The Gradient Acceptability in Mandarin Nonword Judgment

Shuxiao Gong & Jie Zhang
University of Kansas
Theoretical Background

• Speakers offer gradient judgment on novel words (e.g. blick > bwick > bnick in English), which is guided by their phonotactic knowledge (Berent et al. 2007, Hayes & White 2013).

• Such gradient acceptability is reported mainly in Indo-European languages like English with complex syllable structures.

• It is worthwhile to explore the nature of gradient phonotactic acceptability in languages with a less complex syllable structure, for example Mandarin.
Sources of Acceptability Judgment

• Grammatical factors, e.g.,
  – Sonority Sequencing Principle (Berent et al. 2007)
  – Obligatory Contour Principle (Frisch et al. 2004)

• Lexical statistics, e.g.,
  – Neighborhood density (Bailey & Hahn 2001)

• Or, a combination of the two is more likely (Hayes & White 2013; Shademan 2007).

• The interaction of these factors leads to the gradient phonotactic acceptability.
Grammatical Factors

• Non-words violating some principled phonotactic constraints should be labelled as systematic gaps, whereas non-words that do not violate those constraints are accidental gaps (Chomsky & Halle 1965).

• The gradient acceptability implied:
  Real word > Accidental gap > Systematic gap
Systematic vs. Accidental Gaps

- Duanmu & Yi (2015) proposed four principled Mandarin phonotactic constraints.
  - *HH: The vowel feature [+high] cannot occur in succession. (e.g., *[mui] *[tyu])
  - *[Cor]_[Cor]: [Cor] cannot occur in both on-glide and coda. (e.g., *[jai] *[pjei])
  - *[Lab]_[Lab]: [Lab] cannot occur in both on-glide and coda. (e.g., *[wou] *[nwau])
  - Identical articulators cannot occur in succession in onset consonant and on-glide. (e.g., *[tʂjan] *[pwaŋ])
Systematic vs. Accidental Gaps

- All principled constraints are varieties of Obligatory Contour Principle (Leben, 1973; McCarthy, 1986).
- This similarity avoidance effect is grounded in psychological reality because adjacent similar sounds have been shown to present difficulties in production planning (Frisch et al., 2004) and in perception (Woods et al., 2010).
- Typologically, OCP effects in the lexicon are widely reported in many languages such as Arabic (Frisch & Zawaydeh, 2001), Hebrew (Berent & Shimron, 1997), Muna (Coetzee & Pater, 2008), and Quechua (Gallagher, 2010).
Allophonic Restrictions

• Previous studies mostly discussed phonotactic restrictions held on the phonemic level, and few looked into the phonotactic effects of allophonic distributions. (e.g. [spʰik] from English ‘speak’, [ʂwe] from Mandarin ‘说’).

• Because allophones of the same phoneme are often categorized under the same category by speakers due to perceptual similarity (Jaeger 1980).
Allophonic Restrictions

• But, allophonic distinctions can be reliably heard by speakers.
  – Jia et al. (2006) ask native Mandarin speakers to do ABX discrimination on [æ~ə] vowel pair in American English. This contrast is allophonic in Mandarin, yet their discrimination rate reaches 88.9%.

• Therefore, allophonic restrictions will contribute to phonotactic judgment, but the violation of such restrictions will not be as fatal.
Suprasegmental Phonotactics

• In Mandarin, not all existing syllables can freely combine with all of the four tones. Some syllable-tone combinations happen not to exist, e.g. [nei1].

• These missing syllable-tone combinations received significantly lower acceptability than real words in non-word judgment (Kirby & Yu 2007; Myers 2002).

