ACOUSTIC AND VISUAL PARAMETERS UNDERLYING WORD-SHAPE SOUND SYMBOLISM

Simon Lacey[1,2], Yaseen Jamal[1], Sara M. List[1], Kelly McCormick[1], K. Sathian[1,2] & Lynne C. Nygaard[1]

Milton S. Hershey Medical Center, Penn State College of Medicine, Hershey, PA, USA; Emory University, Atlanta, GA, USA

INTRODUCTION

Sound symbolism refers to a non-arbitrary mapping between the sound of a word and its meaning; but it is still unclear which acoustic parameters of a word contribute to meaning.

We used a novel application of representational similarity analysis (RSA) [1] to compare perceptual ratings of the roundedness/pointedness of a large group of pseudowords and shapes to measurements of their acoustic and visual parameters.

RSA computes the pairwise first-order correlation (Pearson's r) between each item and every other item for ratings or parameter measurements.

The correlation values are used to construct a representational dissimilarity matrix (RDM) for the ratings and for each parameter. Pairwise dissimilarity = 1-r and this is the value in each cell of the RDM.

RDMs of stimulus parameters and perceptual ratings can be compared via a second-order correlation (Spearman's r_s) to examine the contribution of a parameter to perception of the pseudowords as rounded or pointed.

METHODS

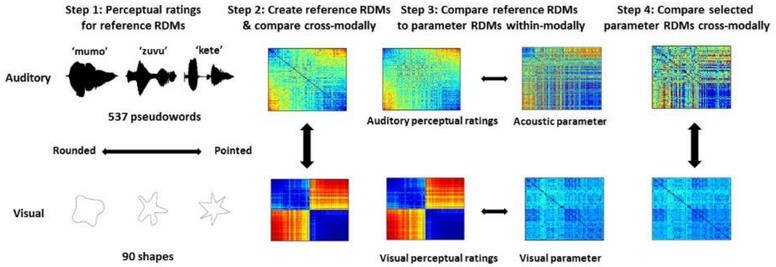


Fig 1. Schematic of the analysis pipeline.

We used ratings of the roundedness/pointedness of 537 CVCV pseudowords that systematically sampled variation in vowel quality, consonant voicing, and articulation [2].

A set of 90 visual shapes were separately rated for roundedness/pointedness [3].

Visual parameters of the shapes were calculated in MATLAB; acoustic parameters of the pseudowords were calculated in MATLAB and from the standard voice report in PRAAT [4]; RSA was implemented in MATLAB.

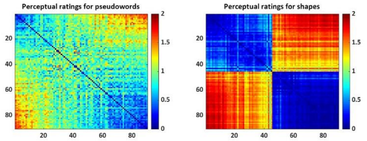


Fig 2. Ratings of the roundedness/pointedness of the pseudowords and shapes were correlated (r_s = .64, p < .001) and thus cross-modally consistent even though they were rated by independent groups.

RESULTS (**NB:** Figs 3 & 5-7: all tests pass Bonferroni correction; Figs 3,6,7 r = Spearman correlation coefficient).

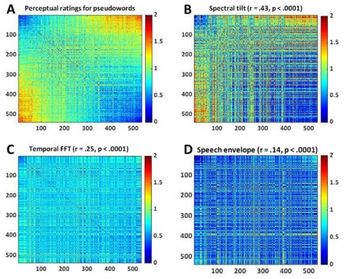


Fig 3. (A) Ratings of the roundedness/pointedness of pseudowords were correlated with:

(B) spectral tilt, which is steeper for rounded and flatter for pointed pseudowords respectively as power migrates from lower to higher frequencies (see Fig 4A);

(C) temporal FFT, which captures more abrupt changes in power for pointed compared to rounded pseudowords (see Fig 4B);

(D) speech envelope, which is continuous/smooth or discontinuous/uneven for rounded and pointed pseudowords respectively (see Fig 4C).

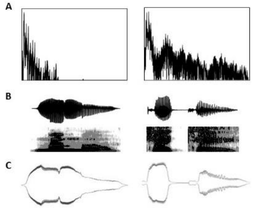


Fig 4. (A) spectral tilt, (B) temporal FFT, and (C) speech envelope for example rounded ('mumo' - left) and pointed ('kete' - right) pseudowords.

