Processing Missing Vowels: 
Effects of Phonological and Phonotactic Knowledge in Japanese

- People have a good idea about their native language (consciously/unconsciously).
- They use knowledge of their native language when they talk (produce) / hear (perceive) utterances.

**Broad Q:** How does the language-specific knowledge affect production or perception of a certain sound?

**Narrow Q:** How does the language-specific knowledge affect perception of voiceless vowels in Japanese?

0. **BACKGROUND: JAPANESE (JPN) PHONOLOGY/PHONOTACTICS**

**Terminology**
phonology: the study of how speech sounds are organized.
phonotactics: permissible combinations of phonemes.

(1) Five vowels: [i, e, a, o, u,]

Table 1: ENG & JPN Vowels

<table>
<thead>
<tr>
<th></th>
<th>ENG</th>
<th>JPN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
<td>F2</td>
</tr>
<tr>
<td>i</td>
<td>403</td>
<td>2917</td>
</tr>
<tr>
<td>e</td>
<td>700</td>
<td>2500</td>
</tr>
<tr>
<td>a</td>
<td>1114</td>
<td>1528</td>
</tr>
<tr>
<td>o</td>
<td>436</td>
<td>915</td>
</tr>
<tr>
<td>u</td>
<td>551</td>
<td>1528</td>
</tr>
</tbody>
</table>
(2) High vowels [i, u] are voiceless in certain consonantal environments.
- In what environments?
  A: between [-voi] Cs  
    *kita* ‘north’ [kitɑ], *futon* [fu tô:n],
    *yakisoba* ‘fried noodles’ [jakʲisوبا]

  B: between a [-voi] C and a pause:  
    *okashi* ‘snacks’ [okaɕi], *dasu* ‘to take out’ [dasu]

(3) What is a ‘voiceless vowel’?
- **devoiced vowels**: no voicing, no clear traces of F1 & F2. Sounds more fricative-like.

  Figure 1: Spectrograms of /hokito/  

- **deleted vowels**: merged with the preceding voiceless consonant, difficult to separate.

  Figure 2: Spectrograms of /mešite/
(4) Why does vowel devoicing occur?
- Vowel duration: high vowels are shorter than non-high vowels
  \[ /u/ = 1.00, /i/ = 1.17, /o/ = 1.26, /e/ = 1.37, \text{ and } /a/ = 1.44 \] (Han 1962a)

<table>
<thead>
<tr>
<th></th>
<th>/paka/</th>
<th>/pika/</th>
<th>/puka/</th>
<th>/peka/</th>
<th>/poka/</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENG</td>
<td>106ms</td>
<td>101ms</td>
<td>112ms</td>
<td>106ms</td>
<td>108ms</td>
</tr>
<tr>
<td>JPN</td>
<td>81ms</td>
<td>57ms</td>
<td>59ms</td>
<td>77ms</td>
<td>58ms</td>
</tr>
</tbody>
</table>

- Speech tempo: vowel devoicing occurs at normal speech tempo (not at super-careful speech)

(5) A phonotactic constraint of Japanese (inhibition of consonant clusters)
- Basically CVCV… syllables
- No consonant clusters are allowed except coda nasals and geminates.
  😊 /kat.ta./ *katta* ‘bought’ /tom.bo./ *tombo* ‘dragon fly’
  😊 /kat/ /tob.mo./
- Loanwords
  😊 *McDonald’s [ma.ku.dō.na.fiu.do]*

I. EXPERIMENT 1

(6) Question Is voiceless [i] detected as quickly and as accurately as fully voiced [i] by native Japanese listeners?

(7) Conditions
A. 3 consonantal environments
   - [-V] (devoicing): /i/ is between two voiceless obstruents. → voiceless [i] is expected.
   - [+V] (voicing): /i/ follows /j/ → voiced [i] is expected.
   - [+Vnas] (voicing): /i/ follows a nasal → voiced [i] is expected.

