

Bisensory association between sound and shape

Yuna Kwak¹, Hyun-Woong Kim¹, Hosung Nam², & Chai-Youn Kim¹

¹ Department of Psychology, Korea University ² Department of English Language and Literature, Korea University



Introduction

Previous studies have demonstrated the robustness of the Bouba-Kiki effect (also known as the Maluma-Takete effect); when asked to match a visual stimulus with a non-word, people consistently label a round shape 'bouba' and a spiky shape 'kiki' [1, 2]. However, debate still exists as to whether phonetic features or sound, and not orthographical and linguistic features, play a crucial role in this non-random association. Therefore, the present study aims to closely examine the influence of sound on this association with synthesized vowel and consonant sounds.

Procedures Expt 1 & 2

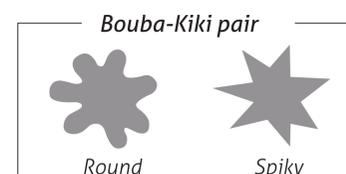
Participants

- 40 participants in each experiment
- 17 males, 23 females in Expt 1
- 15 males, 25 females in Expt 2

Task

- Two-alternative forced choice task: Participants listened to a synthesized sound on each trial and chose the visual shape (round or spiky) that matched the sound.

Visual Stimuli

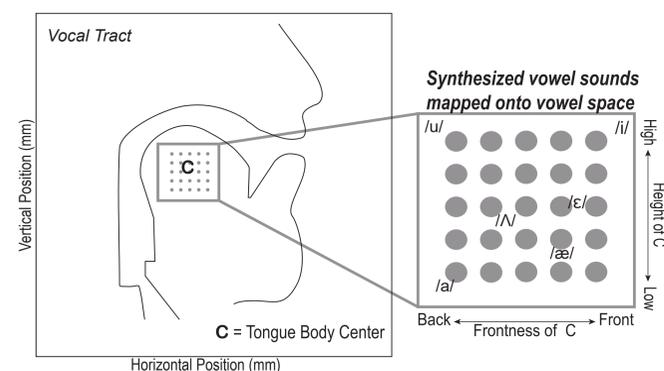


- In Expt 1, two versions of round-spiky pair (between-subjects condition) were used to examine whether difference in the overall shape of the pairs results in different patterns of matching behavior [3].
- In Expt 2, only one pair of visual shapes was used.

Experiment 1: Vowels

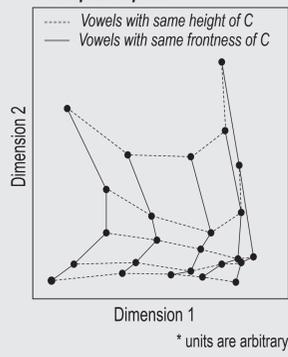
Auditory Stimuli

- 25 vowel sounds were synthesized using Haskins Laboratory Articulatory Synthesizer [4, 5].
- The sounds were not confined to any orthographical or linguistic system, and therefore the influence of sound could be closely examined.



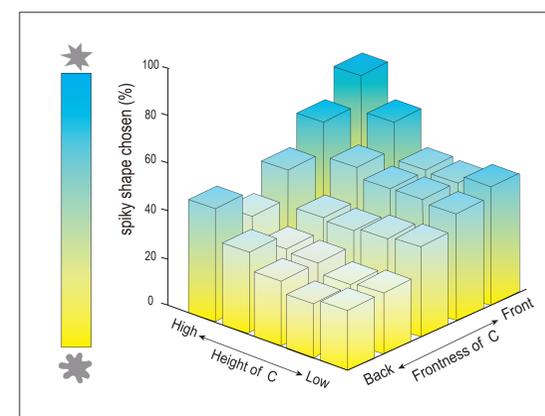
Dissimilarity Rating Results

Perceptual Space



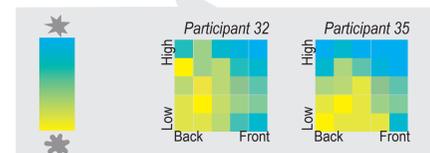
- **Task:** Participants rated the dissimilarity between the 25 synthesized vowel sounds on a scale from 1 to 7.
- **Non-metric Multidimensional Scaling Analysis:**
 - Analysis found the optimal dimensions for explaining the dissimilarity data.
 - Analysis confirmed that the physical space of the synthesized sounds conformed to the perceptual space of the participants, showing the effectiveness of the physical parameters manipulated [stress value = 0.078].

Results



- There was no significant difference between the two pair types of visual shapes, so the results were collapsed for further analysis.
- **Two-way Repeated Measures Anova**
 - significant main effects of frontness [$F(4, 156) = 20.866, p < .001$] and height [$F(4, 156) = 8.887, p < .001$] on the percentage of trials the spiky shape was chosen
 - significant interaction effect between frontness and height on the percentage of trials the spiky shape was chosen [$F(6, 234) = .034, p < .001$].

