Crossmodal correspondence between haptic shape and sound

Yuna Kwak[1], Hosung Nam[2], & Chai-Youn Kim[1]

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

VCN Visual Cognitive Neuroscience @ KU

Introduction

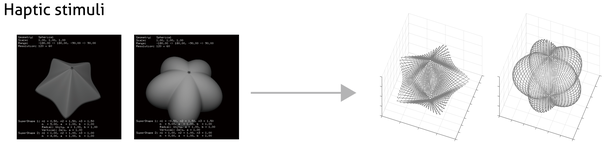
Crossmodal correspondence refers to the tendency of the brain to systematically map stimulus features across different senses [1]. For example, the 'bouba-kiki' effect is one of the most well-known examples demonstrating audiovisual correspondence: people consistently label a round/spiky visual shape 'bouba'/'kiki' [2,3]. Since previous literature is focused on the visual and auditory domain, here we sought to examine the correspondence between haptic shape and sound, by generating three-dimensional objects in which the roundness/spikiness dimension is parametrically manipulated.

Procedures (Expt 1 & Expt2)

- Participants

  Expt 1: 19 participants (5 males, ages: 19-26)
  Expt 2: 23 participants (7 males, ages: 20-28)

- Haptic stimuli



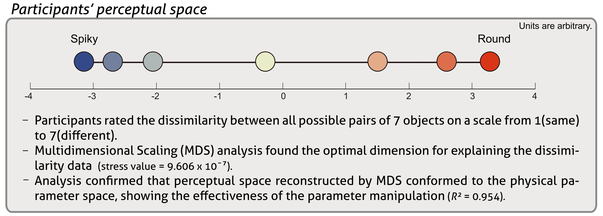
A parametric shape model 'Superformula' was used to generate the round and the spiky 3D shape [4].

X, Y, Z coordinates of 48400 points constituting the surface of each shape were extracted.

Linear interpolation on each corresponding pair of coordinates (in terms of angle formed between the center and the line projected from the center to each coordinate) was conducted to generate the same number of points for 5 intermediate shapes.

3D printed objects

Participants' perceptual space



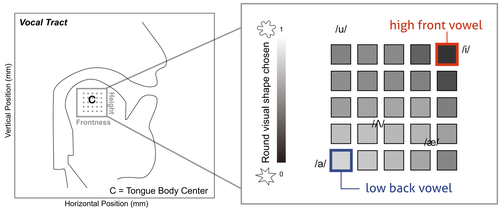
- Participants rated the dissimilarity between all possible pairs of 7 objects on a scale from 1(same) to 7(different).
- Multidimensional Scaling (MDS) analysis found the optimal dimension for explaining the dissimilarity data (stress value = 9.606×10^{-7}).
- Analysis confirmed that perceptual space reconstructed by MDS conformed to the physical parameter space, showing the effectiveness of the parameter manipulation ($R^2 = 0.954$).

- Task

  Participants, while closing their eyes, palpated a three-dimensional object with their right hand for 5 sec on each trial. After listening to two sounds presented sequentially, they made a forced choice to indicate the sound better associated with the object (5 repetitions for each object).

Experiment 1: Vowel sound

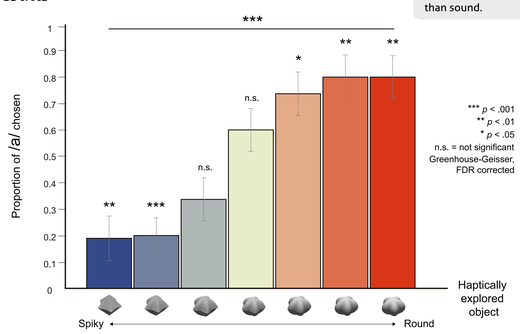
Auditory stimuli



- Twenty-five vowel sounds were synthesized using Haskins Laboratory Articulatory Synthesizer, by manipulating the horizontal and vertical position of the tongue body's center [5].
- Of the vowel sounds, the two that were associated with round and spiky visual shape respectively were presented on each trial [6] .

The vowels were not confined to any orthographical or linguistic system, thereby allowing us to control for factors other than sound.

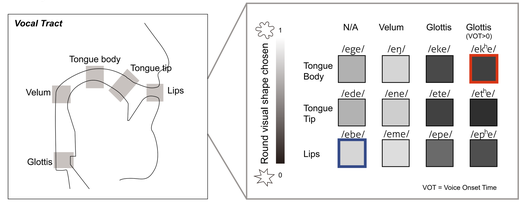
Results



- Repeated Measures ANOVA revealed the main effect of object ($F(1.554, 27.965) = 10.918$).
- The proportion of /a/ choice for each object condition was compared against chance using one sample t-tests (from spiky to round: $t(19) = 3.678$, $t(19) = 4.500$, $t(19) = 2.014$, $t(19) = 1.236$, $t(19) = 2.873$, $t(19) = 3.581$, $t(19) = 3.707$).