• Co-occurrence patterns on segmental and suprasegmental levels are also likely to be noticed by speakers and form a part of their phonotactic grammar.
Suprasegmental Phonotactics

• Tones are perceived as conceptually different forms from other segmental level features (Hyman 2011).
• And the processing of lexical tones is disadvantaged compared to segmental information (Cutler & Chen 1997).
• Prediction on acceptability: lower than real words, but higher than other segmental gaps.
Lexical Statistics: Neighborhood Density

- Speakers may evaluate how similar a word is with other known words.
- One way to quantify this is neighborhood density. It is defined as the number of words generated by substituting, deleting, or adding a single phoneme (Greenberg & Jenkins, 1964).
- Numerous studies have shown that neighborhood density plays an essential role in spoken word perception and production (see Vitevitch & Luce (2016) for a review).
- In acceptability judgment tasks, neighborhood density is positively correlated with ratings (Kirby & Yu, 2007; Myers & Tsay, 2005).
Research Question

• Is non-word acceptability judgment gradient in Mandarin?

• How do the grammatical and lexical statistics factors mentioned above contribute to the judgment variation?

• A syllable well-formedness judgment experiment was conducted to probe into these questions.
Mandarin Overview

• Mandarin shows rich allophony in vowels.
  \[ \begin{align*}
  \varepsilon & \rightarrow o / w _ \# , \text{or} _ u \\
  \varepsilon & \rightarrow e / j , y _ \# , \text{or} _ i \\
  \varepsilon & \rightarrow e / _ n , \eta , \# \\
  \alpha & \rightarrow a / _ i , n , \# \\
  \alpha & \rightarrow a / _ u , \eta \\
  \alpha & \rightarrow e / j , y _ n
  \end{align*} \]

• Mandarin has a (comparably) simple and restrictive syllable structure (namely CGVX), so it is possible to work out a list of all theoretically possible syllables.
  • Onset Consonant: p \ p^h \ m \ f \ t \ t^h \ n \ l \ ts \ tsh \ s \ ts \ tsh^h \ s \ z \ tse \ teh \ s \ k \ k^h \ x
  • Glide: j \ w \ y
  • Surface Vowel: i \ u \ y \ e \ \varepsilon \ o \ a \ \alpha
  • Extra Ending Sound: i \ u \ n \ \eta
Stimuli

• We first listed all theoretically possible Mandarin syllables, both existing and missing syllables.
  – Under the CGVX syllable structure, only the vowel is obligatory. Factorial combination of all sounds gives rise to \((21+1) \times (3+1) \times 8 \times (4+1) = 3,520\) possible syllables, including 384 existing syllables and 3,136 missing syllables.

• Only the high-level Tone 1 was used.
Stimuli

• Given that phonotactically illegal forms may cause perceptual difficulties (Dupoux et al. 1999), and this study is only concerned with the acceptability of the syllables that can be reliably identified.

• To that effect, we implemented two steps:
  – 1,273 syllables that could potentially be confusable with other syllables were removed from the list.
  – An AX discrimination pretest was also designed to check whether participants can differentiate the allophonic differences used in the experiment. The overall accuracy was 91.3%.
Stimuli

• 384 existing syllables
  – 321 real words (e.g., wei)
  – 63 tonal accidental gaps, because they cannot take the high-level tone (e.g., nei)

• 1,863 missing syllables (excl. 1,273 perceptually ambiguous forms)
  – 434 allophonic gaps, gaps that only violate the allophonic rules (e.g., njeu)
  – 1,041 systematic gaps, gaps that violate some of the four major phonotactic constraints of Mandarin (e.g., mui)
  – 388 other segmental accidental gaps, gaps remain unexplained by the four constraints (e.g., nwa)
## Five Stimulus Types

<table>
<thead>
<tr>
<th>All Possible Syllables (3,520)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Existing Syllables</strong> (384)</td>
</tr>
<tr>
<td><strong>Real Words</strong> (321) [wei]</td>
</tr>
<tr>
<td><strong>Tonal Accidental Gaps</strong> (63) [nei]</td>
</tr>
<tr>
<td><strong>Allophonic Gaps</strong> (434) [njeu]</td>
</tr>
<tr>
<td><strong>Segmental Accidental Gaps</strong> (388) [nwa]</td>
</tr>
<tr>
<td><strong>Systematic Gaps</strong> (1,041) [mui]</td>
</tr>
<tr>
<td><strong>Undistinguishable Forms</strong> (1,273)</td>
</tr>
<tr>
<td><strong>Missing Syllables</strong> (3,136)</td>
</tr>
</tbody>
</table>

