

The Gradient Acceptability in Mandarin Nonword Judgment

Syllable well-formedness judgment experiments reveal that speakers possess detailed knowledge of phonotactics and use such knowledge to offer gradient phonotactic acceptability ratings on novel words (e.g., Hayes & White, 2013). Such gradient judgment originates from both grammatical factors and lexical statistics. This study investigates gradient phonotactics stemming from the violations of four types of grammatical constraints in Mandarin Chinese: 1) principled phonotactic constraints, 2) accidental phonotactic constraints, 3) allophonic restrictions, and 4) segmental-tonal cooccurrence restrictions. We first hypothesize that violations of principled phonotactic constraints, provided that they can be motivated on typological and functional grounds, will incur lower acceptability ratings than accidental phonotactic constraint violations. Moreover, given the experimental findings that listeners tend to be less attuned to allophonic differences than phonemic differences (e.g., Jaeger, 1980), and that the processing of lexical tones is disadvantaged compared to segmental information (e.g., Cutler and Chen, 1997), we also hypothesize that violations of allophonic and segmental-tonal cooccurrence restrictions will be more acceptable than principled and accidental phonotactic violations. The lexical statistics factor that we investigate is neighborhood density, which is hypothesized to be positively correlated with acceptability. But crucially, we hypothesize that the grammatical effects cannot be subsumed under the neighborhood density effect. A syllable well-formedness judgment experiment was conducted with native Mandarin speakers to examine how the grammatical and lexical statistics factors contribute to the variation in phonotactic acceptability judgment.

Mandarin provides an excellent test case for these hypotheses for the following reasons. First, Mandarin has a considerably simpler syllable structure than the Indo-European languages where the majority of the gradient phonotactic results were obtained: to form a Mandarin syllable, consonants, glides, and vowels are drawn to fill in a restrictive syllable structure CGVX (Duanmu, 2007). This on the one hand provides a stronger test for the presence of gradience, and on the other hand allows for an exhaustive study of all potential syllable types in a language. Second, there are relatively clear principled phonotactic constraints in Mandarin, as proposed by Duanmu & Yi (2015) below. These constraints can all be considered as variants of the Obligatory Contour Principle, which finds wide typological support and is rooted in perception and processing (e.g., Frisch et al. 2004).

- a. *HH: The vowel feature [+high] cannot occur in succession (e.g., *[mui] *[tyu]).
- b. *[Cor]_[Cor]: [Cor] cannot occur in both G and X (e.g., *[jai] *[pjei]).
- c. *[Lab]_[Lab]: [Lab] cannot occur in both G and X (e.g., *[wou] *[nwau]).
- d. Identical articulators cannot occur in succession in C and G (e.g., *[tʂjan] *[pwaŋ]).

Third, Mandarin has a rich set of allophonic variations in vowels, where variants [ə e o] belong to the mid vowel /ə/, and [a a e] to the low vowel /a/, as shown below.

- a. ə → ə / _ n, ŋ, # ə → e / j, ɥ _ #, or _ i ə → o / w _ #, or _ u
- b. a → a / _ i, n, # a → a / _ u, ŋ a → e / j, ɥ _

Finally, there are also segmental-tonal cooccurrence restrictions in Mandarin, as there are tonal gaps with existing syllables. E.g., [nei] does not occur with the high-level Tone 1.

To construct the stimuli for the acceptability rating experiment, all theoretically possible syllables under the CGVX structure (both existing and missing) were enumerated, using only the high-level tone. The V slot included all *surface* tones. These syllables were grouped into five types: **systematic gaps** (missing syllables that violate some principled phonotactic constraint), **segmental accidental gaps** (missing syllables without any principled phonotactic

violations), **allophonic gaps** (missing syllables that only violate allophonic restrictions), **tonal accidental gaps** (existing syllables that cannot bear the high-level tone), and **real words**. Forty syllables were then randomly drawn for each type to form 200 test stimuli. The 200 syllables were recorded by a native Beijing Mandarin speaker with phonetic training and normalized for intensity. Thirty native Mandarin speaker participants were asked to judge the acceptability of these auditory stimuli on a scale of 1 (bad) to 7 (good).

