Advanced Second Language Learners’ Perception of Lexical Tone Contrasts

SLRF 2015

Eric Pelzl¹, Robert DeKeyser¹, Ellen Lau², Colin Phillips²

¹ Second Language Acquisition Program, University of Maryland, College Park
² Linguistics Department, University of Maryland, College Park
Aim: Are *lexical* tones difficult for *advanced* L2 learners of Chinese? If so, why?

1. Research context

2. Current study
   - Experiment 1: Syllables
   - Experiment 2: Words
   - Experiment 3: Sentences

3. Discussion
Growing interest in how difficulties with phonological distinctions in a L2 might result in lexical ambiguity (Broersma & Cutler, 2011; Sebastián-Gallés & Díaz, 2012)

Example: Japanese learners of English

/r/ /l/ confusion — Did he say “rock” or “lock”?

“Phonolexical” ambiguities explored in variety of L1-L2 pairs (Broersma & Cutler 2011; Chrabaszcz & Gor, 2014; Darcy et al. 2013; Díaz et al., 2012; White et al., 2015)

Do lexical tones cause such difficulties?
Four tones of Mandarin Chinese

mā
‘mom’

má
‘hemp’

mǎ
‘horse’

mà
‘scold’
L2 learners of Mandarin

- Initial difficulties, and perhaps long-term difficulties with confusion of T2 and T3 (Wang et al., 1999; Lee, Tao, & Bond, 2010; Sun, 1997)

- Promising results for tone training (dependent on pitch aptitude) (e.g., Wang et al., 1999; Wong & Perrachione, 2007)

- Limited research with more advanced learners & more complex stimuli suggests advanced learners may still struggle with tones to some degree (Hao, 2008; Sun, 1997; Zhang, 2011)
Current Study
Participants

- **13 advanced L2 speakers** of Mandarin Chinese (6F)
  - Recruited in Beijing and Maryland
  - All native English speakers
  - Screening measures to insure high proficiency

- **20 native speakers** of Mandarin (14F)
  - Recruited in Beijing
### Participants

*Background & Proficiency scores for L2 participants (n=13)*

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>range</th>
<th>sd</th>
</tr>
</thead>
<tbody>
<tr>
<td>age at testing</td>
<td>28.2</td>
<td>22-45</td>
<td>5.9</td>
</tr>
<tr>
<td>age of onset</td>
<td>18.7</td>
<td>16-24</td>
<td>2.0</td>
</tr>
<tr>
<td>semesters of study</td>
<td>7.4</td>
<td>2-19</td>
<td>4.1</td>
</tr>
<tr>
<td>years in immersion</td>
<td>2.2</td>
<td>0.4-6</td>
<td>2.2</td>
</tr>
<tr>
<td>total years learning</td>
<td>9.5</td>
<td>4-27</td>
<td>6.3</td>
</tr>
<tr>
<td>Can Do score</td>
<td>105.6</td>
<td>84-117</td>
<td>10.4</td>
</tr>
<tr>
<td>vocab score</td>
<td>81.9</td>
<td>60.9-97.1</td>
<td>10.8</td>
</tr>
</tbody>
</table>
Experiment 1: Tones in monosyllables

Are advanced L2 learners of Mandarin Chinese able to identify tones on isolated syllables and how does their performance compare to that of L1 speakers?

- **Tone identification**: hear a word, identify the tone; no concern for word meaning

- **Stimuli**
  - 60 monosyllabic real words
  - 15 of each tone
  - Spliced from sentences to increase task difficulty

*Example:* gāo / xióng / běn / qù
Experiment 1: tone identification trials

- "beep" 350ms
- 200ms silence
- "hóu" 800ms silence
- response (1, 2, 3, 4)
- "beep" 350ms
- 200ms silence
- "hóu" response
- etc.
Experiment 1: Tones in monosyllables

Are advanced L2 learners of Mandarin Chinese able to identify tones on isolated syllables and how does their performance compare to that of L1 speakers?

- Predicted results: L2 group will perform like native speakers for T1 and T4, with a slight drop in performance for T2 and T3
# Experiment 1: Tone ID results

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>L1</td>
<td>0.92</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.90</td>
<td>0.30</td>
</tr>
<tr>
<td>T2</td>
<td>L1</td>
<td>0.93</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.76</td>
<td>0.43</td>
</tr>
<tr>
<td>T3</td>
<td>L1</td>
<td>0.79</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.79</td>
<td>0.44</td>
</tr>
<tr>
<td>T4</td>
<td>L1</td>
<td>0.85</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.83</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Means and standard deviations for accuracy by tone condition in the Tone ID task
## Experiment 1: Tone ID results

