

Cross-sensory correspondences and cross talk between dimensions of connotative meaning: Visual angularity is hard, high-pitched, and bright

Peter Walker

Published online: 26 July 2012
© Psychonomic Society, Inc. 2012

Abstract Higher-pitched sounds are judged to be, among other things, sharper, harder, and brighter than lower-pitched sounds. Following Karwoski, Odbert, and Osgood (Journal of General Psychology 26:199–222, 1942), such cross-sensory correspondences are proposed to have a semantic basis, reflecting extensive bidirectional cross-activation among dimensions of connotative meaning. On this basis, the same core set of correspondences should emerge whichever sensory feature is used to probe it. More angular (sharper) shapes should, for example, be higher-pitched and have the same cross-sensory features as higher-pitched sounds. Experiments 1–3 employed a speeded classification task designed to reveal cross-sensory correspondences having a semantic basis. With words as to-be-classified stimuli and with shapes varying in angularity as concurrent incidental stimuli, congruity effects between angularity and each of hardness, pitch, and brightness were confirmed. Correspondences with a semantic basis need not be cross-modality in nature. Experiment 4 confirmed this by reproducing the brightness–angularity congruity effect when contrasting values for both features were encoded nonverbally within the visual modality. The varying nature and origins of cross-sensory correspondences and the basis on which they induce congruity effects in speeded classification are explored.

Keywords Multisensory processing · Cross-sensory Correspondences · Connotative meaning · Congruity effects · Speeded classification · Visual angularity

Different sensory channels can provide *corresponding* information (what Marks, 1978, also referred to as analogous information) about an object, as when vision and audition both confirm its connotative brightness, the former by registering its surface

brightness, the latter by registering that it makes relatively bright (i.e., high-pitched) sounds (*NB*: The connotative meaning of any kind of stimulus is what it suggests, implies, or invokes, rather than what it explicitly or directly denotes). The provision of corresponding connotative information through different sensory channels is the focus of the present study.¹

Early evidence for cross-sensory correspondences comes from studies of the universal connotative meanings of elementary stimulus features (e.g., Karwoski, Odbert, & Osgood, 1942; Lundholm, 1921; Osgood, 1960; Poffenberger & Barrows, 1924; Scheerer & Lyons, 1957). For example, when people draw the visual imagery they experience when listening to short musical selections, they draw lines and forms that tend to be thinner, brighter, smaller, and more angular (sharper) the higher in pitch and/or faster in tempo the music is (Karwoski et al., 1942). And when presented with simple sounds, they judge higher-pitched sounds to be brighter,² faster, harder, higher in space, lighter in weight, sharper, and smaller than lower-pitched sounds (Collier & Hubbard, 2001; Eitan & Timmers, 2010; Hubbard, 1996; Marks, 1974, 1975, 1978; Mondloch & Maurer, 2004; Perrott, Musicant, & Schwethelm, 1980; Walker & Smith, 1984).³

¹ Sensory channels are rather loosely conceived to be those parts of the sensory systems responsible for encoding dissociable stimulus features, such as color, texture, and movement in vision. The term *sensory channel* is preferred over *sensory modality* because correspondences exist among feature dimensions within the same modality (e.g., brightness and angularity/pointedness in vision). For this reason, the slightly broader term *cross-sensory correspondences* is preferred over *cross-modality correspondences* (e.g., Spence, 2011).

² Pitch is not the sole determinant of the brightness of a simple sound. Loudness contributes also (see Marks, 1978). Pitch sometimes combines with other features to influence the brightness of a sound through its timbre. That said, pitch remains a powerful determinant of the brightness of a sound.

³ Here and elsewhere in the present article, the features listed as being possessed by a stimulus are not exhaustive. The connotative features examined in previous studies have always been limited in number and have differed from one study to the next.

P. Walker (✉)
Department of Psychology, Lancaster University,
Lancaster LA1 4YF, UK
e-mail: p.walker@lancaster.ac.uk

Speeded classification tasks have proven especially useful for exploring cross-sensory correspondences. This is because people classify elementary stimulus features more quickly and accurately when task-irrelevant stimuli in other sensory channels have corresponding (congruent), rather than noncorresponding (incongruent), cross-sensory features (for reviews, see Marks, 2004; Spence, 2011). For example, people classifying simple visual stimuli according to their brightness or angularity respond more easily when a concurrent auditory tone has corresponding pitch (i.e., high for bright and angular, low for dark and curved) than when it does not have corresponding pitch (i.e., low for bright and angular, high for dark and curved) (Marks, 1987). Because pitch is incidental to the requirements of the task in this situation, registration of its cross-sensory connotations of brightness and angularity is thought to be automatic (see, e.g., Walker & Smith, 1984).

As part of their commitment to the notion that elementary stimulus features have connotative meanings, Karwoski et al. (1942) proposed an early precursor of more recent claims that cross-sensory correspondences can reflect the semantic coding of such features (e.g., Martino & Marks, 1999; Walker & Smith, 1984). Karwoski et al. suggested that cross-sensory correspondences arise from extensive cross-activation between corresponding places on dimensions of connotative meaning that are aligned with each other. Figure 1 illustrates how several dimensions appear to be aligned when people judge the cross-sensory features possessed by sounds contrasting in auditory pitch (see above). The figure also highlights the three correspondences involving sharpness/angularity that are the focus of the present study.

If interactions among dimensions of connotative meaning are bidirectional (cf. Martino & Marks, 2001), as well as extensive, then it follows from Karwoski et al.'s (1942) proposal that the same dimensional alignments and the same correspondences should be revealed whichever sensory feature is used to probe the correspondences. For example, if higher-pitched sounds are perceived to be sharper than lower-pitched sounds, sharper (more angular) shapes should be judged to have the same cross-sensory connotations as higher-pitched sounds. Thus, in addition to being judged to be higher in pitch, sharper shapes should also be brighter, faster, harder, higher in space, lighter in weight, and smaller than smoother (more curved) shapes. Furthermore, in light of the evidence from speeded-classification tasks (e.g., Marks, 2004), these cross-sensory features should be encoded even when they are an incidental feature of the task situation. The present study was designed to test these predictions.

There is already evidence identifying some of the cross-sensory features, other than higher auditory pitch, associated with visual angularity. For example, when people draw simple lines to represent verbally specified concepts, they draw

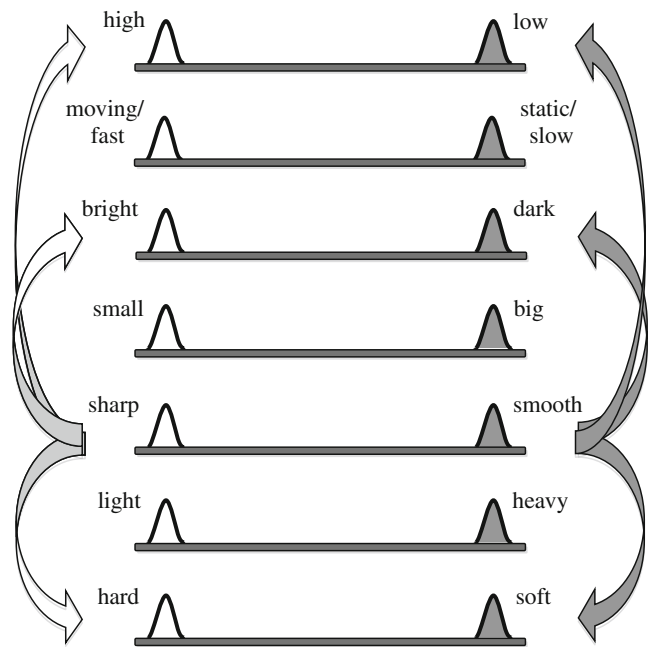


Fig. 1 How the correspondences observed when people make judgments about the cross-sensory features possessed by sounds contrasting in pitch could arise from the alignment, *en bloc*, of several dimensions of connotative meaning (based on Karwoski, Odbert, & Osgood, 1942). Those instances of cross-activation emanating from sharpness (angularity) examined in the present study are highlighted (with the arrows of influence). Other possible related dimensions not investigated in Experiments 1, 2, and 3 are included for completeness.

curved lines to represent the concepts of *delicate*, *quiet*, and *weak*, and angular lines to represent *hard*, *loud*, and *powerful* (Lundholm, 1921; Poffenberger & Barrows, 1924; Scheerer & Lyons, 1957). When they place verbally specified concepts on a scale defined by two shapes contrasting in angularity, they choose the more angular shape to represent the concepts of being *more active*, *faster*, *higher in space* (i.e., *up* rather than *down*), *lighter in weight*, *louder*, and *stronger* (Osgood, 1960). And when they place verbally specified concepts on a scale defined by the contrast between a white circle (or chip) and a black circle (or chip), not only do they choose brighter (white) to be *sharper* than black, they also choose it to represent the concepts of being *faster*, *harder*, *higher* (i.e., *up* rather than *down*), *lighter in weight*, *quieter*, *thinner*, and *weaker* (Osgood, 1960). Finally, Elliott and Tannenbaum (1963) confirmed that visual angularity has connotations of *hardness*. They asked participants to say what qualities were possessed by each of a varied set of randomly generated shapes. Factor analysis revealed that the type of contrast illustrated in Fig. 2 was captured best by the concepts of *angularity* and *hardness*, with more angular shapes being judged to be harder. Notwithstanding the somewhat inconsistent associations between surface lightness and strength/



Fig. 2 Two of the forms Elliott and Tannenbaum (1963) found contrasted most heavily in both angularity and hardness, with more angular forms also being harder

intensity (i.e., loud–quiet and strong–weak), which might reflect influence from the lightness contrast between the color samples and the background against which they appeared, these findings are in line with the prediction that sharper shapes will share their cross-sensory features (connotations) with higher-pitched sounds. Encouraged by these findings, the present study exploits the congruity effects observed in speeded classification to confirm that visual angularity enters into the same cross-sensory correspondences as higher-pitched sounds and that these correspondences have a semantic basis.

