Perceptual identification of high vowels in Taiwan Mandarin

Chenhao Chiu[1, 2], Jian-zhi Huang[1], Po-Hsuan Huang[1]
[1]Graduate Institute of Linguistics, National Taiwan University, [2]Neurobiology and Cognitive Science Center, National Taiwan University
chenhaochiu@ntu.edu.tw; r10142007@ntu.edu.tw; r09142003@ntu.edu.tw

Introduction

- Lip rounding is more of phonetic classification than phonemic identification (Lisker & Rossi, 1992).
  → Lip rounding can be individually variable: some speakers do not have rounded feature for English /ʃ/.
- Feature [round] is associated with lip rounding and protrusion (Catford, 1988; Lisker & Rossi, 1992; Jackson & McGowan, 2012, etc.).
- Three high vowels in Taiwan Mandarin: /i/, /u/, /y/.

	[back]	[round]
/i/	−	−
/u/	+	+
/y/	−	+

- [round] provides more visual cues than [back] for the tongue (Lisker & Rossi, 1992).
- The +/− values are dichotomic, implying that sounds with the same feature values share the same articulatory gestures.
  → (articulatorily) true? visually distinct?

Research Question

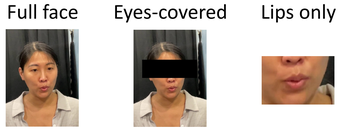
Can Taiwan Mandarin listeners reliably identify the three high vowels from natural speech when no acoustic signal was available?

Methods

- Participants: 60 (31 females) native Taiwan Mandarin speakers, aged 20-33 (mean = 21.7; SD = 3.1).
- Materials: 504 photos of natural speech /i, u, y/ (168 each) from 12 talkers (6 females).

- Forced-choice identification task /i, u, y/.
- Between-subject conditions:

Full face | Eyes-covered | Lips only



$$\text{Accuracy} = \frac{(hit)}{(hit+miss)}$$

d' = z(hit rate) - z(false alarm rate)

- Two-way ANOVA:
  - Accuracy ~ Vowel + Condition + Vowel:Condition
  - d-prime ~ Vowel + Condition + Vowel:Condition

Results

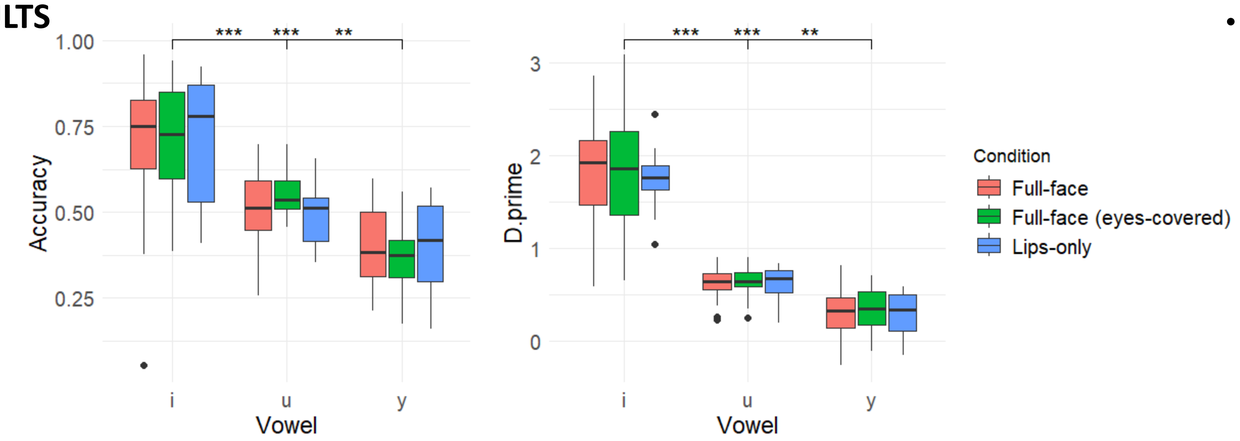


Figure 1. Accuracy (left) and d-prime (right) for high vowels /i, u, y/ across different conditions.

- Main effects of vowel on both accuracy [$F(2, 171) = 81.744$, $p < .001$] and d-prime [$F(2, 171) = 318.087$, $p < .001$], but no effect of condition nor vowel-condition interactions.