B. 2 voicing status of vowel
   - voiceless [i]
   - voiced [i]
Table 3: Experimental Conditions

<table>
<thead>
<tr>
<th>Environment</th>
<th>Voiced</th>
<th>Voiceless</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-V]:</td>
<td>C[-voi] ____</td>
<td>C[-voi]</td>
</tr>
<tr>
<td></td>
<td>☺ [hokito]</td>
<td>☺ [hokito]</td>
</tr>
<tr>
<td>[+V]:</td>
<td>/j/ ____ C[+/- voi]</td>
<td>☺ [tajziga]</td>
</tr>
<tr>
<td></td>
<td>☺ [tajziga]</td>
<td>☺ [tajziga]</td>
</tr>
<tr>
<td>[+Vnas]:</td>
<td>/N/ ____ C[+/- voi]</td>
<td>☺ [kedâŋiga]</td>
</tr>
<tr>
<td></td>
<td>☺ [kedâŋiga]</td>
<td>☺ [kedâŋiga]</td>
</tr>
</tbody>
</table>

(8) Stimuli (see Appendix 1)
A. 180 experimental items (CVCiCV or CVCVCiCV). All nonsense words.
B. 30 [i] fillers + 300 fillers + 10 practice items
C. All words were produced by a female native speaker of the Tokyo dialect of Japanese.

(9) Procedure
A. I made two sets of word list.
B. Participants were randomly assigned to one of the two sets.
C. Participants were positioned in a quiet place.
D. Participants were asked to listen to stimuli and hit a button on a response box as soon as they heard /i/.
E. When subjects hit a button on a response box, the next word came up after a 1000ms pause. When subjects did not respond within 3000ms, the next word came up after a 1000ms pause.
F. Response time (RT) was recorded by a computer.

(10) Participants: 47 native speakers of the Tokyo dialect of Japanese.

(11) RT Results

Figure 3: Mean RTs for JPN Listeners in Phoneme Monitoring Task

The symbol ‘*’ indicates that the difference was statistically significant in the environment.
A. [-V] (devoicing) environment: no significant difference
   → voiceless [i] was detected as quickly as voiced [i].

Table 4: [-V] Devoicing Environment (/i/ between voiceless consonants)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>− (weak)</td>
</tr>
<tr>
<td>phonotactics: vowel should be between consonants</td>
<td>+ (facilitory)</td>
<td>+ (facilitory)</td>
</tr>
<tr>
<td>phonology: combination of environment &amp; vowel</td>
<td>− (bad)</td>
<td>+ (good)</td>
</tr>
<tr>
<td>results</td>
<td>detection speed</td>
<td>no difference</td>
</tr>
</tbody>
</table>

B. [+V] (voicing) environment: voiceless [i] was detected more slowly than voiced [i].

Table 5: [+V] Voicing Environment (/i/ follows /j/)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>− (weak)</td>
</tr>
<tr>
<td>phonotactics: vowel should be between consonants</td>
<td>+ (facilitory)</td>
<td>+ (facilitory)</td>
</tr>
<tr>
<td>phonology: combination of environment &amp; vowel</td>
<td>+ (good)</td>
<td>− (bad)</td>
</tr>
<tr>
<td>results</td>
<td>detection speed</td>
<td>faster</td>
</tr>
</tbody>
</table>

C. [+Vnas] (voicing) environment: voiceless [i] was detected more slowly than voiced [i].

Table 6: [+Vnas] Voicing Environment (/i/ follows a nasal)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>− (weak)</td>
</tr>
<tr>
<td>phonotactics: a vowel doesn’t need to be after a nasal</td>
<td>− (inhibitory)</td>
<td>− (inhibitory)</td>
</tr>
<tr>
<td>phonology: combination of environment &amp; vowel</td>
<td>+ (good)</td>
<td>− (bad)</td>
</tr>
<tr>
<td>results</td>
<td>detection speed</td>
<td>faster</td>
</tr>
</tbody>
</table>
(12) Error Rate Results (error = the number of times when /i/ was ignored.)

Figure 4: Mean Error Rates for JPN Listeners in Phoneme Monitoring Task

The symbol ‘*’ indicates that the difference was statistically significant in the environment.

A. [-V] (devoicing) environment: no significant difference

→ voiceless [i] was detected as accurately as voiced [i].

Table 7: [-V] Devoicing Environment (/i/ between voiceless consonants)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiceless [i]</th>
<th>voiced [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>– (weak)</td>
</tr>
<tr>
<td>phonotactics: vowel should be between consonants</td>
<td>+ (facilitory)</td>
<td>+ (facilitory)</td>
</tr>
<tr>
<td>phonology: combination of environment &amp; vowel</td>
<td>– (bad)</td>
<td>+ (good)</td>
</tr>
<tr>
<td>results errors</td>
<td></td>
<td>no difference</td>
</tr>
</tbody>
</table>

B. [+V] (voicing) environment: no significant difference

→ voiceless [i] was detected as accurately as voiced [i].