In a recent tutorial review, Spence (2011) proposed three types of correspondence, none of which is regarded as having a semantic basis. First, among dimensions concerned with the magnitude or strength of stimuli,⁴ correspondences are thought to reflect the presence of a common basis for coding stimulus magnitude (e.g., the rate of neural firing). Spence refers to this type of correspondence as being *structural* in nature. Second, Spence proposes that other correspondences reflect the perceptual learning of natural co-occurrences among features, such as the co-occurrence between auditory pitch and size. Spence refers to this type of correspondence as having a *statistical* basis. Finally, some correspondences are proposed to arise because the same words are commonly used to mark contrasting values on different dimensions of sensory experience (e.g., the words *high* and *low* mark contrasting levels of pitch, spatial frequency, and spatial elevation). Spence regards these

correspondences as having a *linguistic* basis, although *lexical* is perhaps a more appropriate term.⁵

Spence (2011) also proposed that there is a link between the basis for a nonsemantic correspondence and the level(s) of information processing at which the correspondence takes effect. He suggests, for example, that congruity effects based on structural and statistical correspondences arise at relatively low (early) levels of information processing, allowing the correspondences to have a direct impact on perceptual encoding. In contrast, congruity effects based on linguistic correspondences necessarily reflect interactions at higher (later) levels of processing. This would be the case when participants in a speeded classification task monitor their lexical representations to inform their stimulus classification and response selection (i.e., decisional processes) (e.g., the lexical representations of *high* and *low* during a spatial elevation classification task; see Melara & Marks, 1990).

These nonsemantic bases for cross-sensory correspondences need to be taken into account (i.e., precluded) when attempting to confirm that a particular correspondence has a semantic basis. In rehearsing how this might be achieved, it is useful to assume that two conditions need to be in place for a particular correspondence to induce a congruity effect in speeded classification. First, the classification decision needs to be based on the same type of representation as the correspondence. Second, encoding of the criterial and incidental stimulus features needs to converge on this type of representation.

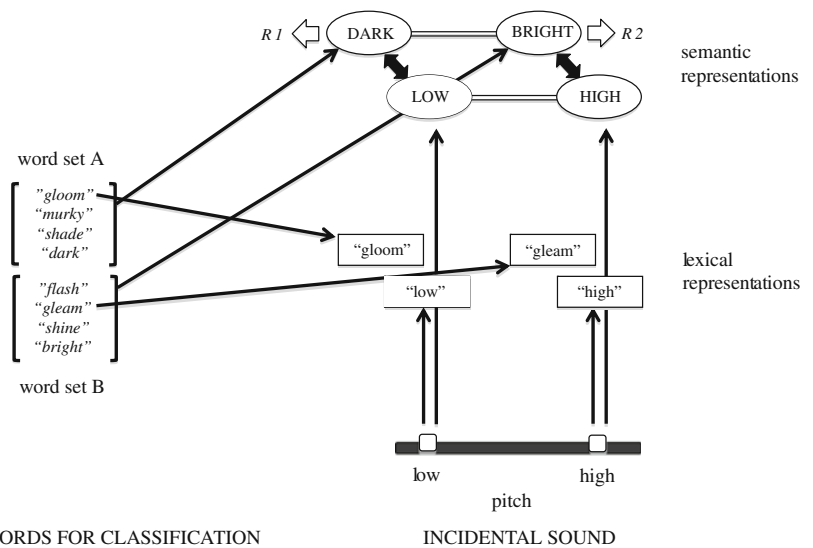
With these conditions in mind, Fig. 3 illustrates the processes and types of representation hypothesized to be involved in a version of the speeded classification task, devised by Walker and Smith (1984, 1985, 1986), for which only a semantic correspondence of the kind proposed by Karwowski et al. (1942) can induce a congruity effect. Taking brightness classification for the purpose of illustration, individual test words are classified according to whether they refer to something bright or to something dark. They are selected from two sets of words, established for the purpose of the experiment, with one set referring to bright things and the other set referring to dark things. The presentation of each test word is accompanied by an incidental sound whose pitch is randomly determined to be high or low. On the basis of the known correspondence between auditory pitch and

⁴ The distinction between magnitude dimensions and other types of dimension of sensory experience (what Stevens, 1957, called *prothetic* and *metathetic* continua, respectively) is, in essence, a distinction between quantity and quality. According to Stevens (1957, 1975), prothetic dimensions, such as loudness, are concerned with how much of a type of sensation is being experienced and are linked to variations in the level of excitation of a sensory channel. Metathetic dimensions are concerned with qualitative differences in sensory experience and are linked to variations in the location of maximum excitation in a sensory channel. Auditory pitch is a metathetic dimension, as indicated by the fact that we do not talk about sounds having *more* or *less* pitch (i.e., it is not a prothetic dimension).

⁵ *Linguistic* is too broad a term in this context. The term *lexical* refers to those features of a word (typically its unique orthography or phonology) that are used to find the word's entry in a dictionary, without reference to any of its meanings. Only with this restricted view of a word can it be said, for example, that *lightness* in weight and *lightness* in color share the same word (i.e., *light*). Once word meaning is brought into play, *light* in weight and *light* in color are different words. This distinction between lexical and semantic means that in the *lexical decision* task, people do not need to refer to a word's meaning in order to confirm its status as a word. Indeed, in some cases of semantic dementia, a word's status as a word can be confirmed even when its meaning is no longer available (e.g., Ward, Stott, & Parkin, 2000).

Fig. 3 A framework in which a congruity effect will reflect a semantic correspondence involving interactions among dimensions of connotative meaning (see the text for explanation)

CROSS-SENSORY CORRESPONDENCES



visual brightness, a congruity effect is predicted in which bright (dark) words are responded to more quickly and accurately when accompanied by a high-pitched (low-pitched) sound.

According to Spence (2011), when contrasting values for an elementary stimulus feature are communicated verbally, structural and statistical correspondences are precluded from contributing a congruity effect. This is because these types of correspondence are argued to reflect interactions contained in early levels of sensory-perceptual encoding. Lexically based correspondences remain a possibility, however, but only if the words being classified are also commonly used labels for the contrasting levels of the incidental stimulus feature. For example, Melara and Marks (1990) asked participants to respond differentially to the words *hi* and *lo*, and arranged for each test word to be accompanied by a high-pitched or low-pitched incidental tone. If participants were to monitor their internal lexicon to decide how to respond and if incidental tones automatically access the lexical representations for their common labels, a congruity effect having a lexical basis would be expected.⁶ On congruent trials, the test word and tone would converge on the same lexical representation, leaving no uncertainty about how to respond. On incongruent trials, however, the word and tone would access opposed lexical representations, creating uncertainty about how to respond. Resolving this uncertainty would require additional processing to ascertain which lexical representation was being accessed by the test

word and which was being accessed by the tone. Melara and Marks observed just such a congruity effect.

Because the test words included in Fig. 3 are not commonly used verbal labels for contrasting values of the incidental stimulus feature (i.e., for pitch), a correspondence based on lexical overlap should be minimized, if not precluded. This was the situation in a study reported by Martino and Marks (1999) in which a congruity effect was observed when participants classified the words *night* and *day* according to the level of brightness with which they are associated. Each test word was accompanied by a high- or low-pitched incidental tone. Because the words *day* and *night* are not verbal labels for contrasting levels of pitch, their lexical representations would not be accessed automatically by the incidental tone. The congruity effect could not, therefore, be reflecting a correspondence with a lexical basis. Likewise, because the brightness-related words indicated in Fig. 3 do not serve as common labels for contrasting levels of pitch, a lexical basis for a congruity effect is ruled out here also.

There is an aspect of the framework illustrated in Fig. 3 that makes it unlikely that participants will base their response selection on lexical representations. Because each test word is selected from one of two contrasting sets of words, participants are discouraged from basing response selection on the lexical identity of each individual test word. Thus, with two sets of four words, participants can choose to try and keep in mind how each of eight individual words is assigned to one of the two response alternatives, or they can choose simply to remember how the binary contrast in one aspect of their meaning is so assigned. Figure 3 assumes they opt to do the latter and base their response selection (i.e., $R1$ or $R2$) on semantic representations of brightness. Although contrasting levels of auditory pitch will have most direct access to semantic representations of height (and not

⁶ Complicating matters is evidence, referred to in the General Discussion section, that nonverbal stimuli (e.g., pictures of objects) do not activate verbal representations directly, but only indirectly via semantic representations.

brightness), the proposed interconnections among dimensions of connotative meaning will ensure that they converge on the same semantic representations of brightness accessed by the test words. A congruity effect is expected on this basis. On congruent trials, the semantic coding of both stimuli converges on the same value for brightness (i.e., *dark* or *bright*), leaving no uncertainty about how to respond. On incongruent trials, the semantic coding of the stimuli engages opposed values for brightness (i.e., *dark* and *bright*), and resolving the resulting uncertainty about how to respond requires additional processing to check which brightness pole is linked to the test word and which to the incidental sound.⁷

Creating the type of situation illustrated in Fig. 3, with which Walker and Smith (1984) confirmed correspondences between auditory pitch and the semantic coding of size, speed, and spatial elevation, was the strategy adopted in the present study to isolate semantically based congruity effects from the nonsemantic alternatives outlined by Spence (2011), but with contrasting levels of shape angularity serving as the incidental stimulus feature. Assuming that shapes with contrasting levels of angularity have their connotations of sharpness encoded, then any other dimensions of connotative meaning aligned with connotative sharpness will be revealed through their induction of a congruity effect. With this in mind, the connotative dimensions of *hardness*, *elevation (pitch)*, and *brightness* were selected as test cases, and experimenter-defined sets of test words whose meanings were associated with contrasting values on these dimensions served as to-be-classified stimuli in different experiments. To minimize the involvement of lexical correspondences, none of the words served as labels marking contrasting levels of visual angularity (e.g., *pointy*, *sharp*, *curved*, *blobby*, *rounded*).

The level of connotative agreement (congruity) between the angularity of the shape and the relevant semantic feature value associated with the test word it accompanied varied across different trials. On some trials, these two aspects were congruent (e.g., angular/*hard*, curved/*soft*, angular/*bright*, curved/*dark*), whereas on other trials, they were incongruent (e.g., angular/*soft*, curved/*hard*, angular/*dark*, curved/*bright*). It was anticipated that a significant effect of connotative congruity on the speed of classification would confirm that, even as task-irrelevant features, angular shapes, relative to curved shapes, are encoded as being connotatively hard, high (pitched), and bright, thus confirming the cross talk between these dimensions and the dimension of connotative sharpness.

Westbury (2005) has demonstrated that the angularity of a task-irrelevant shape can influence the ease with which

people classify letter strings as being words or nonwords (i.e., it can impact on lexical decision). He did this by arranging for each letter string to appear inside either an angular or a curved outline shape and then observing an effect of the congruity between the angularity of the shape and the phonological “angularity” of the nonword it framed (i.e., whether the consonants were *plosive* or *continuant* in nature).⁸ There was no such effect of the angularity of the surrounding shape when it was a word whose status was being confirmed. The success of Westbury’s strategy prompted the adoption of the same in the first three experiments reported here. The lexical decision task was replaced by a word classification task in which words appeared within an angular or curved outline shape and were responded to on the basis of a specified aspect of their meaning.