Individual Data



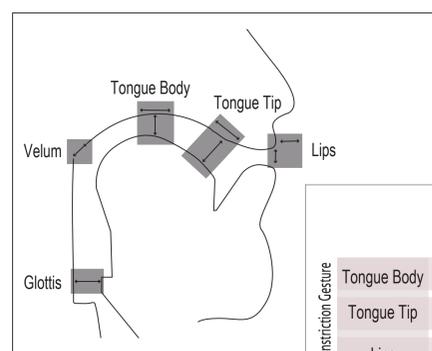
Summary

Non-arbitrary relationship exists between vowel sounds and shape; front and high vowels tend to be associated with the spiky shape more often.

Experiment 2: Consonants

Auditory Stimuli

- 12 vowel-consonant-vowel (vcv) sounds were generated using Haskins Laboratory Articulatory Synthesizer.
- Constriction gestures at 5 speech organs (**lips, tongue tip, tongue body, velum, glottis**) were parametrically manipulated with the vowel gesture fixed at its rest position [6].
- The sounds were not confined to any orthographical and linguistic system, and participants lacked awareness of the consonantal nature of sounds.



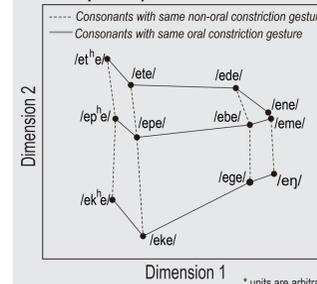
Stimulus Conditions

Oral Constriction Gesture	Non-oral Constriction Gesture			
	Velum	N/A	Glottis	Glottis VOT>0
Tongue Body	/en/	/ege/	/eke/	/ek ^h e/
Tongue Tip	/ene/	/ede/	/ete/	/et ^h e/
Lips	/eme/	/ebe/	/epe/	/ep ^h e/

* VOT = Voice Onset Time

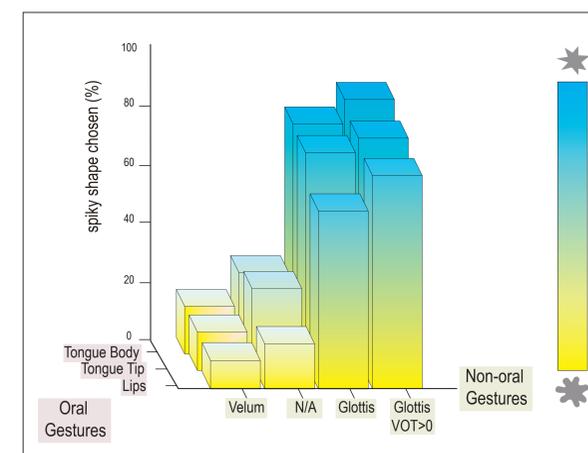
Dissimilarity Rating Results

Perceptual Space



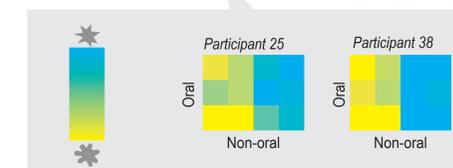
- **Task** was identical to that of the dissimilarity rating in Expt 1.
- **Non-metric Multidimensional Scaling Analysis:**
 - Consonants in the same non-oral/oral constriction gesture condition were grouped closely together along dimension 1 and 2 in perceptual space, showing that the physical space of the synthesized consonants conformed to the participants' perceptual space [stress value = 0.054].

Results



- **Two-way Repeated Measures Anova**
 - significant main effects of non-oral [$F(4, 156) = 20.866, p < .001$] and oral [$F(4, 156) = 8.887, p < .001$] constriction gestures on the percentage of trials the spiky shape was chosen
 - no interaction effect between non-oral and oral constriction gesture

Individual Data



Summary

Non-arbitrary relationship exists between consonant sounds and shape; consonants generated with glottal gestures are associated with the spiky shape more often.

Conclusion

Given that the synthesized vowels and consonants were not confined to any linguistic or orthographical system, we could control factors other than sound that might contribute to the Bouba-Kiki effect. Therefore, these results suggest a non-arbitrary relationship between sound and shape, and we conclude that sound indeed plays a vital role in the consistency observed in shape matching.

References & Acknowledgements

- [1] Köhler (1947). Gestalt Psychology.
 - [2] Ramachandran & Hubbard (2001). Journal of Consciousness Studies. 8(12), 3-34.
 - [3] D'Onofrio (2013). Language and Speech. 0(0), 1-27.
 - [4] Nam et al. (2013). Journal of Phonetics. 41(2), 63-77.
 - [5] Kim et al. (Accepted). Multisensory Research.
 - [6] Kim et al. (2016). VSS 2016.
- This work was supported by NRF-2016R1A2B4011267.