What sound does that taste? Cross-modal mappings across gustation and audition

Julia Simner, Christine Cuskley¶, Simon Kirby¶
Department of Psychology (¶ Department of Linguistics and English Language), School of Psychology, Philosophy and Language Sciences, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, Scotland, UK; e-mail: j.simner@ed.ac.uk
Received 5 October 2009, in revised form 20 January 2010

Abstract. All people share implicit mappings across the senses, which give us preferences for certain sensory combinations over others (eg light colours are preferentially paired with higher-pitch sounds; Ward et al, 2006 *Cortex* **42** 264 – 280). Although previous work has tended to focus on the cross-modality of vision with other senses, here we present evidence of systematic cross-modality between taste and sound. We created four sound continua representing to varying extents phonetic qualities of speech (F1, F2, voice discontinuity, and spectral balance). Sixty-five participants selected their preferred sound to accompany each of the basic tastes of sweet, sour, bitter, and salty, at two different concentrations. We found significant shared preferences among our participants to map certain acoustic qualities to certain types of tastes (eg sweet tastes tend to be mapped to a lower spectral balance than sour tastes). We also found a preference for mapping certain sound qualities to different taste concentrations. Together our data provide the first detailed analysis of how phonetic features map systematically to different tastants and concentrations. We examine the roots of these mappings, and discuss how such associations might guide the ways in which human languages are used to name objects with taste.

1 Introduction

All people are predisposed to prefer certain sensory combinations over others. For example, we tend to pair higher-pitch sounds with smaller, brighter, and spikier shapes than lower-pitch sounds, and we also map higher-pitch sounds to higher positions in space (Bernstein and Edelstein 1971; Karwoski et al 1942; Marks et al 1987). Such cross-modal mapping may serve a functional purpose, and one recent study, for example, suggests they might facilitate multisensory integration (Parise and Spence 2009). The majority of empirical studies in cross-modality have tended to examine correspondences within the visual domain (eg between space, brightness, and size—Nicholls et al 1999) or correspondences between vision and the other senses (eg between colour and touch—Simner and Ludwig 2009). However, humans experience cross-modality across a range of sensory domains, and our aim in this study is to show systematic cross-sensory mappings between taste and sound. We present empirical evidence of correspondences between qualities of pure tastants, and the acoustic qualities of sound, particularly those related to speech and language. We also demonstrate that these systematic mappings can be relatively explicit, in that they can be detected when participants freely associate across these dimensions. In presenting this evidence, we ask whether the naming of objects with taste is an arbitrary process, or whether it might be guided by certain underlying cross-modal principles.

More so than any of the traditional five senses, taste is most often experienced cross-modally because of the sensory interactions associated with eating (see Auvray and Spence 2008; Delwiche 2004 for reviews). When we consume food, we exploit chemoreceptors on our tongue, but also smell the odour of the food, see its form, texture, and colour, and feel its form and texture in the mouth. However, the experience of flavour is primarily the combination of odour and taste (Djordjevic et al 2004). Indeed, the absence or addition of odorants can affect perceived taste quality.

Patients with anosmia (loss of sense of smell), for example, report that their ability to taste is also often affected (Doty 2009). Conversely, adding odorants to a taste solution can enhance or detract from a given taste: a strawberry odour can enhance perceived sweetness, and a caramel odour can reduce perceived sourness (Stevenson et al 1999). In reverse, taste can also enhance the perception of odour. For example, a sweet tastant will increase the perception of a fruity odour (Verhagen and Engelen 2006).

Taste perception can also be influenced by manipulating visual and tactile qualities. Colour can affect perceived taste, since flavour discrimination is adversely affected when the colour of a solution does not correspond with its taste (Zampini et al 2007). Perceived sweetness can also be heightened by textural changes, such as increases in viscosity (Bult et al 2007; Christensen 1980). Several studies also show differences in perceived taste or flavour when manipulating the sound produced by the consumption of food. Zampini and Spence (2004) showed that changing the amplitude or frequency components within the sound produced when biting into crisps (potato chips) influenced ratings of their crunchiness or freshness. Other studies have manipulated the sound of carbonated drinks, or the sounds produced when chewing (Christensen and Vickers 1981; Zampini and Spence 2005). This body of literature (see Spence and Zampini 2006 for an overview) examines the acoustic properties generated by the food itself. In response, Crisinel and Spence (2009) asked whether there are cross-modal relationships between taste and specific auditory qualities per se. Since our own work takes its lead from this study, we describe it in detail below.

Crisinel and Spence (2009) described early studies by Holt-Hansen (1968, 1976), which showed that participants associated two different brands of beer with different acoustic pitches. Crisinel and Spence asked whether the roots of these findings may lie in the different qualities of sour and bitter within the two brands, and they showed empirically that people hold systematic implicit associations between certain pitches and certain food names. Their participants categorised food names as denoting flavours with either bitter qualities (coffee, beer, tonic water, dark chocolate) or sour qualities (lime, lemon juice, vinegar, pickles). In the same session, and using the same buttons, they also categorised sound stimuli as being either high pitch or low pitch. In the experimental manipulation the way that buttons were shared across tasks was varied such that participants responded either [high-pitch/sour; low-pitch/bitter] or [high-pitch/bitter; low-pitch/sour]. Response latency and accuracy showed a preferential pairing of [high-pitch/sour; low-pitch/bitter]. This is an intriguing finding, and suggests that sound and taste may cross-map in the same way as other combinations of the senses. However, there are certain features of Crisinel and Spence's particular choice of stimuli which raise questions about the interpretation of their study, and we address these issues here.