In Experiments 1–3 of the present study, the criterial stimuli were words whose meanings related to contrasting levels of hardness (e.g., *granite*), pitch (e.g., *squeak*), and brightness (e.g., *dark*), respectively. The test word on a trial was selected at random from two sets of words contrasting in the level of hardness (e.g., *granite* vs. *fur*), the level of pitch (e.g., *squeak* vs. *drone*), or the level of brightness (e.g., *glister* vs. *gloom*) to which one aspect of their meaning referred. The angularity of the shape within which each word appeared was selected independently of the word, the only restriction being that every word appeared equally often in a connotatively congruent shape (e.g., *rock*/angular, *drone*/curved, *gloom*/curved) and a connotatively incongruent shape (e.g., *dough*/angular, *squeak*/curved, *dark*/angular). Participants were required to classify each word as quickly as possible according to whether it referred to something suggesting hardness or to something suggesting softness (when hardness was the criterial feature—i.e., Experiment 1), to something suggesting a high-pitched sound, or to something suggesting a low-pitched sound (when pitch was the criterial feature—i.e., Experiment 2), or to something suggesting brightness or to something suggesting darkness (when brightness was the criterial feature—i.e., Experiment 3). For reasons to be explained later, Experiment 4 adopted a slightly different approach by adopting a brightness classification task in which contrasting levels of brightness were presented visually and nonverbally. Specifically, the angular and curved shapes were themselves filled in with one of four different levels of achromatic color varying from black to white.

⁸ Producing stop consonants requires the airflow in the vocal tract to be interrupted (stopped). Plosives are a type of stop consonant in which this interruption is complete. A continuant is any sound that is not a stop consonant—that is, a sound produced without interrupting the airflow in the vocal tract. Using symbols from the International Phonetic Alphabet, examples of nonwords containing only plosive consonants are /kɪb/, /gaet/, /tɛg/, /dɒp/, and /pɛd/, and examples of nonwords containing only continuant consonants are /mɪŋ/, /nɪm/, /laen/, /ɪɛŋ/, and /wɒl/ (taken from Monaghan, Mattock, & Walker, *in press*).

⁷ There is also the possibility, of course, that inhibitory connections across opposite ends of the dimensions exist and are able to contribute to reduced levels of performance in the presence of incongruity.

Stimulus development

Preliminary work was required to create and validate the shapes to be used in the experiments, to select and assess the words to be used in the classification tasks of Experiments 1–3, and to select a neutral typeface in which words would appear. An account of this preliminary work is reported here, before the individual experiments are introduced. Whereas details of the shapes to be used in all four experiments also are provided here, information about the words selected for use as to-be-classified stimuli (i.e., their identity and their visual and phonological angularity) is provided in the account of the individual experiment in which they are used. The 59 undergraduate students providing judgments about the stimuli under consideration ($n = 12$ in relation to the shapes to be used, 15 in relation to the typeface in which the words would appear, 13 in relation to the words to be classified, and 19 in relation to the visual angularity of the words) did not contribute to the evaluation of more than one type of stimulus and did not participate in any of the speeded-classification experiments. None received payment or any other form of credit for participating.

Angular and curved shapes

Two sets of outline shapes contrasting in angularity, but not differing in other respects (e.g., perceived size) that might link directly to other connotative meanings, were based on a quasirandomly selected set of points that were connected either by straight lines or by curves, using standard options available in Microsoft Word. The lines were drawn in black on a white background, and the resulting shapes were able to enclose each of the test words to be used in Experiments 1–3. The three angular and three curved shapes finally selected for use are shown in Fig. 4. Although all the shapes were initially created to fit snugly within a virtual 8.4×5 cm (landscape) frame, they were rescaled slightly to ensure that they were not perceived to differ in size.⁹

Twelve students rated each of the six shapes on five 7-point scales anchored at their ends by antonyms referring to

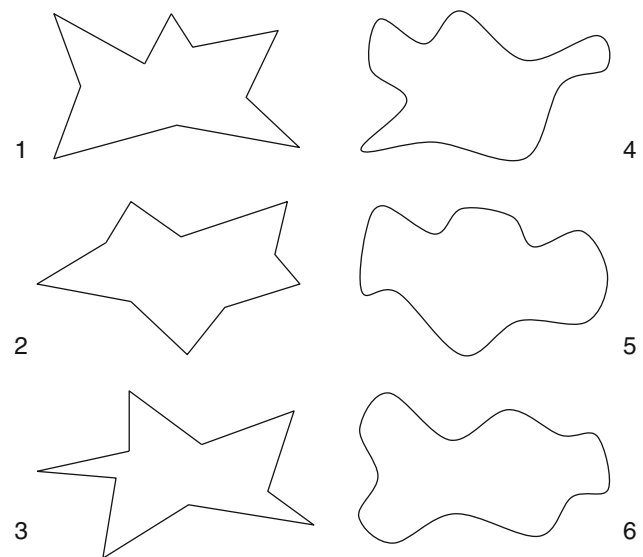


Fig. 4 The six shapes used in the experiments

different dimensions of connotative meaning. Two of these pairs (*big–small*, *simple–complex*) referred to feature dimensions along which it was intended the angular and curved shapes would match. The other three pairs (*soft–hard*, *high–low* (pitch), *dark–bright*) referred to the connotative meanings predicted to be associated with visual angularity (i.e., with the angular shapes being judged to be harder, higher in pitch, and brighter than the curved shapes). The scales on which the shapes were rated were represented visually with a horizontal line along which there were seven equally spaced vertical marks. Positioned immediately above successively less extreme marks were the verbal labels *extremely*, *very*, and *slightly*. The neutral midpoint position was labeled *neither*. Participants were asked to circle the mark that best represented a shape’s position on the scale. An integer score of 1–7 was recorded, with 1 and 7 being assigned to the two *extremely* ratings (i.e., 1 for *extremely soft*, *low*, *dark*, *big*, and *simple*; 7 for *extremely hard*, *high*, *bright*, *small*, and *complex*).

The shapes were presented for rating in a different random order for each participant and, on each occasion, were equally likely to be flipped vertically, or not, and horizontally, or not.¹⁰ The order in which the scales were presented

⁹ The procedure by which the shapes were rescaled, so that the angular and curved versions were matched for perceived size, involved pairing each angular shape with the most similar of the curved shapes. These pairs were then presented, side by side, on separate Powerpoint slides, and 10 students used the rescaling function to adjust the size of the curved shape until it appeared to match the size of the angular shape with which it was paired. There was very little variation in the final settings from different participants, and no curved shape needed to be altered by more than 3 % its original size to achieve a match. The most frequently selected size for each curved shape was adopted for use in the study.

¹⁰ During this preliminary work, and also in the speeded classification tasks reported below, the orientation of a shape was determined randomly for each presentation. Averaged across trials, stimuli, and participants, therefore, the results should be independent of shape orientation. The reason for randomizing selection of shape orientation for each presentation was to guard against the possibility that shape orientation would not be neutral. In particular, asymmetries in the shapes could mean that they appear to point in different directions depending on their orientation (e.g., pointing up has connotations of high rather than low). Random determination of shape orientation is one way of dealing with this potential problem.

Table 1 The mean rating (with *SD* in parentheses) and associated single-sample *t* value for each shape (numbered as in Fig. 4) on each of the six scales

scale	Angular shapes			Curved shapes		
	1	2	3	4	5	6
Soft-hard	5.75 (0.62)	4.83 (0.39)	6.00 (0.60)	2.50 (0.90)	2.67 (1.07)	2.17 (0.72)
	9.75***	7.42***	11.49***	-5.75***	-4.30***	-8.85***
Low-high pitch	5.42 (1.58)	4.92 (0.79)	6.25 (0.62)	2.75 (0.62)	2.15 (0.51)	2.50 (0.90)
	3.56**	4.00***	12.54***	-6.97***	-10.65***	-5.75***
Dark-bright	4.92 (1.51)	4.75 (1.00)	6.08 (0.51)	3.50 (1.31)	3.17 (1.64)	3.50 (1.73)
	2.11*	2.46*	14.02***	-1.32	-1.76	-1.00
Big-small	3.75 (1.05)	3.92 (0.67)	4.50 (0.90)	3.83 (0.72)	3.75 (1.06)	3.92 (1.16)
	-0.82	-0.43	1.92	-0.80	-0.82	-0.25
Simple-complex	4.41 (1.16)	4.42 (0.90)	4.58 (1.38)	3.92 (1.00)	4.25 (1.22)	4.25 (1.29)
	1.24	1.60	1.47	-0.29	0.71	0.67

Note. Mean ratings less (more) than 4.0 indicate that the first (second) mentioned antonym is most appropriate. Mean ratings significantly different from the neutral value of 4.0 are indicated

* $p < .05$

** $p < .01$

*** $p < .001$

and the left–right ordering of each antonym pair were both randomized. Before rating any of the shapes, participants familiarized themselves with all the shapes to be rated. Regarding the *high–low* (pitch) scale, participants were asked, “Imagine this shape came to life and made a sound, what would be the pitch of this sound?”

Table 1 presents the mean rating for each shape on each scale, together with the *t* value and associated significance level for the corresponding single-sample test (all *dfs* = 11).¹¹ The *t* tests were designed to confirm which deviations from the neutral value of 4.0 (i.e., the *neither* point on each scale) were significant. None of the shapes, angular or curved, was judged to deviate significantly from *neither* on the size and complexity scales. All of the angular shapes were judged to be significantly *hard*, *high-pitched*, and *bright*. All of the curved shapes were judged to be significantly *soft* and *low-pitched*. Although each of them was placed on the dark side of *neither* on the brightness scale, none of the deviations from *neither* was significant. When all the ratings on each scale were submitted to analysis of variance (ANOVA),¹² with the contrast between the set of angular shapes and the set

of curved shapes entered as a within-participants factor, the two different types of shape differed significantly in their *hardness*, $F(1, 11) = 176$, $MSE = 171$, $\eta_p^2 = .94$, $p < .001$, *pitch*, $F(1, 11) = 108$, $MSE = 159$, $\eta_p^2 = .91$, $p < .001$, and *brightness*, $F(1, 11) = 11.15$, $MSE = 62.35$, $\eta_p^2 = .50$, $p = .007$, but not in their *size*, $F(1, 11) < 1$, or *complexity*, $F(1, 11) = 3.60$, $MSE = 6.12$, $\eta_p^2 = .25$, $p = .08$.