Table 8: [+V] Voicing Environment (/i/ follows /j/)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiceless [i]</th>
<th>voiced [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>– (weak)</td>
</tr>
<tr>
<td>phonotactics: vowel should be between consonants</td>
<td>+ (facilitory)</td>
<td>+ (facilitory)</td>
</tr>
<tr>
<td>phonology: combination of environment &amp; vowel</td>
<td>+ (good)</td>
<td>– (bad)</td>
</tr>
</tbody>
</table>
| results errors                 |               | no difference → WHY???
C.  [+Vnas] (voicing) environment: voiceless [i] was ignored more than voiced [i].

Table 9: [+Vnas] Voicing Environment (/i/ follows a nasal)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>– (weak)</td>
</tr>
<tr>
<td>phonotactics:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a vowel doesn’t need to be after</td>
<td>– (inhibitory)</td>
<td>– (inhibitory)</td>
</tr>
<tr>
<td>a nasal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>phonology:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>combination of environment &amp; vowel</td>
<td>+ (good)</td>
<td>– (bad)</td>
</tr>
<tr>
<td>results</td>
<td>errors</td>
<td>fewer</td>
</tr>
</tbody>
</table>

II. EXPERIMENT 2

(13) Question Is voiceless [i] detected as quickly and as accurately as fully voiced [i] by native listeners of American English?

(14) Task, Conditions, Procedures: same as Experiment 1

(15) Participants: 45 native listeners of American English at U of A.

(16) RT Results

Figure 5: Mean RTs for AE Listeners in Phoneme Monitoring Task
(17) Error Rate Results

Figure 6: Mean Error Rates for AE Listeners in Phoneme Monitoring Task

![Error Rate Results](image)

In all environments, voiceless [j] was detected more slowly and was ignored more than voiced [i]. AE listeners don’t have Japanese-specific phonological and phonotactic knowledge. They relied on acoustic cues.

III. EXPERIMENT 3

(18) Question Are lexical words with voiceless [j] recognized as quickly and as accurately as the same words with fully voiced [i] by native Japanese listeners?

(19) Conditions: same as Experiment 1 but with real words

(20) Conditions: same as Experiment 1

Table 10: Experimental Conditions

<table>
<thead>
<tr>
<th>Environment</th>
<th>Voiced</th>
<th>Voiceless(^1)</th>
<th>Glossary</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-V]:</td>
<td>[acîta]</td>
<td>[acîta]</td>
<td>ashita ‘tomorrow’</td>
</tr>
<tr>
<td>[+V]:</td>
<td>[kujîra]</td>
<td>[kujîra]</td>
<td>kujîra ‘whale’</td>
</tr>
<tr>
<td>[+V(_{nas})]:</td>
<td>[yônîge]</td>
<td>[yônîge]</td>
<td>yônîge ‘to flee by night’</td>
</tr>
</tbody>
</table>

\(^{1}\) Voiceless vowels include both devoiced and deleted vowels.
(21) Stimuli: (see Appendix 2)
   A. 120 experimental items. All real words.
   D. 30 real word fillers + 200 non-word fillers + 10 practice items
   E. All words were produced by a female native speaker of the Tokyo dialect of Japanese.

(22) Procedure: same as Experiment 1 except I asked participants to hit a button on a response box as soon as they heard a real word.

(23) Participants: 47 native speakers of the Tokyo dialect of Japanese.

(24) RT Results

Figure 7: Mean RTs for JPN Listeners in Lexical Decision Task

A. [-V] (devoicing) environment: words with voiceless [i] was detected as quickly as words with voiced [i].

Table 11: [-V] Devoicing Environment (/i/ between voiceless consonants)

<table>
<thead>
<tr>
<th>effects</th>
<th>words with voiced [i]</th>
<th>words with voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>phonotactics: vowel should be between consonants</td>
<td>+ (facilitory)</td>
<td>+ (facilitory)</td>
</tr>
<tr>
<td>phonology: combination of environment &amp; vowel</td>
<td>– (bad)</td>
<td>+ (good)</td>
</tr>
<tr>
<td>results</td>
<td>detection speed</td>
<td>slower</td>
</tr>
</tbody>
</table>

faster \(\Rightarrow\) WHY???
B. [+V] (voicing) environment: voiceless [i] was detected more slowly than voiced [i].