The data were fitted with a linear mixed-effects regression model, using z-score transformed rating scores as the dependent variable and stimulus type, neighborhood density, and stimulus duration as independent variables. Stimulus duration was included as a nuisance variable as the stimuli in different word types had different durations. Step-wise likelihood ratio tests suggested that the best model is $rating \sim type + duration + ND + type:duration + type:ND + (1|item)$. The ANOVA table of the model is shown in Table 1. Crucially, although neighborhood density and stimulus duration both significantly affect the rating, stimulus type not only significantly improves the model independently, but also accounts for the largest amount of variance in the rating result.

	Sum Sq.	Mean Sq.	NumDF	DenDF	F value	p value
<i>Type</i>	51.284	12.8211	4	185	28.5454	<.0001
<i>Duration</i>	6.905	6.9049	1	185	15.3733	.0001
<i>Neighborhood Density</i>	4.029	4.0291	1	185	8.9706	.0031
<i>Type:Duration</i>	4.929	1.2323	4	185	2.7437	.0299
<i>Type:Neighborhood Density</i>	3.455	0.8636	4	185	1.9228	.1084

Table 1 Type III analysis of variance table of the ratings model

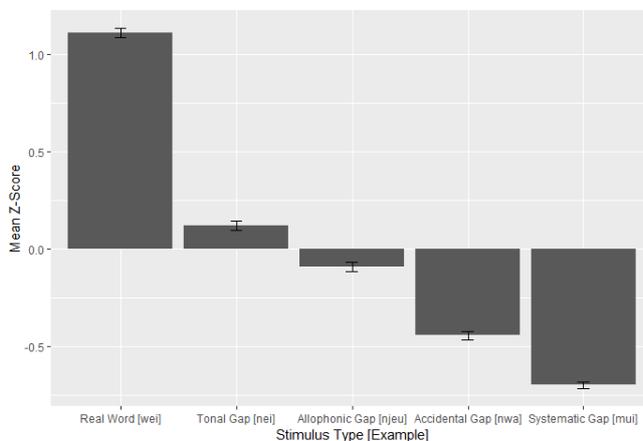


Figure 1 Mean z-scored ratings by stimulus type

The mean z-scored ratings for each stimulus type are shown in Figure 1, which indicates that there is a sharp contrast between the acceptability of real words and non-words, but the speakers' phonotactic judgment is clearly gradient. Switching the baseline from real words to other stimulus types showed that the acceptability difference between tonal and allophonic gaps is not significant, whereas all other pairs are significantly different from each other.

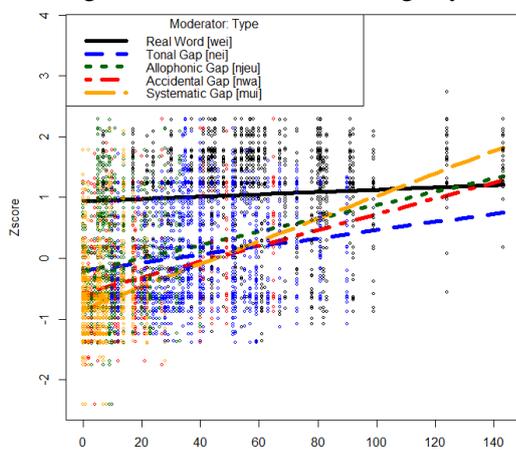


Figure 2 Effects of ND on ratings

Figure 2 shows that, overall, neighborhood density is positively correlated with the rating, but the effect in real words is weaker than in other word types.

We conclude that phonotactic acceptability is gradient in Mandarin nonce syllables. Although part of the gradient can be accounted for by neighborhood density and stimulus duration, the greatest contributor to the gradient is the grammatical factor of syllable type, with the violations of principled constraints leading to the lowest rating and the violations of tonal and allophonic gaps leading to the highest rating.