Typeface

Because a typeface is essentially a family of elementary geometric forms, it was decided that the typeface in which words would appear, in both the preliminary work and the experiments, should be neutral in its perceived hardness, (auditory) pitch, and brightness. With this in mind, Times New Roman was selected. This choice was informed by data from 15 new undergraduate students who rated the full lowercase alphabet in this and four other frequently used text typefaces (Arial, Century, Courier New, Franklin Gothic Book) on the *soft–hard*, *high–low*(pitch), and *dark–bright* scales. With regard to the pitch scale, participants were asked to “Imagine this text came to life and made a sound, what would be the pitch of this sound?” Times New Roman emerged as the most neutral of the typefaces. Its average rating (with *SD* in parentheses) on the hardness, pitch, and brightness scales was 4.13 (0.74), 4.27 (0.70), and 4.06 (0.57), respectively, with associated single-sample *t* values (all *dfs* = 15) of 0.70 ($p = .13$), 1.47 ($p = .16$), and 0.44 ($p = .67$), respectively.

¹¹ Sign tests confirmed the same deviations from the neutral value to be significant.

¹² In this and other studies (e.g., Walker & Walker, *in press*) in which scales of this kind have been used to gather cross-sensory feature judgments, participants’ responses have yielded data for which an ANOVA is an appropriate form of analysis. Although, in a previous study, nonparametric analyses were reported, the picture emerging from ANOVAs was identical (see Walker & Walker, *in press*).

Words

Thirteen new undergraduate students, who also did not contribute to the experiments reported below, rated each of 50 words on the *soft–hard*, *high–low* (pitch), or *dark–bright* scale, whichever was appropriate for the word.¹³ Words with mean ratings on the relevant scales that were significantly different from the neutral (*neither*) value of 4.0, as assessed with single-sample *t* tests, became candidates for use as to-be-classified stimuli in the experiments. Where possible, it was also arranged for the words in any of the sets to be varied in nature (e.g., *squeak*, *flute*, and *mouse* were selected for the set of high-pitched words used in Experiment 2). An ideal situation would have been one in which all the words in a set had nothing in common other than their association with the same polarized value on the relevant connotative dimension. On the basis of these ratings, two sets of five words whose meanings contrasted in hardness emerged as being appropriate test stimuli for Experiment 1. Two sets of six words whose meanings contrasted in pitch became available for Experiment 2, and two sets of four words contrasting in brightness became available for Experiment 3. Details about the words in each set are provided in the account of the experiment in which they were used.

Visual and phonological angularity of the words

In light of Westbury's (2005) study, it was anticipated that confirmation would be needed that any effects induced by the angularity of a surrounding shape arose from its congruity with the meaning of the test word, and not from its congruity with either the shape or the phonology of this word. Once the results from the speeded classification task were to become available, such confirmation was to be obtained by entering an index of the angularity of the shape of each word, and then of its phonological angularity, as a covariate in an item- (word) based analysis of the difference in mean RT between trials involving curved shapes and trials involving angular shapes (i.e., $\text{mean RT}_{\text{curved}} - \text{mean RT}_{\text{angular}}$).

Nineteen new undergraduate students rated each of the words selected for use in Experiments 1–3 for the visual angularity of their shapes as they appeared in Times New

Roman. A 6-point scale was created,¹⁴ anchored at one end by the terms *angular/sharp* and at the other end by the terms *curved/smooth*. Successively less extreme points on the scale were marked by *very*, *quite*, and *slightly*, and there was no neutral midpoint. The average visual angularity ratings for the words used in Experiments 1–3 are provided in the account of the experiment in which they were used.

With regard to phonological angularity, the proportion of consonants in each word that were continuant rather than plosive in nature was determined by a colleague with expertise in this domain.¹⁵ These values, for which a lower proportion indicates greater phonologically angularity, also are provided in the account of the experiment in which the words were used.

Experiment 1: Visual angularity and hardness

Words were classified in Experiment 1 on the basis of the hardness of the material to which they refer. The preliminary rating procedure undertaken to find English words appropriate for such a classification task yielded two sets of five words contrasting in meaning (see Table 2). These words were presented as test stimuli and were classified by participants according to whether the concepts to which they refer have associations with hardness or softness (i.e., refer to a material that is hard or soft).

Method

Participants Fourteen employees of the NHS Business Services Authority in Preston (7 males, 7 females; age range, 18–35 years) completed the hardness classification task.

Design and procedure A within participants 10×2 design was used, with word (10 test words) and congruity (connotatively congruent vs. incongruent) as the two factors.

Participants completed two blocks of trials, within each of which the 10 test words appeared once inside each of the six shapes after this had been flipped horizontally, or not, and vertically, or not, on a random basis. For half of the trials, therefore, the connotative hardness of the shape within which a test word appeared was congruent with the level of hardness suggested by the word (e.g., *angular/granite*).

¹³ The 50 words were the outcome of the author and two colleagues generating commonly used English words referring unambiguously to polarized values along the dimensions. The English language provides varying numbers of such words for different dimensions. Because not all of these words then attracted sufficiently consistent ratings, and because of the need to minimize confounds across contrasting sets of words to be used in each experiment (e.g., avoiding systematic differences in word-initial letter), the speeded classification tasks used in Experiments 1–3 employed varying numbers of polarized test words.

¹⁴ A 6-point scale was used here because the data were gathered in the context of another study in which there was a preference for scales that do not have a neutral midpoint and which, therefore, force participants to make a judgment toward one pole or the other. This switch has no implications for how well the study reported here was able to assess the ideas being proposed in the research.

¹⁵ The author is grateful to Padraic Monaghan for assessing the phonological angularity of each word.

Table 2 Mean ratings on the *soft–hard* dimension (with *SDs* in parentheses) for the *soft* and *hard* words used in Experiment 1, together with single-sample *t* values assessing the significance of deviations from the neutral (*neither*) value of 4.0

SOFT		HARD	
dough	3.69 (0.48) 2.31*	brick	5.38 (0.51) 9.86***
fur	2.69 (0.63) 7.48***	diamond	6.15 (0.55) 14.00***
plasticine	2.31 (0.75) 8.12***	granite	5.35 (0.63) 7.48***
sponge	3.08 (0.86) 3.86**	rock	5.85 (0.55) 12.00***
velvet	3.69 (0.48) 2.31*	steel	6.08 (0.64) 11.69***

Note. Mean ratings less (more) than 4.0 indicate that soft (hard) was judged most appropriate. Mean ratings significantly different from the neutral value of 4.0 are noted

* $p < .05$

** $p < .01$

*** $p < .001$

For the other trials, the connotative hardness of the shape within which the test word appeared was incongruent with the level of hardness suggested by the word (e.g., angular/*fur*). A different random order of presentation for the 60 word–shape combinations in each block was determined online for every participant.

The word and shape appeared and disappeared together as part of the same black-on-white display. Each word appeared in 24-point Times New Roman, and participants classified it, as quickly as possible, according to whether it referred to something that was hard or to something that was soft. They did this by pressing the “z” key or the “/” key with their left and right index fingers, respectively. The assignment of word type (hard or soft) to key/hand was counterbalanced across participants.

Each display remained on the screen until a response was made, at which point it was replaced with a blank white screen. The next display appeared after an interval of 3 s. A brief rest interval was provided between the two blocks of trials.

Stimulus presentation and response monitoring were controlled by a Dual 2-GHz Apple PowerMac G5 (interfaced with an Apple A1038, 1,680 × 1,050 cinema back-lit LCD display), running version 2.1.1 of the PsyScript experiment generator.

Results

The results are summarized in Table 3.

A nonparametric analysis of accuracy focused simply on assessing the presence of a connotative congruity effect. The

main analysis focused on a participant-based ANOVA of correct RT. An alpha level of .05 was used for all statistical tests.

Accuracy When participants’ overall levels of accuracy on congruent and incongruent trials were compared, it was confirmed that accuracy was significantly higher on congruent trials than on incongruent trials, $p(\text{corr}) = .99$ and .97, respectively, Wilcoxon Signed Ranks $z = 2.37$, $p = .009$.

Reaction time To get the RT data fit for analysis, RTs associated with incorrect responses were excluded, and excessively long RTs for each participant (i.e., more than 2 *SDs* above their mean correct RT) were replaced by the cutoff value. A 10×2 repeated measures ANOVA was undertaken with word and congruity as the two factors. Although there was not a significant effect of word, $F(1, 9) = 1.48$, $MSE = 0.01$, $\eta_p^2 = .10$, $p = .17$, there was a significant effect of congruity, $F(1, 13) = 7.84$, $MSE = 0.83$, $\eta_p^2 = .38$, $p = .015$. Participants responded more quickly on congruent trials than on incongruent trials ($M = 754$ and 863 ms, respectively). The word \times congruity interaction was not significant, $F(9, 117) = 1.53$, $MSE = 0.014$, $\eta_p^2 = .10$, $p = .15$.

Item-based analysis and word shape angularity Values for the visual and phonological angularity of the test words are given in Table 4.

A 2×5 repeated measures ANOVA was undertaken on the angularity ratings of the words, with hardness and word as the two factors. Because different words were involved within each level of hardness, word was a dummy variable. There was a significant effect of the level of hardness (i.e., hard vs. soft) referred to in the meaning of a word, $F(1, 18) = 6.16$,

Table 3 Mean correct RT (*SEM* in parentheses) and $p(\text{correct})$ for each of the *soft* and *hard* words in Experiment 1 according to whether it appeared in a connotatively congruent or connotatively incongruent shape

Test word	Congruent	Incongruent	RT difference (inc – con)
dough	750 (30) 1.00	884 (69) 0.98	134
fur	726 (28) 0.99	823 (59) 0.99	97
plasticine	738 (31) 0.99	905 (68) 0.96	167
sponge	746 (43) 0.98	864 (53) 0.94	118
velvet	779 (33) 0.99	920 (64) 0.96	141
brick	800 (40) 0.99	830 (66) 0.98	30
diamond	786 (45) 1.00	858 (47) 0.93	72
granite	759 (35) 1.00	851 (52) 0.96	92
rock	750 (47) 1.00	821 (51) 0.98	71
steel	705 (30) 1.00	875 (67) 1.00	170

Note. Normal and bold text relate to conditions in which test words appeared inside curved and angular shapes, respectively

Table 4 Mean ratings of visual angularity (1 = *very angular/sharp*, 6 = *very curved/smooth*) for the *soft* and *hard* words used in Experiment 1

Soft		Hard	
dough	4.10 (0)	brick	2.16 (.33)
fur	4.68 (1.00)	diamond	3.37 (.50)
plasticine	2.11 (.67)	granite	3.95 (.50)
sponge	4.74 (.75)	rock	3.00 (.50)
velvet	1.47 (.75)	steel	2.79 (.67)

Note. In parentheses is given the proportion of consonants in each word that are continuant rather than plosive (with a lower value indicating greater phonological angularity)

$MSE = 6.45$, $\eta_p^2 = .26$, $p = .023$.¹⁶ Hard words were judged to have shapes that were more *angular/sharp* (mean rating = 3.05) than the shapes of the soft words (mean rating = 3.42). A 2×5 item-based ANOVA was then undertaken on the RT difference between trials where a word was surrounded by a curved shape and trials where the same word was surrounded by an angular shape (i.e., mean $RT_{\text{curved}} - \text{mean } RT_{\text{angular}}$). Hardness and word were treated as two between-item factors (again with word as a dummy variable), and the mean angularity rating for each word's shape was entered as a covariate. The ANOVA confirmed a significant effect of the level of hardness indicated by a word's meaning, $F(1, 7) = 75.43$, $MSE = 1.69$, $\eta_p^2 = .92$, $p < .001$, with the overall value for the mean $RT_{\text{curved}} - RT_{\text{angular}}$ difference being 90 ms for hard words and -134 ms for soft words. The angularity of each word's shape did not significantly moderate the effect of the angularity of the surrounding shape, $F(1, 7) = 1.29$, $MSE = 0.03$, $\eta_p^2 = .16$, $p = .29$.