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>L1</td>
<td>0.92</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.90</td>
<td>0.02</td>
</tr>
<tr>
<td>T2</td>
<td>L1</td>
<td>0.93</td>
<td>0.25</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.76</td>
<td>0.43</td>
</tr>
<tr>
<td>T3</td>
<td>L1</td>
<td>0.79</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.79</td>
<td>0.44</td>
</tr>
<tr>
<td>T4</td>
<td>L1</td>
<td>0.85</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.83</td>
<td>0.38</td>
</tr>
</tbody>
</table>

Means and standard deviations for accuracy by tone condition in the Tone ID task.
Experiment 1: Tone ID results

Percent correct responses

Tone 1

individual score

range of L1 performance

chance
• Generally both L1 and L2 highly accurate
• L2 group differs from L1 in accuracy for T2
Advanced L2 learners of Mandarin can largely perform like natives on a tone identification task, though T2 remains difficult for many L2 learners.
Lexical Decision Task: hear an item, decide if it is a word or not a word

Stimuli
- 240 disyllabic real words, 120 segmental nonwords, 120 tonal nonwords (all spliced from sentences)
- A single list contains: 60 real words, 30 segmental nonwords, 30 tonal nonwords

Example: gāozhōng ‘high school’ / guāzhōng / gǎozhōng

Experiment 2: Tones in words

Are advanced L2 learners of Mandarin Chinese able to use tonal and/or segmental cues to distinguish isolated real words from nonwords and how does their performance compare to that of L1 speakers?
Experiment 2: lexical decision trials

- "beep" 350ms
- 200ms silence
- "gǎozhōng" response (‘yes’, ‘no’)
- 1000ms silence
- "beep" 350ms
- 200ms silence
- "hóuzi" response
- etc.

etc.
Experiment 2: Tones in words

Are advanced L2 learners of Mandarin Chinese able to use tonal and/or segmental cues to distinguish isolated real words from nonwords and how does their performance compare to that of L1 speakers?

- Predicted results: Strong L2 performance on segmental nonwords, but much weaker performance on tonal nonwords
## Experiment 2: lexical decision results

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>real words</td>
<td>L1</td>
<td>0.94</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.91</td>
<td>0.29</td>
</tr>
<tr>
<td>segmental nonwords</td>
<td>L1</td>
<td>0.96</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.82</td>
<td>0.38</td>
</tr>
<tr>
<td>tonal nonwords</td>
<td>L1</td>
<td>0.91</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.29</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Means and standard deviations for accuracy by condition in the LDT.
## Experiment 2: lexical decision results

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>real words</td>
<td>L1</td>
<td>0.94</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.91</td>
<td>0.29</td>
</tr>
<tr>
<td>segmental nonwords</td>
<td>L1</td>
<td>0.96</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.82</td>
<td>0.38</td>
</tr>
<tr>
<td>tonal nonwords</td>
<td>L1</td>
<td>0.91</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.29</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Means and standard deviations for accuracy by condition in the LDT
## Experiment 2: Lexical Decision Results

<table>
<thead>
<tr>
<th>Condition</th>
<th>Group</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>real words</td>
<td>L1</td>
<td>0.94</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.91</td>
<td>0.29</td>
</tr>
<tr>
<td>segmental nonwords</td>
<td>L1</td>
<td>0.96</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.82</td>
<td>0.38</td>
</tr>
<tr>
<td>tonal nonwords</td>
<td>L1</td>
<td>0.91</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>L2</td>
<td>0.29</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Means and standard deviations for accuracy by condition in the LDT
• Many L2ers perform in L1 range for real words & seg. nonwords
• NO L2ers in L1 range for tonal nonwords
Advanced L2 learners of Mandarin can use segmental cues to identify nonwords, but struggle to utilize tonal cues to differentiate words from nonwords.

Effect of *lexical* tones (vs. perception of tones *per se*)

- Effect of *disyllabic* stimuli? (cf. Hao, 2008; Sun, 1997; Zhang, 2011)
- Effect of not knowing targeted *vocabulary*?
- Effect of uncertain/incomplete *explicit tone knowledge*?
Experiment 3: Tones in Sentences

Do advanced L2 learners of Mandarin Chinese display evidence of early sensitivity to tonal and segmental errors in sentential context as indexed by the N400?

- Event-Related Potential (ERP) Experiment
Experiment 3: Tones in Sentences

- Stimuli:
  - 120 sentence quadruplets 480 total tokens
  - Each sentence has four conditions, occurring with:
    1. an **expected** real word
    2. a **semantically** mismatching word
    3. a **segmentally** mismatching nonword
    4. a **tonally** mismatching nonword
  - Each participant hears 120 sentences, 30 per condition
  - Four lists balanced across participants
  - 60 un-manipulated, well-formed filler sentences
Example item

我的父亲有两个姐妹，但没有____。
Wǒde fùqīn yǒu liǎngge jiēmèi, dàn méiyòu _____.
“My father has two sisters, but no ______.”