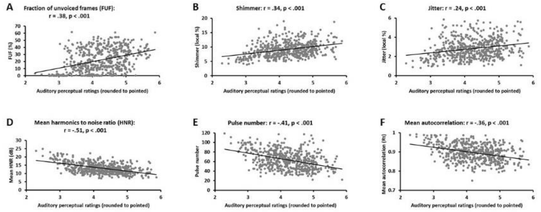


Fig 5. Roundedness/pointedness of pseudowords correlated positively with the fraction of unvoiced frames, shimmer, and jitter, and negatively with the mean harmonics to noise ratio, pulse number, and the mean autocorrelation; in all cases, this reflected pseudowords transitioning from rounded to pointed as vocal variability or roughness increased.

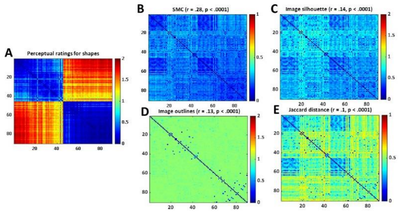


Fig 6. (A) Ratings of the roundedness/pointedness of the shapes were correlated with:

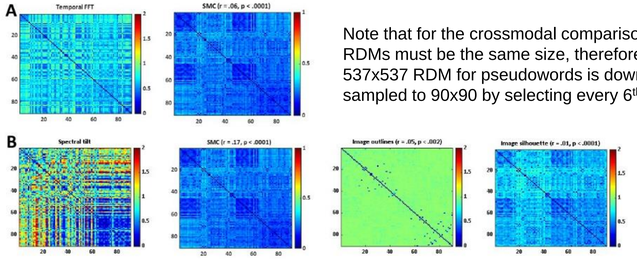
(B) the simple matching coefficient (SMC);

(C) image silhouette;

(D) image outlines;

(E) the Jaccard distance

RESULTS (cont.)



Note that for the crossmodal comparison, the RDMs must be the same size, therefore the 537x537 RDM for pseudowords is down-sampled to 90x90 by selecting every 6th item.

Fig 7. Spectral parameters of the auditory pseudowords (A) temporal FFT and (B) spectral tilt were significantly correlated with global indices of visual shape – SMC, image outlines, and image silhouette.

DISCUSSION

Ratings of the roundedness/pointedness pseudowords and shapes were closely related, reflecting the well-known sound-symbolic crossmodal correspondence between words and shapes [5,6].

Global acoustic parameters (spectral tilt, temporal FFT, and speech envelope) contributed to perception of pseudowords as rounded or pointed.

Voice quality measures involving a single measurement per item had to be binned into sets to calculate pairwise dissimilarity. These were unrelated to pseudoword ratings using RSA because of the loss of statistical power resulting from binning (data not shown). However, many voice quality measures were related to pseudoword ratings via conventional correlational analyses, with ratings transitioning from rounded to pointed with increasing vocal variability or roughness.

Crossmodally, spectral parameters of the pseudowords were related to global indices of visual shape.

This work establishes the utility of RSA for large stimulus sets comprising multiple measurements per item and offers new insights into the stimulus parameters underlying word-shape sound symbolism.

REFERENCES

1. Kriegeskorte, N. et al. (2008) Neuron, 60:1126-1141.
2. McCormick, K. et al. (2015) Proceedings 37th Annual Meeting Cognitive Science Society, 1565-1570.
3. McCormick, K. et al. (unpublished data)
4. Boersma, P. & Weenink, D. (2012) http://www.praat.org
5. Köhler, W. (1929) *Gestalt Psychology*. Liveright, NY.
6. Köhler, W. (1947) *Gestalt Psychology: An Introduction to New Concepts in Modern Psychology*. Liveright, NY.

FUNDING

This work was supported by grants to KS & LCN from the National Eye Institute at the NIH (R01EY025978) and Emory University Research Council. Support to KS from the Veterans Administration and to SML from the Laney Graduate School is also acknowledged.