Item-based analysis and phonological angularity The five *soft* and five *hard* words had mean phonological angularity indices of 0.63 and 0.50, respectively, indicating that the *hard* words tended to be phonologically more angular than the *soft* words. The same item-based ANOVA on the mean $RT_{\text{curved}} - RT_{\text{angular}}$ difference, but with the phonological angularity of the words replacing their visual angularity as the covariate, again confirmed a significant effect of the level of hardness indicated by a word's meaning, $F(1, 7) = 75.61$, $MSE = 1.67$, $\eta_p^2 = .92$, $p < .001$. The phonological angularity of the test words did not significantly moderate the effect of the angularity of the surrounding shape, $F(1, 7) = 1.42$, $MSE = 1.67$, $\eta_p^2 = .92$, $p = .2$.

Discussion

When words are classified according to whether they suggest hardness or softness, they are classified more quickly and

accurately when they appear concurrently with (and inside) shapes with congruent connotations of hardness based on their angularity. Specifically, words suggesting hardness (softness) are classified relatively quickly when they appear in an angular (curved) shape. This effect is not mediated by the visual or phonological angularity of the words.

Experiment 2: Visual angularity and auditory pitch

In Experiment 2, words were classified on the basis of the pitch of the sound to which they refer. The rating procedure undertaken to find English words appropriate for such a classification task yielded two sets of six words contrasting in meaning (see Table 5). Following the same procedure as in Experiment 1, these words had to be classified by participants according to whether the concepts to which they refer have associations with high-pitched or low-pitched sounds.

Method

Participants Twelve employees of the NHS Business Services Authority in Preston (6 males, 6 females; age range, 19–33 years) completed the pitch classification task.

Design and procedure The design and procedure were the same as those for Experiment 1.

Table 5 Mean ratings on the high–low pitch dimension (with SDs in parentheses) for the high- and low-pitched words used in Experiment 2, together with single-sample *t* values assessing the significance of deviations from the neutral (*neither*) value of 4.0

HIGH PITCHED		LOW PITCHED	
crack	3.39 (0.87) 2.55*	boom	5.15 (0.55) 7.50***
flute	2.92 (0.76) 5.11***	cow	4.69 (0.95) 2.64*
mouse	3.54 (0.66) 2.52*	growl	4.54 (0.66) 2.94**
scream	2.15 (0.80) 8.31***	rumble	5.85 (0.55) 12.00***
snap	3.46 (0.88) 2.21*	thunder	5.15 (0.69) 6.04***
squeak	1.85 (0.55) 14.00***	tuba	6.08 (0.86) 8.68***

Note: *** $p < .001$, ** $p < .01$, and * $p < .05$

Note. Mean ratings less (more) than 4.0 indicate that high (low) was judged most appropriate. Mean ratings significantly different from the neutral value of 4.0 are noted

¹⁶ This was the only result in Experiments 1–4 that was not confirmed with the equivalent nonparametric analysis.

Results

The results are summarized in Table 6.

Accuracy Overall levels of accuracy were not significantly different across congruent and incongruent trials, $p(\text{corr}) = .99$ and $.98$, respectively, Wilcoxon signed ranks $z = 1.38$, $p = .08$.

Reaction time A 12×2 repeated measures ANOVA was undertaken, with word (12 test words) and congruity (connotatively congruent vs. incongruent) as the two factors. There were significant main effects of word, $F(11, 121) = 2.15$, $MSE = 0.07$, $\eta_p^2 = .16$, $p = .02$, and of congruity, $F(1, 11) = 15.52$, $MSE = 1.90$, $\eta_p^2 = .59$, $p = .002$. Participants responded more quickly on congruent trials than on incongruent trials ($M = 910$ and $1,072$ ms, respectively). The word \times congruity interaction was not significant, $F < 1$.

Item-based analysis and word shape angularity Values for the visual and phonological angularity of the test words are given in Table 7.

A 2×6 repeated measures ANOVA was undertaken on the angularity ratings of the words, with pitch and word as the two factors. Word was again a dummy variable. There was not a significant effect of the level of pitch (i.e., high vs. low) indicated by the meaning of a word, $F(1, 18) < 1$. The shapes of the high-pitched words had a mean *angular/sharp* rating of 3.81, and the shapes of the low-pitched words had a mean rating of 3.77. A 2×6 item-based ANOVA was undertaken on the RT difference between trials where each word was

Table 6 Mean correct RT (*SEM* in parentheses) and $p(\text{correct})$ for each of the low-pitched and high-pitched words in Experiment 2 according to whether it appeared in a connotatively congruent or connotatively incongruent shape

Test word	Congruent	Incongruent	RT difference (inc – con)
boom	954 (99) 0.97	1108 (97) 0.97	154
cow	968 (69) 1.00	1224 (104) 0.94	256
growl	938 (71) 0.99	1119 (92) 0.99	181
rumble	887 (64) 1.00	1066 (102) 1.00	179
thunder	926 (93) 1.00	1095 (102) 0.97	169
tuba	921 (80) 1.00	1064 (98) 1.00	143
crack	903 (73) 0.99	1038 (94) 0.99	135
flute	889 (70) 0.97	1097 (87) 0.99	208
mouse	959 (67) 0.99	1037 (76) 0.99	78
scream	885 (59) 1.00	1023 (91) 0.99	138
snap	884 (78) 0.99	1038 (105) 0.97	154
squeak	801 (62) 1.00	956 (90) 1.00	155

Note. Normal and bold text relate to conditions in which test words appeared inside curved and angular shapes, respectively

Table 7 Mean ratings of visual angularity (1 = *very angular/sharp*, 6 = *very curved/smooth*) for the words used in Experiment 2

HIGH PITCHED		LOW PITCHED	
crack	2.79 (.33)	boom	5.16 (.50)
flute	3.00 (.67)	cow	3.84 (0)
mouse	5.47 (1.00)	growl	2.89 (.67)
scream	5.00 (.75)	rumble	4.21 (.67)
snap	3.79 (.67)	thunder	2.79 (.75)
squeak	2.79 (.50)	tuba	3.74 (.33)

Note. In parentheses is given the proportion of consonants in each word that are continuant rather than plosive (with a lower value indicating greater phonological angularity)

surrounded by a curved shape and trials where the same word was surrounded by an angular shape (i.e., mean $RT_{\text{curved}} - \text{mean } RT_{\text{angular}}$). Pitch and word were treated as two between-item factors (again with word as a dummy variable), and the mean angularity rating for each word's shape was entered as a covariate. The ANOVA confirmed a significant effect of the level of pitch indicated by a word's meaning, $F(1, 9) = 195.6$, $MSE = 3.81$, $\eta_p^2 = .96$, $p < .001$, with the overall value for the mean $RT_{\text{curved}} - RT_{\text{angular}}$ difference being 145 ms for high-pitched words and -180 ms for low-pitched words. The angularity of each word's shape did not significantly moderate the effect of the angularity of the surrounding shape, $F(1, 9) = 1.30$, $MSE = 0.03$, $\eta_p^2 = .13$, $p = .18$.

Item-based analysis and phonological angularity The six high-pitched and six low-pitched words had mean phonological angularity indices of 0.65 and 0.49, respectively, suggesting that the high-pitched words were, if anything, phonologically less angular than the low-pitched words. The same item-based ANOVA on the mean $RT_{\text{curved}} - RT_{\text{angular}}$ difference, but with the phonological angularity of the words replacing their visual angularity as the covariate, confirmed a significant effect of the level of pitch indicated by a word's meaning, $F(1, 9) = 150.4$, $MSE = 3.30$, $\eta_p^2 = .94$, $p < .001$. The phonological angularity of the test words did not significantly moderate the effect of the angularity of the surrounding shape, $F < 1$.

Discussion

When words are classified according to whether they suggest high-pitched or low-pitched sounds, they are classified more quickly when they appear concurrently with (and inside) shapes with congruent connotations of pitch based on their angularity. Specifically, words suggesting high-pitched (low-pitched) sounds are classified relatively quickly when they appear in an angular (curved) shape. This effect is not mediated by the visual or phonological angularity of the words.

Experiment 3: Visual angularity and brightness conveyed verbally

In Experiment 3, words were classified on the basis of the level of brightness to which they refer (i.e., whether this is bright or dark). The rating procedure undertaken to find English words appropriate for such a classification task yielded two sets of four words contrasting in meaning (see Table 8). Following the same procedure as in Experiment 1, these words had to be classified by participants according to whether the concepts to which they refer have associations with brightness or with darkness.

Method

Participants Seventeen undergraduate students at Lancaster University (10 males, 7 females; age range, 19–24 years) completed the brightness classification task.

Design and procedure The design and procedure were essentially the same as those for Experiments 1 and 2.

Results

The results are summarized in Table 9.

Accuracy Overall, levels of accuracy were not significantly different across congruent and incongruent trials, $p(\text{corr}) = .96$ and $.94$, respectively, Wilcoxon signed ranks $z = 1.07$, $p = .28$.

Reaction time An 8×2 repeated measures ANOVA was undertaken, with word and congruity as the two factors.