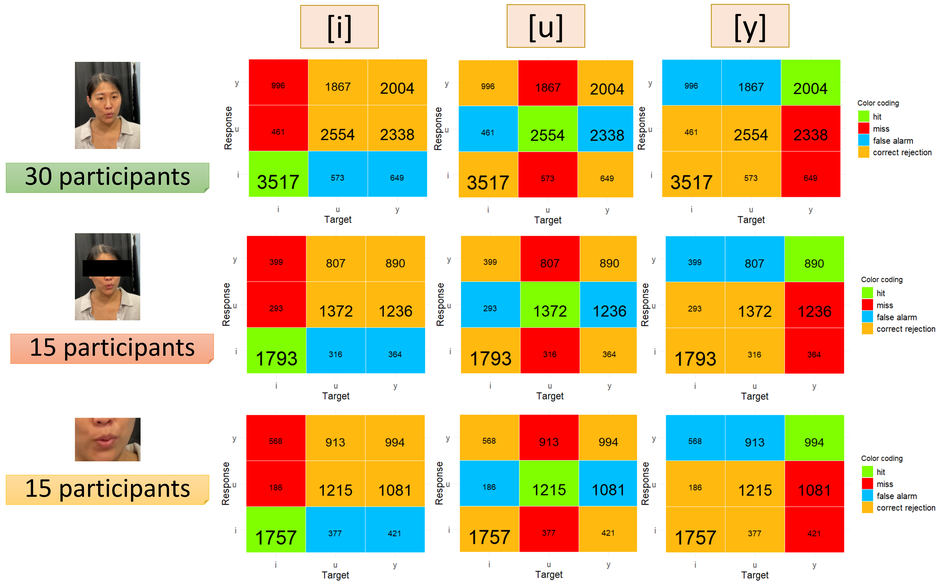


Figure 2. Confusion matrices for each target high vowels /i, u, y/ across different conditions.

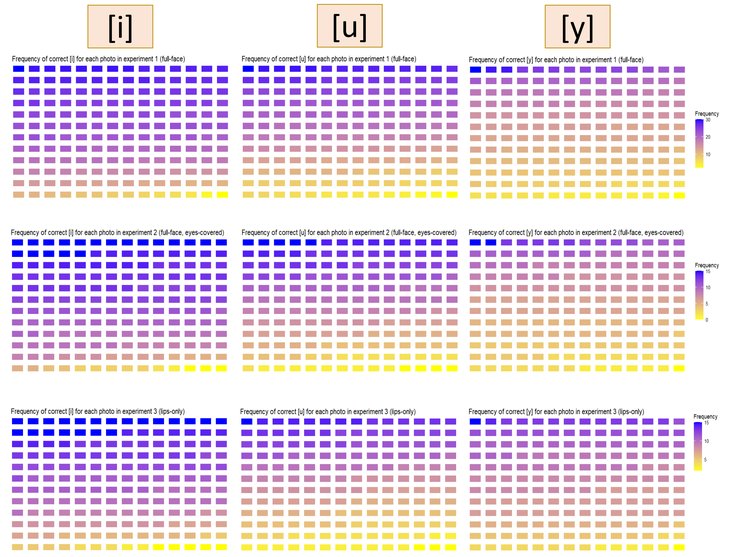


Figure 3. Heatmaps of the frequency of correct identification in each photo for each target high vowels /i, u, y/ across different conditions.

Discussion

- Taiwan Mandarin native listeners can reliably identify three high vowels that contrast in lip postures, with the highest accuracy for /i/, followed by /u/, and then /y/.
- Target /y/ was more likely to be mis-identified as /u/ than target /u/ being mis-identified as /y/ (miss cells in Figure 2). → *Labeling of /y/ and /u/ may not be a mirrored mapping.*
- Item-by-item analyses of the condition revealed that listeners could identify /y/ more accurately when provided with only lip information.
  → *Limited visual information force listeners to discern the subtle differences in lip postures.*

References

[1] Lisker, L., & Rossi, M. (1992). Auditory and visual cueing of the [±rounded] feature of vowels. *Language and speech, 35*(4), 391-417.
[2] Catford, J. C. (1988). *A Practical Introduction to Phonetics*. Oxford: Clarendon Press.
[3] Jackson, M. T.-T., McGowan, R. S. (2012). A study of high front vowels with articulatory data and acoustic simulations. *JASA, 131*, 3017–3035.
[4] Ladefoged, P., Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford, England: Blackwell.
[5] Lin, Y. H. (2007). *The Sounds of Chinese*. Cambridge University Press.

Acknowledgement

This work has been supported by the fund (MOST 111-2410-H-002-49-MY2) granted to the first author.

Lab website