Table 12: [+V] Voicing Environment (/i/ follows /j/)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>– (weak)</td>
</tr>
<tr>
<td><strong>phonotactics:</strong> vowel should be between consonants</td>
<td>+ (facilitory)</td>
<td>+ (facilitory)</td>
</tr>
<tr>
<td><strong>phonology:</strong> combination of environment &amp; vowel</td>
<td>+ (good)</td>
<td>– (bad)</td>
</tr>
<tr>
<td><strong>results</strong> detection speed</td>
<td>faster</td>
<td>slower</td>
</tr>
</tbody>
</table>

C. [+Vnas] (voicing) environment: voiceless [i] was detected more slowly than voiced [i].

Table 13: [+Vnas] Voicing Environment (/i/ follows a nasal)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>– (weak)</td>
</tr>
<tr>
<td><strong>phonotactics:</strong> a vowel doesn’t need to be after a nasal</td>
<td>– (inhibitory)</td>
<td>– (inhibitory)</td>
</tr>
<tr>
<td><strong>phonology:</strong> combination of environment &amp; vowel</td>
<td>+ (good)</td>
<td>– (bad)</td>
</tr>
<tr>
<td><strong>results</strong> detection speed</td>
<td>faster</td>
<td>slower</td>
</tr>
</tbody>
</table>

(25) Error Rates Results

Figure 8: Mean Error Rates for JPN Listeners in Lexical Decision Task
A.  [-V] (devoicing) environment: words with voiceless [i] had fewer errors than words with voiced [i].

Table 14: [-V] Devoicing Environment (/i/ between voiceless consonants)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>– (weak)</td>
</tr>
<tr>
<td>phonotactics:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vowel should be between consonants</td>
<td>+ (facilitory)</td>
<td>+ (facilitory)</td>
</tr>
<tr>
<td>phonology:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>combination of environment &amp; vowel</td>
<td>– (bad)</td>
<td>+ (good)</td>
</tr>
<tr>
<td>results</td>
<td>errors</td>
<td>more</td>
</tr>
</tbody>
</table>

B.  [+V] (voicing) environment: words with voiceless [i] had more errors than words with voiced [i].

Table 15: [+V] Voicing Environment (/i/ follows /j/)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>– (weak)</td>
</tr>
<tr>
<td>phonotactics:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>vowel should be between consonants</td>
<td>+ (facilitory)</td>
<td>+ (facilitory)</td>
</tr>
<tr>
<td>phonology:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>combination of environment &amp; vowel</td>
<td>+ (good)</td>
<td>– (bad)</td>
</tr>
<tr>
<td>results</td>
<td>errors</td>
<td>fewer</td>
</tr>
</tbody>
</table>

C.  [+Vnas] (voicing) environment: words with voiceless [i] had more errors than words with voiced [i].

Table 16: [+Vnas] Voicing Environment (/i/ follows a nasal)

<table>
<thead>
<tr>
<th>effects</th>
<th>voiced [i]</th>
<th>voiceless [i]</th>
</tr>
</thead>
<tbody>
<tr>
<td>acoustic cues</td>
<td>+ (strong)</td>
<td>– (weak)</td>
</tr>
<tr>
<td>phonotactics:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a vowel doesn’t need to be after a nasal</td>
<td>– (inhibitory)</td>
<td>– (inhibitory)</td>
</tr>
<tr>
<td>phonology:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>combination of environment &amp; vowel</td>
<td>+ (good)</td>
<td>– (bad)</td>
</tr>
<tr>
<td>results</td>
<td>errors</td>
<td>less</td>
</tr>
</tbody>
</table>
IV. CONCLUSION

- Listeners used phonotactic and phonological knowledge of their native language for detecting vowel /i/.