- Randomly select 40 syllables from each of the five types as the test stimuli.
- Test stimuli were recorded by a native Beijing Mandarin speaker with phonetic training, and then normalized for intensity.
Participants & Procedure

- 30 native Mandarin speakers born and raised in Northern China. Average age: 24 years old.

- Experiment procedure:
  - 5 practice trials and 200 test trials.
  - Participants heard stimuli from earphones.
  - The instruction asked the participants to rate whether the stimulus sound like Mandarin or not, on a scale from 1 (bad) to 7 (good).
  - Participants clicked on one of the seven rating buttons.
  - 500 ms pause between two trials.
  - Rating responses were recorded.
Data Analysis

• Linear mixed-effects regression model
  – Dependent variable: z-score transformed ratings out of each subject
  – Random effects: intercept of **Item**
  – Fixed effects: *five stimulus types; neighborhood density; duration*
    • *Duration* was also z-score transformed to avoid excessively distinct scaling
Data Analysis

- Both forward and backward step-wise algorithms were attempted for searching for the best model, and the two types of algorithms agreed on the same final model.

- rating ~ type + ND + duration + type:ND + type:duration + (1|item)
Switching the baseline from Real Word to other stimulus types indicates that the acceptability difference between tonal gaps and allophonic gaps is not significant, whereas all other type pairs are.
Duration by Type

- The stimuli tend to be longer for more ungrammatical stimulus types.
Stimulus Type : Duration Interaction

- But, the effect of duration varies for different stimulus types.
Stimulus Type: ND Interaction

- Neighborhood density, on the other hand, is positively correlated with all five stimulus types overall.
- Less correlation in real words (Kirby & Yu 2007; Myers & Tsay 2005).
The ANOVA table illustrates that, after duration and neighborhood density are taken into account, the five stimulus types still stand out as the most significant predictor of acceptability judgment.
Further Gradience

- The five stimulus types can be further subcategorized.
- Result in further gradience within each type.
Tonal Gap

• Among the 63 Tone 1 tonal gaps covered in this study, 40 of them start with a sonorant onset [m n l ʐ].
• This is because sonorant onset syllables with Tone 1 lack historical sources.
  – Modern Mandarin sonorant onsets [m n l ʐ] are predominantly derived from Mid Chinese voiced sounds. Yet, the syllables carrying Tone 1 in modern Mandarin develop from Mid Chinese syllables with voiceless onsets only.
• Rating results suggest that speakers are aware of this trend in the lexicon.
Tonal Gap
Systematic Gap

• Plot the mean acceptability of systematic gaps by their constraint violation.
  – Only violate *HH (e.g., [mui])
  – Only violate *[Cor]_[Cor] (e.g., [jai])
  – Only violate *[Lab]_[Lab] (e.g., [wou])
  – Only violate Articulator Dissimilation (e.g., [tʃjan])
Systematic Gap

Mean Z-Scores of Systematic Gaps' Ratings by Constraint Violation

- HH [mui]
- ArtDiss [tʃjan]
- *UU[wou]
- *l [jai]
Summary

- Mandarin non-word judgment is gradient: Real words > Tonal Gaps > Allophonic Gaps > Accidental Gaps > Systematic Gaps
- Multiple grammatical factors contribute to the gradience, including OCP-based phonotactic constraints, allophonic restrictions, syllable-tone combination restrictions.
- Lexical statistics contributes to the gradient acceptability as well. But, most variation is explained by the grammatical factors.
Thank you!
References