Expected:  xiōngdì  “brothers”
Semantic mismatch:  càidān  “menu”
Segmental mismatch:  xuēdī  [nonword]
Tonal mismatch:  xióngdì  [nonword]
The N400

- Previous studies found an N400 response to tone mismatches with monosyllabic stimuli (Mandarin: Brown-Schmidt & Canseco-Gonzalez, 2004; Cantonese: Schirmer et al., 2005; Li, Yang, & Hagoort, 2008)

- N400 displays an increase in amplitude inversely related to the expectedness of a target word (e.g., Kutas and Hillyard, 1980; Kutas and Hillyard, 1984; Kutas & Federmeier, 2000; DeLong, Urbach, and Kutas, 2005)
Experiment 3: Tones in Sentences

- Event-Related Potential (ERP) Experiment
  Recorded while listening, locked to critical words

- Behavioral Task
  After each sentence, judge whether it sounded ‘OK’
  Practice training stimuli to understand ‘OK’
  Not told which word is manipulated
  Location of critical words varied across sentences
Do advanced L2 learners of Mandarin Chinese display evidence of early sensitivity to tonal and segmental errors in sentential context as indexed by the N400?

Predicted results: L2 learners will show N400 responses to semantic and segmental mismatches, but not to tonal mismatches.
Main effect of group (L1>L2) \[ z=-6.80; \ p<.001 \]
(some of) L2 group had difficulty understanding stimuli
Experiment 3: ERP results

L1 (n=18)

-2
2
-100 100 300 500 700
200-450ms 500-800ms

expected
semantic
segmental
tonal

我的父亲有两个姐妹，但没有兄弟。
“My father has two sisters, but no brothers.”
Experiment 3: ERP results

L1 (n=18)

expected
semantic
segmental
tonal

Cz

-100 100 300 500 700

-2 2

我的父亲有两个姐妹，但没有菜单。
“My father has two sisters, but no menu.”
Experiment 3: ERP results

L1 (n=18)

black: expected
light green: semantic
light blue: segmental
red: tonal

```
我的父亲有两个姐妹，但没有xuēdì。
“My father has two sisters, but no [brathers].”
```
**Experiment 3: ERP results**

L1 (n=18)

```
Cz
```

```
<table>
<thead>
<tr>
<th>expected</th>
<th>semantic</th>
<th>segmental</th>
<th>tonal</th>
</tr>
</thead>
</table>
```

```
My father has two sisters, but no [bróthers].
```

我的父亲有两个姐妹，但没有xióngdì。
Experiment 3: ERP results

L1 (n=18)
Experiment 3: ERP results

L2 (n=12)

expected
semantic
segmental
tonal

Cz

我的父亲有两个姐妹，但没有兄弟。
“My father has two sisters, but no brothers.”
Experiment 3: ERP results

L2 (n=12)

expected
semantic
segmental
tonal

Cz

-2

-100 100 300 500 700

我的父亲有两个姐妹，但没有菜单。
“My father has two sisters, but no menu.”
Experiment 3: ERP results

L2 (n=12)

Cz

expected
semantic
segmental
tonal

我的父亲有两个姐妹，但没有xuēdì。
“My father has two sisters, but no [brathers].”
Experiment 3: ERP results

`xióngdì`.

“My father has two sisters, but no [bróthers].”
Experiment 3: ERP results

L2 (n=12)

- expected
- semantic
- segmental
- tonal
Discussion

- Perceiving tones *per se* may not be a challenge for advanced L2 learners.

- Utilizing tones to recognize words (at least disyllabic words) appears to remain a major challenge even for advanced learners.

- As in some previous studies, we see a seeming discontinuity between performance on a phonetic and phonolexical tasks (e.g., Darcy et al., 2013; Díaz et al. 2012).
Questions/suggestions are welcome and appreciated!

Thanks to advisors:

Robert DeKeyser
Ellen Lau
Colin Phillips

Thanks to international collaborators:

Taomei Guo
Junjie Wu, Di Lu, Chunyan Kang, Yongben Fu
Shuai Hao, Wenli Zhang

Thanks to classmates and undergrad assistants:

Man Li, Helen Zhou, Jiankun Liu, Brendan Cone

Thanks to participants past (and future)!

Research supported in part by:
NSF-IGERT grant #0801465
NSF-EAPSI grant #1514936