Table 8 Mean ratings on the dark–bright dimension (with SDs in parentheses) for the *dark* and *bright* words used in Experiments 3, together with the single-sample t values assessing the significance of deviations from the neutral (*neither*) value of 4.0

Dark		Bright	
dark	2.54 (0.52) 10.16***	bright	5.84 (0.55) 12.00***
gloom	1.92 (0.64) 11.69***	flash	6.23 (0.60) 13.42***
murky	2.54 (0.52) 10.16***	gleam	5.62 (0.65) 8.95***
shade	2.46 (0.53) 10.69***	shine	5.46 (0.52) 10.16***

Note. Mean ratings less (more) than 4.0 indicate that dark (bright) was judged most appropriate. Mean ratings significantly different from the neutral value of 4.0 are noted

* $p < .05$

** $p < .01$

*** $p < .001$

Table 9 Mean correct RT (*SEM* in parentheses) and $p(\text{correct})$ for each *dark* and *bright* word in Experiment 3 according to whether it appeared in a connotatively congruent or incongruent shape

Test word	Congruent	Incongruent	RT difference (inc – con)
dark	653 (50) 0.98	679 (59) 0.98	26
gloom	684 (41) 1.00	746 (64) 0.92	62
murky	671 (41) 0.99	678 (39) 1.00	7
shade	717 (47) 0.82	719 (47) 0.89	2
bright	679 (47) 0.97	712 (60) 0.93	33
flash	666 (55) 0.98	705 (54) 0.96	39
gleam	680 (52) 0.92	812 (75) 0.89	132
shine	668 (65) 0.98	698 (50) 0.98	30

Note. Normal and bold text relate to conditions in which test stimuli appeared inside curved and angular shapes, respectively

There were significant main effects of word, $F(7, 112) = 2.50$, $MSE = 0.02$, $\eta_p^2 = .14$, $p = .02$, and of congruity, $F(1, 16) = 28.01$, $MSE = 0.11$, $\eta_p^2 = .64$, $p < .001$. Participants responded more quickly on congruent trials than on incongruent trials ($M = 677$ and 719 ms, respectively). The word \times congruity interaction also was significant, $F(7, 112) = 2.13$, $MSE = 0.02$, $\eta_p^2 = .12$, $p = .015$. Inspection of the mean RTs in Table 9 suggests that this interaction reflects the fact that the effect of congruity was stronger for the brighter words than for the darker words.

Item-based analysis and word shape angularity Values for the visual and phonological angularity of the test words are given in Table 10.¹⁷

A 2×4 repeated measures ANOVA was undertaken on the angularity ratings of the words, with brightness and word as the two factors. Word was again a dummy variable. There was not a significant effect of the level of brightness (bright vs. dark) indicated by the meaning of a word, $F < 1$. The shapes of the *bright* words had a mean *angular/sharp* rating of 3.51, and the shapes of the *dark* words had a mean rating of 3.62. A 2×4 item-based ANOVA was undertaken on the mean $RT_{\text{curved}} - \text{mean } RT_{\text{angular}}$ difference. Brightness and word were treated as two between-item factors (with word as a dummy variable), and the mean angularity rating for each word's shape was entered as a covariate. The ANOVA confirmed a significant effect of the level of brightness indicated by a word's meaning, $F(1, 5) = 7.44$, $MSE = 0.24$, $\eta_p^2 = .60$, $p = .04$, with the overall mean value for the $RT_{\text{curved}} - RT_{\text{angular}}$ difference being 59 ms for *bright* words and -24 ms for *dark* words. The angularity of each word's

¹⁷ Across all 30 words used in Experiments 1–3, there was not a significant association between visual and phonological angularity (Spearman's $\rho = .27$, $p = .15$, two-tailed).

Table 10 Mean ratings of visual angularity (1 = *very angular/sharp*, 6 = *very curved/smooth*) for the words used in Experiment 3

Dark		Bright	
dark	2.79 (0)	bright	2.26 (.33)
gloom	4.79 (.67)	flash	3.79 (1.00)
murky	3.11 (.50)	gleam	4.42 (.67)
shade	3.79 (.50)	shine	3.58 (1.00)

Note. In parentheses is given the proportion of consonants in each word that were continuant rather than plosive (with a lower proportion indicating greater phonological angularity)

shape did not significantly moderate the effect of the angularity of the surrounding shape, $F < 1$.

Item-based analysis and phonological angularity The sets of *dark* and *bright* words had mean phonological angularity indices of 0.42 and 0.75, respectively, suggesting that the *bright* words were phonologically less angular than the *dark* words. The same item-based ANOVA on the mean $RT_{\text{curved}} - RT_{\text{angular}}$ difference, in which the phonological angularity of the words replaced their visual angularity as the covariate, confirmed a significant effect of the level of brightness indicated by a word's meaning, $F(1, 5) = 6.40$, $MSE = 0.24$, $\eta_p^2 = .56$, $p = .05$. The phonological angularity of the test words did not significantly moderate the effect of the angularity of the surrounding shape, $F < 1$.

Discussion

When words are classified according to whether they suggest brightness or darkness, they are classified more quickly and accurately when they appear concurrently with shapes with congruent connotations of brightness based on their angularity. Specifically, words suggesting brightness (darkness) are classified relatively quickly when they appear in an angular (curved) shape. This effect is not mediated by the visual or phonological angularity of the words. Experiment 4 provides an additional check that the visual and phonological characteristics of the test words were not essential contributors to this congruity effect. It does so by presenting to-be-classified levels of brightness nonverbally, thereby removing word characteristics from the stimulus situation.

Experiment 4: Visual angularity and brightness conveyed visually

Evans and Treisman (2010) observed congruity effects based on cross-modality correspondences between auditory pitch and a number of visual features (i.e., visual size, visuospatial height, and visuospatial frequency). They did

not observe a congruity effect between visual size and visuospatial height (where small is assumed to correspond with high), and this raises the possibility that correspondences are exclusively cross-modality in nature, with the congruity effects they induce being confined to situations in which stimuli are encoded in different modalities. The same is implied by Spence's (2011) preference for the term *cross-modality correspondences*. However, if correspondences are mediated by the semantic coding of elementary stimulus features, it should not matter whether two concurrent features are encoded in the same or different modalities. All that should matter is that they can have separate connotative meanings. The term *cross-sensory correspondences* allows for this.

Because the to-be-classified words in Experiment 3 were presented visually, it might be argued that their interaction with visual angularity demonstrates a within-modality congruity effect. However, the feature being classified was conveyed via word meaning independently of the modality through which it is accessed (i.e., in principle, the same results would be expected had the words been presented aurally). A more compelling demonstration of a within-modality congruity effect would involve presenting to-be-classified levels of brightness visually but nonverbally. Experiment 4 was designed to do this.

Experiment 3 was modified to create a situation in which contrasting levels of brightness were presented visually and nonverbally. The same geometric shapes were presented as solid achromatic forms whose brightness was set at one of four levels (i.e., the shapes "enclosed" a level of perceived visual brightness, rather than a word referring to a level of brightness). Participants classified each shape, as quickly as possible, according to whether it was perceived to be brighter or darker than the mid-gray background against which it appeared. The angularity of the shapes was again a task-irrelevant feature. The expectation was that a congruity effect would be observed whereby participants would respond more quickly and accurately when the angularity of the shape was in correspondence with the brightness of the shape relative to the mid-gray background.

When a test shape is brighter or darker than the mid-gray background against which it appeared, its surface brightness takes on one of two values. This noticeable variation in surface brightness, which has no implications for response selection, might itself interact with shape angularity to yield a congruity effect. That is, within the levels of surface brightness associated with a particular task-defined category of brightness (i.e., within those brightness levels that are higher than the background and within those levels that are lower than the background), participants might respond more quickly when higher (lower) levels of brightness are paired with the more angular (curved) shapes. If the predicted congruity effect originates in levels of processing at

or beyond the level at which stimuli are categorized for purposes of response selection, a congruity effect arising from these within-category variations in brightness would not be expected. This would be in line with Martino and Marks's (2001) claim that cross-sensory correspondences (what they call *weak synaesthesia*) are largely based on the context-sensitive coding of stimulus features, rather than on their absolute feature values, provided that context sensitivity refers not just to the other stimuli being presented, but also to the specific requirements of the task (e.g., how stimuli are being classified).

It will be noticed that whereas the brightness words used in Experiment 3 concerned perceived *illuminant brightness*, the manipulation of brightness in this experiment concerns perceived *surface brightness* (cf. Marks, 1987, for evidence that these are distinct types of brightness and that it is the latter that is being manipulated in Experiment 4).¹⁸ Unlike perceived illuminant brightness, perceived surface brightness is not a magnitude dimension (not *prothetic* in Stevens's, 1957, terminology) (see Marks, 1974, 1982, 1987; Smith & Sera, 1992). Because of this, these two types of brightness can show different patterns of correspondence (e.g., both forms of brightness interact with auditory pitch to give rise to congruity effects in speeded classification, but only perceived illuminant brightness interacts with loudness to yield such effects; see Marks, 1987). In the context of the present study, both types of brightness are expected to behave in the same way because angularity, like auditory pitch, is not a magnitude dimension and will not enter into correspondences on that basis.

Method

The angular and curved shapes were randomly flipped vertically, or not, and horizontally, or not, before each presentation. Two of the achromatic colors in which they appeared (*black* and *dark gray*, 2 and 15 cd/m², respectively) were darker than the mid-gray background (70 cd/m²). Two (*light gray* and *white*, 165 and 320 cd/m², respectively) were brighter than the background. Participants had to classify each test shape as quickly as possible according to whether it was brighter or darker than the mid-gray background against which it appeared. It was predicted that participants would respond more accurately and/or more quickly when

the connotative brightness of the shape, based on its angularity, was congruent, rather than incongruent, with its surface brightness, relative to the background. With four levels of achromatic color for a shape and two types of shape, there are eight combinations, four of which are congruent (i.e., *angular/white*, *angular/light gray*, *curved/dark gray*, and *curved/black*), and four of which are incongruent (i.e., *angular/black*, *angular/dark gray*, *curved/light gray*, and *curved/white*). Variations in brightness within a level mapping on to one of the responses were not expected to interact with shape angularity to induce a congruity effect.

Participants Eighteen undergraduate students at Lancaster University (9 males, 9 females; age range, 19–27 years) completed this version of the speeded classification task.

Design and procedure The design and procedure were essentially the same as those for the previous experiments.

A within-participants 4 × 2 design was used, with achromatic color (*black*, *dark gray*, *light gray*, and *white*) and congruity (connotatively congruent vs. incongruent) as the two factors.

Participants completed two blocks of trials, within each of which each of the six shapes appeared once at every level of brightness. For half of the trials in each block, therefore, the connotative brightness of the shape, based on its angularity, was congruent with its surface brightness relative to the background (e.g., *angular/white*). For the other trials, the connotative brightness of the shape was incongruent with its surface brightness relative to the background (e.g., *angular/dark gray*). A different random order of presentation for the 24 shape–brightness combinations in each block was determined online for every participant.