Experiment 2: Consonant sound

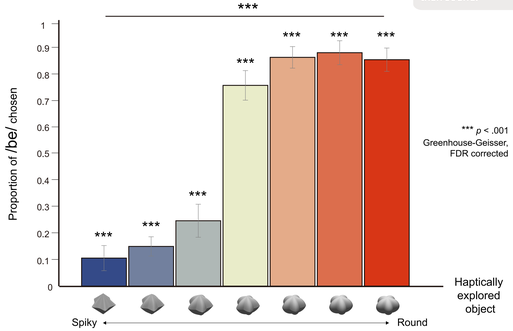
Auditory stimuli



- Twelve vowel-consonant-vowel (vcv) sounds were generated using Haskins Laboratory Articulatory Synthesizer [7].
- Of the consonant sounds, the two (cv) that were associated with round and spiky visual shape respectively were presented on each trial [6].

The consonants were not confined to any orthographical or linguistic system, thereby allowing us to control for factors other than sound.

Results



- Repeated Measures ANOVA revealed the main effect of object ($F(2.287, 50.318) = 55.199$).
- The proportion of /be/ choice for each object condition was compared against chance using one sample t-tests (from spiky to round: $t(22) = 8.449$, $t(22) = 9.771$, $t(22) = 4.168$, $t(22) = 4.496$, $t(22) = 8.762$, $t(22) = 8.277$, $t(22) = 7.900$).

Preliminary Brain Imaging

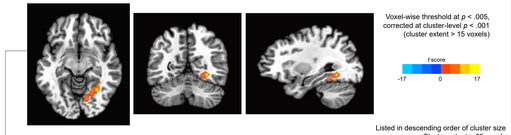
- Participants 8 participants (4 males, ages: 18-27)
- Design & Analysis
  - Participants, while closing their eyes, palpated the 7 objects with their right hand inside the scanner.
  - 20 repetitions per object within each run
  - Whole brain analysis was conducted with AFNI's amplitude modulation function (3dDeconvolve_AM2): mean regressor (the mean response), amplitude modulation regressor (variations in response with the user-supplied amplitude factor - i.e., haptic roundness/spikiness), head motion regressors.

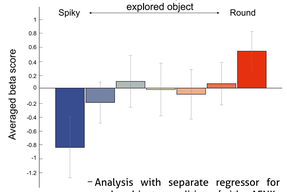
Scan parameters:
TR = 3000 msec, TE = 30 msec, slice thickness = 4.0 mm x 4.0 mm, in-plane resolution = 2.0 mm x 2.0 mm

Results

- Clusters with statistically significant beta score of the amplitude modulation regressor (brain regions in which activity is parametrically modulated by the roundness/spikiness dimension in the haptic modality)

Rt. Fusiform Gyrus (FG) [29 -56 -7 TAL]





- Analysis with separate regressor for each object condition (with AFNI's 3dDeconvolve), beta score for peak voxel averaged across participants

Listed in descending order of cluster size
Cluster extent > 25 voxels

Regions showing amplitude modulation

Rt. Fusiform Gyrus	[29 -56 -7 TAL]
Lt. Fusiform Gyrus	[-49 64 -19 TAL]
Rt. Parahippocampal Gyrus	[26 -13 -28 TAL]
Rt. Visual Association Area	(BA 18) [26 -85 11 TAL]
Rt. Visual Association Area	(BA 19) [35 -76 -10 TAL]
Lt. Anterior Prefrontal Cortex	[-22 53 -4 TAL]
Lt. Medial Parietal Cortex	[-4 52 53 TAL]

- Activity in FG is increased for round objects, compared to spiky objects.
- The results are in line with previous studies examining object-specific activation in haptics, audition, and vision, and curvature processing in vision [8,9].

Conclusion

The results of the current work extend sound-shape correspondence to the auditory and haptic modality: round objects are associated with sounds sounding similar to /a/ and /be/, while spiky object tends to be associated with sounds sounding similar to /i/ and /ke/, consistent with the 'bouba-kiki' effect. Furthermore, the stimuli used in the present study enable us to conclude that the results reflect the correspondence between sound and shape per se, and not any other factors.

References & Acknowledgements

This work was supported by NRF-2016R1A2B4011267.

[1] Spence (2011). *Attention, Perception, & Psychophysics* 73, 971-995.
[2] Ramachandran & Hubbard (2001). *Journal of Consciousness Studies* 8, 3-34.
[3] Köhler (1947). Gestalt Psychology.
[4] Lee Masson et al. (2016). *Cerebral Cortex* 26, 3402-3412.
[5] Nam et al. (2013). *Journal of Phonetics* 41, 63-77.
[6] Kwak et al. (2017). IMRF 2017.
[7] Kim et al. (2016). VSS 2016.
[8] Kassuba et al. (2011). *NeuroImage* 56, 1566-1577.
[9] Yue et al. (2014). *PNAS* 111, 3467-3475.