- Three factors effect differently in each experiment. [sāŋ:kuːfoːrisuŋŋiu] ☺

SELECTED REFERENCES


APPENDIX 1: Experimental Items for Experiments 1 & 2

- Devoicing (D-) Environment (30 words)

  - hokita
  - sekite
  - mekito
  - seshita
  - moshito
  - meshite
  - wachika
  - machike
  - nahika
  - mahiko
  - tekisa
  - kachiho
  - nehisa
  - hokito
  - yachite
  - samokika
  - toyakiko
  - nagahita
  - kotashike
  - hasechito
  - tadahika
  - saneshita
  - moneshito
  - wagashite
  - yawachike
  - motahike
  - nosahiko
  - kotashite
  - notochiko
  - menahisa

- Non-Devoicing (ND-) Environment (20 words)

  - tajida
  - najida
  - wajide
  - tejido
  - yojima
  - wajina
  - tajiga
  - tajida
  - kojiga
  - sojime
  - rejiza
  - kejiso
  - hojizo
  - mejina
  - yajide
  - natajima
  - kanojime
  - sasajina
  - ketajiba
  - mokojibe
  - narajizo
  - merajime
  - warejida
  - masajido
  - kanajida
  - warejina
  - tasajiba
  - yatejima
  - yotajino
  - towajina
### Non-Devoicing-Nasal (Nas-) Environment (20 words)

<table>
<thead>
<tr>
<th>wanide</th>
<th>tenido</th>
<th>saniza</th>
<th>kenima</th>
<th>nanime</th>
<th>tanina</th>
</tr>
</thead>
<tbody>
<tr>
<td>tanimo</td>
<td>yonida</td>
<td>waniba</td>
<td>nenigo</td>
<td>seniba</td>
<td>hanima</td>
</tr>
<tr>
<td>soniga</td>
<td>tonize</td>
<td>maniga</td>
<td>hasanina</td>
<td>yosanine</td>
<td>kasanide</td>
</tr>
<tr>
<td>naraniza</td>
<td>hokonimo</td>
<td>kedaniga</td>
<td>katonino</td>
<td>wasenida</td>
<td>tarenido</td>
</tr>
<tr>
<td>tasonibe</td>
<td>ketanime</td>
<td>samonima</td>
<td>yoranide</td>
<td>nasaniba</td>
<td>wadonima</td>
</tr>
</tbody>
</table>

### APPENDIX 2: Experimental Materials for Experiment 3

#### Devoicing (D-) Environment (20 words)

- akikan ‘empty can’
- hashika ‘measles’
- kachiku ‘domestic animals’
- tsukisasu ‘stub’
- akisu ‘robber’
- hisan ‘misery’
- akita ‘the name of a place’
- ashita ‘tomorrow’
- fukitsu ‘ill omen’
- onshitsu ‘a green house’
- koshitsu ‘a single room’
- hakike ‘nausea’
- fukikesu ‘to blow off’
- kakikomu ‘to write down’
- sekikomu ‘to cough’
- ochikomu ‘to get depressed’
- tekisetsu ‘appropriate’
- yakisoba ‘fried noodle’
- soshite ‘and then’
- tekitoo ‘reasonable’

#### Non-Devoicing (ND-) Environment (20 words)

- ajiwau ‘to taste’
- oojisama ‘prince’
- nekojita ‘too sensitive to hot food’
- sanjigen ‘three dimensions’
- kajiru ‘to bite’
- kejime ‘to be distinguishable’
- genjitsu ‘real world’
- kujira ‘whale’
- kujikeru ‘to be discouraged’
- tejina ‘conjuring tricks’
- majime ‘serious’
- sujigaki ‘plot’
- tojiru ‘to close’
- fujisan ‘Mt. Fuji’
- betsujin ‘different person’
- hajiku ‘to snap’
- soojiki ‘vacuum cleaner’
- najimu ‘to become familiar’
- nejiru ‘to wrench’
- hajimeru ‘to begin’

#### Non-Devoicing-Nasal (Nas-) Environment (20 words)

- onigiri ‘rice ball’
- tanigoe ‘crossing over a valley’
- tenimotsu ‘luggage’
- gyuunniku ‘beef’
- nanimono ‘whoever’
- wanigawa ‘crocodile skin’
- ganimata ‘bowlegged’
- inisharu ‘initial’
- tanima ‘ravine’
- kaniza ‘the Cancer’
- onigokko ‘playing tag’
- butaniku ‘pork’
- uraniwa ‘back yard’
- aniki ‘elder brother’
- banira ‘vanilla’
- hiniku ‘sarcasm’
- minikui ‘ugly’
- monitaa ‘monitor’
- yonige ‘to flee by night’
- kuinige ‘to run away without paying one’s bill’