:

> EXISFAITH-[NASALCONSONANT]: Assign a violation for every input nasal consonant which has *none* of its input features in correspondence with an output **consonant.**

This constraint is satisfied by any candidate which retains a surface nasal consonant, whether faithful or place-changed, since both of those output nasals will retain the input segment's [nasal] feature. In the case of what we have described as nasal segment "deletion" – which clearly involves at least some degree of fusion, with the [nasal] feature surfacing on the preceding vowel, see footnote 6 – this existential constraint can still be satisfied if the input nasal's *place* feature can be mapped to a surface consonant. When the input's /N+C/ sequence shares place, as in (2), this can be achieved by more fusion, fusing the input nasal's place feature with its following stop. Thus, the surface [t] in candidate (2ii) is somehow the bearer of the input /n/'s coronal place. However: if the input's /N+C/ sequence does not share place, then no such fusion is possible: a surface sequence of [æ̃.k] as in (3iii) or even [ɑ̃.k] in (7iii) has not retained any of the input nasal's place OR nasal features on an output *consonant* – and so they violate this Existential-Faith.

We do not present this constraint as a perfect specimen of cross-linguistic typology within a universal CON; rather, we use it as a tool to investigate how our production data may be captured using a MaxEnt grammar and the rest of our constraint set. Ongoing work will continue to refine this account.

6   Discussion

Compared to other weighted-constraint frameworks such as Harmonic Grammar, MaxEnt allows variation among multiple possible optima in the output and directly models probability distributions that are comparable to the observed proportions in production data. The formal analysis of section 5 demonstrates the extent to which our MaxEnt grammar is able to simulate the observed variabilities in the two varieties of Mandarin L1 and English L2 output probabilities of loVN and loVN.C. For instance, L1 group differences carry over to L2 in terms of nasal place change: NM speakers applied more place change in L1 Mandarin loVN + mismatched C than CM speakers, and the same pattern was observed in L2 English illegal loVN. In our corresponding analysis, NM speakers start with a lower weight of IDENT-PLACE compared to the CM speakers, which is then transferred and even slightly amplified in L2 learning.

This work in progress also attempts to capture an asymmetry in the grammar of legal vs. illegal VN with mismatched NC in English compounds. The constraints and weighting conditions translated from the legal context generate the expected preference for the fully place matched VNC output sequences in the illegal context, yet not to the observed magnitude. One solution might lie in proposing additional constraint targeting segment-specific mapping when taking into account of all loVN combinations – that is, imposing separate penalties for alveolar → velar vs velar → alveolar mappings. We also propose a constraint along the brand of existential faithfulness to model the nasal deletion rates for loVN with a mismatched NC, and ongoing work continues to elaborate on this analysis.

The current work extends the existing literature of formal modeling of L1 phonological learning to the domain of L2, and make falsifiable predictions about phonological development. We also recognize that some of our observed L2 patterns cannot be just explained by transfer or learning – namely, more deletion of nasal coda in legal VN in L2 English as compared to L1 Mandarin. Future work can further investigate the sources of these non-transfer effects and explore how these unexpected behaviors emerge and stabilize as L2 phonological grammars evolve over time.

7   References

Becker, Michael, and Anne-Michelle Tessier (2011). Trajectories of faithfulness in child-specific phonology. Phonology 28, 163–196. doi: 10.1017/S0952675711000133

Broselow, Ellen (2004). Unmarked structures and emergent rankings in second language phonology. *International Journal of Bilingualism*, *8*(1), 51-65.

Chen, Marilyn Y. (2000). Acoustic analysis of simple vowels preceding a nasal in Standard Chinese. *J. Phon.* 28, 43–67.

Coetzee, Andries. W., and Joe Pater (2011). The place of variation in phonological theory. *The handbook of phonological theory*, 401-434.

Curtin, Suzanne, and Kie Zuraw (2002). Explaining constraint demotion in a developing system. Proc. Boston Univ. Conf. Lang. Dev. 26, 118–129.

