## COLOR SOUND SYMBOLISM IN NATURAL LANGUAGES

TABLE	2	The eleve	n innest	igated	color	concepts

Color concept	WCS chip <sup>1</sup>	RGB	CIELAB	Luminance	Saturation <sup>2</sup>
RED	G1	186.9, 28.7, 71.3	41.2, 61.4, 17.9	41.2	64.0
GREEN	F17	0, 146.2, 69.9	51.6, -63.3, 29.0	51.6	69.6
YELLOW	C9	253.7, 193.4, 0	81.4, 7.3, 109.1	81.4	109.4
BLUE	F29	0, 129, 205	51.6, -3.4, -48.1	51.6	48.2
BLACK	J0	0, 0, 0	0, 0, 0	0	0
WHITE	A0	255, 255, 255	100, 0, 0	100	0
GRAY	_	119, 119, 119	50, 0, 0	50	0
NIGHT	_	59, 59, 59	25, 0, 0	25	_
DAY	_	185, 185, 185	75, 0, 0	75	_
DARK	_	59, 59, 59	25, 0, 0	25	_
LIGHT	_	185, 185, 185	75, 0, 0	75	_

Notes: <sup>1</sup> World Color Survey chip corresponding to focal color (Regier et al., 2005); <sup>2</sup> saturation was calculated as Euclidean distance from the achromatic central spindle of the CIELAB space:  $\sqrt{(a^2 + b^2)}$ .

## 3. Results

## 3.1. VOWELS

We discovered a significant association between the luminance of a color and the sonority of vowels in the word for this color (Figure 4A). The average sonority of vowels was predicted to be 0.4 points (95% CI [0.2, 0.6]) higher on a scale of 12 to 16 in words for WHITE (luminance = 100) than in words for BLACK (luminance = 0). Luminance also predicted the subjective brightness of vowels in a color's name (Figure 4B): the average brightness rating of vowels in words for WHITE was predicted to be 12% (95% CI [4, 20]) higher than in words for BLACK. The association between spectral centroid and luminance was not statistically significant, but it was in the predicted direction (83 Hz, 95% CI [-34, 199]). In addition, luminance predicted an increase in the frequency of F1, but not F2 or F3 (Figure 4D, 4E, 4F). It is worth pointing out that all these vowel characteristics tend to be positively correlated. For example, according to the acoustic analysis of IPA recordings in the pilot study, F1 correlates with both vowel sonority (r = .41) and vowel brightness (r = .47), and sonority is also positively associated with brightness (r = .37). In other words, the large picture is that there is a tendency for both bright and sonorous vowels (which are largely the same) to occur in the words for light colors, while dark and less sonorous vowels are more common in the words for darker colors.

In order to ascertain that the observed association between color luminance and the 'brightness' of phonemes in the corresponding words is not caused by another color characteristic, it would be desirable to perform multiple regression controlling for saturation and hue. Unfortunately, this is not possible with only a few color words. Looking at univariate effects of