Results

The results are summarized in Table 11.

Accuracy Overall levels of accuracy were not significantly different across congruent and incongruent trials, $p(\text{corr}) = .98$ and $.97$, respectively, Wilcoxon signed ranks $z = 0.72$, $p = .47$.

Reaction time A 2 × 2 × 2 repeated measures ANOVA was undertaken. The first factor was between-category brightness, defined according to whether a shape was brighter or darker than the background. The second factor, within-category brightness, was defined according to the level of brightness within a level of categorical brightness—that is, whether the achromatic color was relatively bright (i.e., *white* and *dark gray*) or dark (i.e., *light gray* or *black*) within its level of categorical brightness. The third factor was congruity (connotatively congruent vs. connotatively incongruent).

¹⁸ The perception of illuminance (i.e., the amount of light provided by a source) and the perception of surface luminance (i.e., the amount of light coming from a surface) are both often referred to as perceived *brightness*. Although most studies of cross-sensory correspondences involve the perception of surface luminance, a few involve the perception of illuminance (see Lewkowicz & Turkewitz, 1980; Marks, 1978; Stevens, 1975). It is perception of the former, surface luminance, here referred to as *brightness* or *surface brightness*, which is of concern in Experiment 4.

Table 11 Mean correct RT (*SEM* in parentheses) and *p*(correct) for each level of brightness in which a shape appeared, relative to the brightness of the background, according to whether it appeared as a connotatively congruent or connotatively incongruent shape

Shape brightness	Congruent	Incongruent	RT difference (inc – con)
Darker than mid-gray background			
black	481 (27) 0.97	501 (35) 0.99	20
dark gray	534 (36) 0.98	569 (50) 0.96	35
Brighter than mid-gray background			
light gray	530 (30) 0.99	541 (28) 0.97	11
white	482 (27) 0.99	508 (29) 0.98	26

Note. Normal and bold text relate to conditions in which the different brightness levels appeared as curved and angular shapes, respectively

There was not a significant main effect of between-category brightness, $F < 1$, or of within-category brightness $F(1, 16) = 1.09$, $MSE = 0.003$, $\eta_p^2 = .06$, $p = .31$. There was, however, a significant interaction between these two factors, $F(1, 16) = 27.63$, $MSE = 0.09$, $\eta_p^2 = .63$, $p < .001$, reflecting the fact that participants were quicker to respond the higher the brightness contrast was between the shape and the background (i.e., participants were quicker to classify black and white shapes than to classify dark gray and light gray shapes). The main effect of congruity was significant, $F(1, 16) = 12.90$, $MSE = 0.017$, $\eta_p^2 = .45$, $p = .002$, with participants responding more quickly on congruent trials than on incongruent trials ($M = 507$ and 530 ms, respectively). Congruity did not interact significantly with either between-category brightness or within-category brightness (both F s < 1), nor did it enter into a three-way interaction with these two factors, $F = 0$. The last result confirms that the level of brightness contrast between a circle and the mid-gray background did not interact with angularity to yield a congruity effect.

Item-based analysis A single-factor item-based ANOVA was undertaken on the RT difference between trials where each level of brightness appeared as a curved shape and trials where the same level of brightness appeared as an angular shape (i.e., mean $RT_{\text{curved}} - \text{mean } RT_{\text{angular}}$). When between-category brightness was the between-item factor, there emerged a significant effect of the level of brightness, confirming the congruity effect observed with the participant-based analysis, $F(1, 2) = 17.55$, $MSE = 0.04$, $\eta_p^2 = .90$, $p = .05$. The overall mean value for the $RT_{\text{curved}} - \text{mean } RT_{\text{angular}}$ difference was 18 ms for the brighter levels of achromatic color (i.e., *white* and *light gray*) and -27 ms for the darker levels of achromatic color (i.e., *dark gray* and *black*). There was no such effect of brightness when it was within-category brightness (i.e., *white* + *dark gray* vs. *light gray* + *black*) that was entered as the between item factor, $F = 0$.

Discussion

A congruity effect involving brightness and angularity was observed when contrasting values for both features were encoded visually and nonverbally. This result, which was predicted on the basis that the cross-sensory correspondences explored in the present study have a semantic basis, indicates that cross-sensory correspondences are not necessarily cross-modality in nature. Instead, it is proposed that correspondences can involve any two elementary stimulus features that have their own connotative meanings, regardless of the sensory channels over which they are encoded.

A subsidiary aspect of the results from Experiment 4 provides further indication that the brightness–angularity congruity effect arose from interactions at relatively late stages of information processing, at or beyond the level at which the stimuli were classified according to their brightness in support of response selection. Thus, a congruity effect was not induced within those levels of brightness converging on the same response (i.e., *brighter* or *darker*). This is in line with Martino and Marks's (2001) claim that cross-sensory correspondences are largely based on the context-sensitive coding of stimulus features, rather than on their absolute feature values.

Finally, it is noted that whereas the to-be-classified words in Experiment 3 referred to different levels of illuminant brightness, the to-be-classified colors in Experiment 4 involved different levels of surface brightness. Together, therefore, these two experiments reveal that both forms of perceived visual brightness enter into correspondence with visual angularity in the same way—that is, with more angular being aligned with brighter.

General discussion

The four experiments reported here support the proposal that cross-sensory correspondences, and the congruity effects they induce, can reflect extensive and bidirectional cross-activation among dimensions of connotative meaning (see Fig. 1). On this basis, the same core correspondences should emerge whichever sensory feature is used to probe them, confirming that the *en bloc* alignment of the dimensions is context invariant. For example, given that higher-pitched sounds appear to be sharper than lower-pitched sounds, sharper (i.e., more angular) shapes should be higher in pitch than curved shapes and should share their other cross-sensory features with higher-pitched sounds. Selecting the cross-sensory features of hardness, pitch, and brightness as test cases, the results from the preliminary observations and Experiments 1–3 confirm these predictions and the premises on which they are based.

When angular and curved outline shapes were judged for their cross-sensory features, angular shapes were deemed to be harder, higher-pitched, and brighter than curved shapes. When the same shapes served as incidental stimuli in a speeded-classification task, their angularity induced the predicted congruity effects. Words whose meanings suggested contrasting levels of hardness, pitch, and brightness were classified more quickly and correctly when the connotations of the incidental shape within which they appeared were congruent, rather than incongruent, with their meaning. Because of the nature of the speeded-classification task in the present study and the fact that the shapes and their angularity were incidental stimuli, the connotations linked to shape angularity must have been registered quickly and automatically.

In agreement with Karwowski et al.'s (1942) proposal that cross-sensory correspondences involve the connotative meanings of elementary stimulus features, the congruity effects reported here were observed after steps had been taken to guarantee a semantic basis for the correspondences on which they were based, largely by minimizing, if not precluding, the types of nonsemantic correspondence identified by Spence (2011). These steps involved specifying to-be-classified feature values verbally, using words from contrasting sets as the stimuli to be classified, and ensuring that none of the test words marked contrasting levels of angularity/sharpness in everyday language. The fact that the nonsemantic features of the test words (i.e., their visual and phonological angularity) did not moderate any of the congruity effects is consistent with participants responding to the words on a semantic basis.

Cross-sensory correspondences having a semantic basis (e.g., in the connotative meanings of elementary stimulus features) and the congruity effects they induce should not be restricted to situations in which stimulus features are encoded in different sensory modalities. Instead, any stimulus features having connotative meanings should be able to enter into correspondences, whether the sensory channels over which they are encoded are in the same modality or in different modalities. For this reason, proposing a semantic basis for correspondences entails viewing them as being essentially cross-sensory, rather than cross-modality, in nature. Experiment 4 provided direct support for this view by presenting both the criterial and incidental feature values visually (and nonverbally). More specifically, the brightness classification task used successfully in Experiment 3 was adapted to allow contrasting levels of visual brightness to be presented visually. Despite confining the encoding of angularity and brightness to the visual modality, and despite shifting the focus from illuminant brightness to surface brightness, a brightness–angularity congruity effect was observed, with increases in angularity again being congruent with increases in brightness.

With regard to the possibility that the correspondences explored here could have a lexical basis, as defined by Spence (2011), this has been dismissed on the grounds that there are no commonly used verbal labels in English that mark contrasting feature values on more than one of the dimensions.¹⁹ Although the author is aware of two exceptions, these relate to specialized fields of activity. First, contrasting types of sand used in building work in the U.K. are occasionally referred to as being *sharp* and *soft*, the latter term implying that sharp sand is also hard sand. Second, in Western musical notation, modest raising and lowering of the pitch of a note is communicated with symbols labeled *sharp* and *flat*, respectively. But these instances of lexical overlap are both modest and specialized, and so, for the moment at least, it seems appropriate to accept that there is little, if any, potential for the three correspondences examined here to have the kind of lexical basis envisaged by Spence. Nevertheless, a more general role for verbal recoding in the present experiments needs to be considered.

It might be argued that the angularity of each incidental shape in the speeded classification task was automatically recoded verbally, thereby engaging the same lexical representations accessed by the criterial stimuli (e.g., *bright*, *high*, and *hard* for the angular shapes, and *dark*, *low*, and *soft* for the curved shapes). This would allow interactions among these lexical representations, rather than among semantic representations, to be responsible for any congruity effects. It is important to note, however, that pictorial stimuli seem not to access verbal representations (e.g., the names of pictured objects or their features) directly, but only indirectly via semantic representations (Garrard, Lambon Ralph, Patterson, Pratt, & Hodges, 2005; Lambon Ralph, McClelland, Patterson, Galton, & Hodges, 2001; McGregor, Friedman, Reilly, & Newman, 2002; Morton, 1985). Thus, lexical representations linked to shape angularity would need to be activated in a top-down manner by semantic representations of angularity. But these semantic representations of angularity would not activate lexical representations concerning hardness, elevation, and brightness. Instead, such activation would come about through top-

¹⁹ Spence (2011) suggested that a correspondence with a semantic basis could emerge from a situation where the same verbal labels are used to mark contrasting values on different dimensions of perceptual experience (e.g., *high* and *low* for pitch and spatial elevation). The evidence suggests otherwise, however, and suggests that it is the close association between dimensions at the level of their connotative meaning that encourages them to share the same verbal labels. Thus, when Eitan and Timmers (2010) examined the terms used to label contrasting levels of pitch in different cultures, their equivalent of *high* and *low* was not always used to mark contrasting levels of auditory pitch. In these cases, their terms for *light/heavy*, *sharp/blunt*, *small/big*, *thin/thick*, and *weak/strong* were often used instead. With the exception of *weak/strong*, these are just the terms that the core set of connotative dimensions discussed here would predict.

down influence from semantic representations of hardness, elevation, and brightness, respectively. This line of reasoning assumes that cross-activation occurs between representations of connotative sharpness and representations of connotative hardness, elevation, and brightness, which is just the position being promoted in the present study. It will be appreciated, therefore, that even if evidence were to be forthcoming confirming the automatic verbal recoding of the angularity of a task-irrelevant shape,²⁰ the correspondences responsible for congruency would still reside in the functional organization of the connotative meanings of elementary stimulus features.