Duanmu, San (2007). *The Phonology of Standard Chinese.* Oxford: Oxford University Press.

Escudero, Paola, and Paul Boersma (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition* 26, 551–585. doi: 10.1017/S0272263104040021.

Goldwater, Sharon, and Mark Johnson (2003). Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the workshop on variation within Optimality Theory* (pp. 111-120).

Hayes, Bruce (2004). "Phonological acquisition in optimality theory: the early stages," in Constraints in Phonological Acquisition, eds R. Kager, J. Pater and W. Zonneveld (Cambridge: CUP), 158–203.

Hayes, Bruce, and Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *Ling. Inq.* 39:3. doi: 10.1162/ling.2008.39.3.379

Itô, Junko (1988). *Syllable Theory in Prosodic Phonology.* Doctoral Dissertation, University of Massachusetts, Amherst.

Jarosz, Gaja. (2010). Implicational markedness and frequency in constraint-based computational models of phonological learning. J. Child Lang. 37, 565–606. doi: 10.1017/S0305000910000103

Jesney, Karen (2016). On the relationship between learning sequence and rate of acquisition. In G. Hansson, et al (eds) *Proceedings of AMP 2015.*

Kiparsky, Paul (1985). Some consequences of lexical phonology. *Phonology, 2*(1), 85-138.

Levelt, Clara C., Niels O. Schiller, and Willem J. M. Levelt (1999). A developmental grammar for syllable structure in the production of child language. Brain Lang. 68, 291–299.

Liao, Sishi, *et al.* (2022). Nasal Coda Loss in the Chengdu Dialect of Mandarin: Evidence from RT-MRI. *INTERSPEECH.* 1347-1351.

Luo, Mingqiong (2020). Processes and Consequences of Co-Articulation in Mandarin V1N.(C2) V2 Context: Phonology and Phonetics. *INTERSPEECH*, 641-645.

Luo, Mingqiong, Jeroen Van de Weijer, and Marjoleine Sloos (2020). A perceptually-based OT approach to opacity in Mandarin Chinese nasal rhymes, in *Explorations of Chinese Theoretical and Applied Linguistics*, eds D. Chen and D. Bell, Cambridge Scholar Publishing, 8–43.

Neergaard, Karl, Hongzhi Xu, and J.S. German (2022). Database of word-level statistics for Mandarin Chinese (DoWLS-MAN). *Behav Res, 54*, 987–1009. https://doi.org/10.3758/s13428-021-01620-7.

Schwartz, Bonnie D., and Rex A. Sprouse (1996). L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12, 40–72.

Shi, Xiujuan, Qibin Ran, and Feng Shi (2011). An Analysis of Sonorant Nasality in Beijing Mandarin. *ICPhS.* 17.

Struijke, Caro (2000*). Existential Faithfulness: A study of reduplicative TETU, feature movement and dissimilation*. Doctoral dissertation, University of Maryland College Park.

Vaden, Kenny, Harry Halpin and Greg Hickok (2009). *Irvine Phonotactic Online Dictionary*, Version 2.0. Available from http://www.iphod.com.

van Leussen, Jan Willem, and Paola Escudero (2015). Learning to perceive and recognize a second language: the L2LP model revised. *Frontiers in Psychology* 6:1000. doi: 10.3389/fpsyg.2015.01000.

Zhang, Sijia, and Anne-Michelle Tessier (2024). Modeling the consequences of an L1 grammar for L2 production: simulations, variation, and predictions. *Frontiers in Language Sciences,* 3, 1327600.

Zhang, Sijia (forthcoming). *Phonological learning in a second language: first language variability and computational simulations.* Doctoral dissertation, University of British Columbia.

Zhou, Chao, and Silke Hamann (2024). Modelling the acquisition of the Portuguese tap by L1-Mandarin learners: A BiPhon-HG account for individual differences, syllable-position effects and orthographic influences in L2 speech. *Glossa: a journal of general linguistics 9*(1). pp. 1–39. DOI: https://doi.org/10.16995/glossa.9692