References

Stimuli

• However, not all syllables will be considered. Many sequences will very likely be misperceived as other existing forms or hard to distinguish from each other. I propose following criteria to rule out ambiguous forms.
  – Y-Spreading: the [y] vowel must spread [+front] and [+round] to the preceding glide; therefore, glide distinctions are all neutralized before [y] (e.g., [jy]=[wy]=[ʊy]).
  – U-Spreading: the [u] vowel must spread [+round] to the preceding glide (e.g., [ju]=[ʊu]).
  – Palatal Spreading: Palatals must cooccur with coronal glides (e.g., [tʃ]=[tʃj], [tʃw]=[tʃʊ]).
  – [oŋ] = [uŋ] because the pronunciation of the high back vowel before the velar nasal is more open and lax (Chao, 1968), which might be confused with [o].
  – Studies from loanword phonology show that the allophonically inappropriate [aŋ] will always be perceived as [an] by native speakers, and [ɑn] as [ɑŋ] (Hsieh, Kenstowicz, & Mou, 2009).

• These criteria mark 1,273 syllables as undistinguishable from some other syllables. Therefore, the remaining list contains 1,863 missing syllables and 384 existing syllables.
## Results

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std Error</th>
<th>t value</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>0.9396</td>
<td>0.1449</td>
<td>6.485</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>Tonal Gap</td>
<td>-1.1405</td>
<td>0.2023</td>
<td>-5.637</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>Allophonic Gap</td>
<td>-1.1582</td>
<td>0.1720</td>
<td>-6.734</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>Accidental Gap</td>
<td>-1.5142</td>
<td>0.1697</td>
<td>-8.924</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>Systematic Gap</td>
<td>-1.7807</td>
<td>0.1746</td>
<td>-10.198</td>
<td>&lt; .0001*</td>
</tr>
<tr>
<td>Duration</td>
<td>-0.0982</td>
<td>0.0675</td>
<td>-1.455</td>
<td>.1472</td>
</tr>
<tr>
<td>Neighborhood Density</td>
<td>0.0019</td>
<td>0.0023</td>
<td>0.826</td>
<td>.4100</td>
</tr>
<tr>
<td>Tonal Gap : Duration</td>
<td>-0.1089</td>
<td>0.0993</td>
<td>-1.096</td>
<td>.2744</td>
</tr>
<tr>
<td>Allophonic Gap : Duration</td>
<td>0.0371</td>
<td>0.0956</td>
<td>0.388</td>
<td>.6988</td>
</tr>
<tr>
<td>Accidental Gap : Duration</td>
<td>-0.0983</td>
<td>0.0983</td>
<td>-1.000</td>
<td>.3186</td>
</tr>
<tr>
<td>Systematic Gap : Duration</td>
<td>0.1881</td>
<td>0.0998</td>
<td>1.886</td>
<td>.0609</td>
</tr>
<tr>
<td>Tonal Gap : ND</td>
<td>0.0048</td>
<td>0.0039</td>
<td>1.256</td>
<td>.2106</td>
</tr>
<tr>
<td>Allophonic Gap : ND</td>
<td>0.0091</td>
<td>0.0066</td>
<td>1.367</td>
<td>.1733</td>
</tr>
<tr>
<td>Accidental Gap : ND</td>
<td>0.0111</td>
<td>0.0053</td>
<td>2.089</td>
<td>.0381*</td>
</tr>
<tr>
<td>Systematic Gap : ND</td>
<td>0.0168</td>
<td>0.0010</td>
<td>1.682</td>
<td>.0942</td>
</tr>
</tbody>
</table>
An Additive Effect

- Systematic gaps and accidental gaps may or may not violate allophonic restrictions.
  - For example, the systematic gap [nwau] not only violates the labial cooccurrence constraint, but also is allophonically inappropriate since the low vowel /a/ before the off-glide [u] should surface as the back [ɑ], not the front [a].

- Rating data suggest that gaps that also violate allophonic restrictions are judged worse than those who do not.
An Additive Effect