Could the correspondences explored in the present study have, in some other situations, a structural and/or statistical basis? Although illuminant brightness can function as a magnitude dimension (see, e.g., Marks, 1987), pitch and surface brightness appear not to (see, e.g., Marks, 1974, 1982, 1987; Smith & Sera, 1992), and so correspondences involving them are unlikely to have a structural basis. Because it is unclear whether angularity and hardness also can function as magnitude dimensions, it remains uncertain whether they can enter into correspondences having a structural basis.

With regard to the correspondences perhaps having a statistical basis in some situations, although it is very difficult to find natural co-occurrences supportive of the correspondence between angularity and brightness, co-occurrences can be found for the correspondences between angularity and each of hardness and pitch.²¹ First, there is a natural co-occurrence between angularity and hardness, because angular (sharp/pointy) objects tend to be formed from relatively hard materials. Second, this co-occurrence could, in turn, mediate an association between angularity and pitch, because objects made from harder materials tend to make higher-pitched sounds when struck (Freed, 1990; Klatzky, Pai, & Krotkov,

2000; van den Doel & Pai, 1998). It is proposed here, however, that any cross-sensory correspondences acquired through exposure to such natural co-occurrences will be semantic, rather than nonsemantic, in nature, which would take them outside Spence's (2011) account of statistical correspondences. This is because the co-occurrences themselves will be conceptual in nature, with the *category* of harder materials co-occurring with the *categories* of sharper objects and higher-pitched sounds.

This raises the interesting possibility, for which there is some evidence, that cross-sensory correspondences having a semantic basis can be responsible for the perceptual element in some congruity effects. For example, Roffler and Butler (1968) demonstrated that the pitch of a sound influences a person's perception of the spatial elevation of the sound, even for children who are not yet familiar with the application of the labels *high* and *low* to pitch. Parise and Spence (2008, 2009) demonstrated that the correspondence between auditory pitch and each of visual size and visual angularity induces perceptual congruity effects. Specifically, when brief visual and auditory stimuli are presented in close temporal proximity, congruity among their features (e.g., high pitch with visual pointedness) increases the probability that the stimuli will be perceived to be co-occurring. Finally, Walker, Francis, and Walker (2010) showed that the correspondence between surface brightness and weight can alter how heavy an object feels when hefted, with a darker object feeling lighter in weight than an equally heavy brighter object. If statistical correspondences have a semantic (conceptual) basis, there is a real possibility that one or more of these effects of cross-sensory correspondence on perception have a semantic basis.

It seems that Spence (2011) did well to anticipate that his typology of cross-sensory correspondences would probably need to be expanded to include correspondences having a semantic basis.

Author Note The author is grateful to Diane Bisset for contributing to the collection of data for Experiments 1 and 2.

References

- Collier, W. G., & Hubbard, T. L. (2001). Judgements of happiness, brightness, speed and tempo change of auditory stimuli varying in pitch and tempo. *Psychomusicology*, 17, 36–55.
- Eitan, Z., & Timmers, R. (2010). Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. *Cognition*, 114, 405–422.
- Elliott, L. L., & Tannenbaum, P. H. (1963). Factor-structure of semantic differential responses to visual forms and prediction of factor-scores from structural characteristics of the stimulus shapes. *The American Journal of Psychology*, 76, 589–597.
- Evans, K., & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision*, 10, 1–12.

²⁰ There is much to mitigate against such automatic verbal recoding. For example, it would be competing for resources with concurrent verbal processing of the critical stimulus and verbal components of response selection. In addition, automatic verbal recoding would be expected to occur only in relation to overlearned verbal labels for familiar visual material, and not for aspects of novel pictorial stimuli (i.e., novel shapes).

²¹ With regard to the possibility that a correspondence between two features reflects the existence of a natural co-occurrence between them, the absence of solid evidence one way or the other means that contrasting claims can be made (i.e., absence of evidence is not evidence for absence). A real need exists for a comprehensive program of research to confirm the existence of natural co-occurrences, even those that are generally assumed to occur (e.g., between size and pitch), and to assess the consistency with which they support a particular correspondence. For example, although several natural co-occurrences might link angularity and brightness, they might do so in contradictory ways, some aligning angular with bright, others aligning it with dark. If so, then further work would be required to determine the relative impact each one has on the overall correspondence, perhaps reflecting the varying degrees to which people are exposed to the different co-occurrences.

- Freed, D. J. (1990). Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events. *Journal of the Acoustical Society of America*, 87, 311–322.
- Garrard, P., Lambon Ralph, M. A., Patterson, K., Pratt, K. H., & Hodges, J. R. (2005). Semantic feature knowledge and picture naming in dementia of Alzheimer's type: A new approach. *Brain and Language*, 93, 79–94.
- Hubbard, T. L. (1996). Synesthesia-like mappings of lightness, pitch, and melodic interval. *The American Journal of Psychology*, 109, 219–238.
- Karwowski, T. F., Odber, H. S., & Osgood, C. E. (1942). Studies in synesthetic thinking: II. The role of form in visual responses to music. *The Journal of General Psychology*, 26, 199–222.
- Klatzky, R. L., Pai, D. K., & Krotkov, E. P. (2000). Perception of material from contact sounds. *Presence*, 9, 399–410.
- Lambon Ralph, M. A., McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R. (2001). No right to speak? The relationship between object naming and semantic impairment: Neuropsychological evidence and a computational model. *Journal of Cognitive Neuroscience*, 13, 341–356.
- Lewkowicz, D. J., & Turkewitz, G. (1980). Cross-modal equivalences in early infancy: Auditory–visual intensity matching. *Developmental Psychology*, 16, 597–607.
- Lundholm, H. (1921). The affective tone of lines: Experimental researches. *Psychological Review*, 28, 43–60.
- Marks, L. E. (1974). On associations of light and sound: The mediation of brightness, pitch, and loudness. *The American Journal of Psychology*, 87, 173–188.
- Marks, L. E. (1975). On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin*, 82, 303–331.
- Marks, L. E. (1978). *The unity of the senses: Interrelations among the modalities*. New York: Academic Press.
- Marks, L. E. (1982). Bright sneezes and dark coughs, loud sunlight and soft moonlight. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 177–193.
- Marks, L. E. (1987). On cross-modal similarity: Auditory–visual interactions in speeded discrimination. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 384–394.
- Marks, L. E. (2004). Cross-modal interactions in speeded classification. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *Handbook of multisensory processes* (pp. 85–105). Cambridge: MIT Press.
- Martino, G., & Marks, L. E. (1999). Perceptual and linguistic interactions in speeded classification: Tests of the semantic coding hypothesis. *Perception*, 28, 903–923.
- Martino, G., & Marks, L. E. (2001). Synesthesia: Strong and weak. *Current Directions in Psychological Science*, 10, 61–65.
- McGregor, K. K., Friedman, R. M., Reilly, R. M., & Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research*, 45, 332–346.
- Melara, R. D., & Marks, L. E. (1990). Processes underlying dimensional interactions: Correspondences between linguistic and non-linguistic dimensions. *Memory & Cognition*, 18, 477–495.
- Monaghan, P., Mattock, K., & Walker, P. (in press). The role of sound symbolism in word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.
- Mondloch, C. J., & Maurer, D. (2004). Do small white balls squeak? Pitch–object correspondences in young children. *Cognitive, Affective, & Behavioral Neuroscience*, 4, 133–136.
- Morton, J. (1985). Naming. In S. Newman & R. Epstein (Eds.), *Current perspectives in dysphasia* (pp. 217–230). Edinburgh: Churchill Livingstone.
- Osgood, C. E. (1960). The cross-cultural generality of visual-verbal synesthetic tendencies. *Behavioral Sciences*, 5, 146–169.
- Parise, C., & Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neuroscience Letters*, 442, 257–261.
- Parise, C., & Spence, C. (2009). ‘When birds of a feather flock together’: Synesthetic correspondences modulate audiovisual integration in non-synesthetes. *PLoS One*, 4, e5664. doi:10.1371/journal.pone.0005664
- Perrott, D. R., Musicant, A., & Schwethelm, B. (1980). The expanding-image effect: The concept of tonal volume revisited. *The Journal of Auditory Research*, 20, 43–55.
- Poffenberger, A. T., & Barrows, B. E. (1924). The feeling value of lines. *Journal of Applied Psychology*, 8, 187–205.
- Roffler, S. K., & Butler, R. A. (1968). Localization of tonal stimuli in the vertical plane. *Journal of the Acoustical Society of America*, 43, 1260–1266.
- Scheerer, M., & Lyons, J. (1957). Line drawings and matching responses to words. *Journal of Personality*, 25, 251–273.
- Smith, L. B., & Sera, M. D. (1992). A developmental analysis of the polar structure of dimensions. *Cognitive Psychology*, 24, 99–142.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971–995.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153–181.
- Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural, and social prospects*. London: Wiley.
- van den Doel, K., & Pai, D. K. (1998). The sounds of physical shapes. *Presence*, 7, 382–395.
- Walker, P., Francis, B. J., & Walker, L. (2010). The brightness–weight illusion: Darker objects look heavier but feel lighter. *Experimental Psychology*, 57, 462–469.
- Walker, P., & Smith, S. (1984). Stroop interference based on the synaesthetic qualities of auditory pitch. *Perception*, 13, 75–81.
- Walker, P., & Smith, S. (1985). Stroop interference based on the multimodal correlates of haptic size and auditory pitch. *Perception*, 14, 729–736.
- Walker, P., & Smith, S. (1986). The basis of Stroop interference involving the multimodal correlates of auditory pitch. *Perception*, 15, 491–496.
- Walker, P., & Walker, L. (in press). Size–brightness correspondence: Crosstalk and congruity among dimensions of connotative meaning. *Attention, Perception, & Psychophysics*. doi:10.3758/s13414-012-0297-9
- Ward, J., Stott, R., & Parkin, A. J. (2000). The role of semantics in reading and spelling: Evidence for the ‘summation hypothesis’. *Neuropsychologia*, 38, 1643–1653.
- Westbury, C. (2005). Implicit sound symbolism in lexical access: Evidence from an interference task. *Brain and Language*, 93, 10–19.