Crisinel and Spence's use of food names to map cross-modal properties gives rise to certain problems in interpretation, some of which they address themselves in their discussion. Crisinel and Spence point out that there are clear differences between reading food names and experiencing actual tastes, and that different people may interpret food names in different ways (eg some may find beer more bitter than others). More importantly, however, there is a problem that arises from the very nature of linguistic stimuli, and we draw attention to this here. The food names of Crisinel and Spence may have triggered unwanted cross-modal influences entirely unrelated to taste. There is now a growing cross-modal literature showing that people have inherent sensory preferences for certain linguistic features (eg graphemes, phonemes, lexical frequency). For example, all people share implicit preferences for the colours of graphemes (A tends to be red; L tends to be yellow, etc—Simner et al 2005) and they also select more luminant shades of colour for graphemes of higher frequency (Smilek et al 2007).

People also pair linguistic sequences such as months and numbers to specific positions in peripersonal space (eg early months and low numbers to left-sided space for speakers or left-to-right writing systems—Dehaene et al 1993; Gevers et al 2003). They also show underlying mappings between other linguistic properties and the visual dimension of shape. One demonstration of this is in the classic 'bouba-kiki' paradigm of Ramachandran and Hubbard (2001; see also Davis 1961; Kohler 1929, 1947; Maurer et al 2006): people tend to label spiky shapes as *kikis* and rounded shapes as *boubas* (see Sapir 1929 for related sound symbolism paradigms). Finally, Klink (2000) showed that non-words containing voiceless consonants are judged to be smaller, faster, lighter, sharper, and softer than non-words with voiced consonants. All these findings show that people implicitly map sensory qualities to the linguistic features of words. In turn, this suggests that the linguistic stimuli of Crisinel and Spence may have exerted unwanted influences themselves, and that this might have played a role in the cross-modal mapping those authors attributed to taste.

To assess this issue, we performed a posteriori analysis on the words used by Crisinel and Spence (2009) and found significant differences in linguistic qualities across conditions (in addition to the taste manipulation planned by Crisinel and Spence). Words in their sour condition were more linguistically marked in that they were significantly lower in frequency than words in the bitter condition (respective group means were 9.1 and 107.6—Kucera and Francis 1967; $t_6 = 2.4$, $p = 0.05$). In addition, the bitter-food names contained a significantly higher proportion of rounded graphemes (those with curved contours in their visual form, such as *o*, *e*, *a*) compared to sour-food names, as a function of overall word length. The mean ratio of rounded graphemes in bitter-food names was 0.91, and in sour-food names 0.65 ($t_6 = 2.9$, $p = 0.03$). It is therefore possible that the taste mappings (to pitch) presented by Crisinel and Spence were in fact linguistic mappings, arising from differences across conditions in the angularity/rounding of graphemes, or in the lexical markedness/frequency. Indeed, there are several reasons to find this plausible. Participants in the Crisinel and Spence study mapped high pitch to the most angular words, and Karwoski et al (1942) have already shown that people pair the notion of high pitch with the notion of angularity. Furthermore, work from our own lab shows that cross-modal attributions are sensitive to the angularity of graphemes. For example, non-words containing angular consonants such as z (eg zeze) are preferred as names for objects that are angular, while non-words containing rounded consonants such as g (eg gege) are preferred as the names of shapes that are rounded (Cuskley et al 2009). In addition, Smilek et al (2007) show that cross-modal mappings are also sensitive to linguistic frequency. In other words, where Crisinel and Spence presented cross-modal matching between low pitch and bitter, they may instead have demonstrated cross-modal matching between low pitch and rounded graphemes (or, indeed, between low pitch and low markedness in terms of lexical frequency).

One other cause for a re-examination of the Crisinel and Spence findings is that their stimuli also confounded acoustic pitch with timbre. Their high-pitch stimuli were instantiated by notes from piano, clarinet, violin, and trumpet, while their low-pitch stimuli were instantiated by piano, bassoon, cello, and bass trombone. Ward et al (2006) have shown that timbre can play an important role in the cross-modal associations made by the average person, since participants attribute different timbre to different colour qualities (eg significantly more saturated colours are attributed to piano versus pure tones). To address this issue, and the issues described above, we aimed to re-examine the cross-modality of sound and taste using a different method. First, we presented our participants with tastants rather than food names, and we manipulated acoustic sounds without timbre confounds. In presenting our findings, we aimed to show that Crisinel and Spence were in fact correct in their assumptions

about taste – sound correspondence and that, importantly, effects of this type can be seen when all other influences are removed. We also had two further aims for our study. First, we asked whether taste – sound mappings might be explicit, in as much as they could be measured in tasks where participants make explicit judgments (ie when matching tastes in the mouth to sounds played over headphones). Second, we wanted to test whether taste – sound associations are also found in acoustic stimuli related, to varying degrees, to the phonetic features of speech. If humans map tastes to sounds within the acoustic boundaries of speech, this might suggest that the naming of objects with taste could have arisen in a non-arbitrary way. This in turn might have important implications for theories of language evolution and vocabulary construction.

Here we report our empirical study in which participants were asked to match tastes in the mouth to sounds played over headphones. We administered four types of tastant (sweet, salty, bitter, sour) each along two concentrations (medium, high). Taste was administered as drops on the tongue, and participants were then required to match each taste to whatever settings they preferred on four different sound continua. These four types of sound were selected to examine different acoustic qualities: F1, F2, voice discontinuity, and spectral balance. The first three of these relate to properties of speech, and were chosen because they are well motivated in the literature on linguistics (eg Ladefoged 1993), sound symbolism (eg Hinton et al 1994), and cross-modality (eg Maurer et al 2006; Ramachandran and Hubbard 2001, 2005). F1 corresponds perceptually to vowel height (Ladefoged 1993) and is the frequency of the first formant. Vowel height is a phonemic distinction for English speakers, and is perceptually salient (eg Giegerich 2005; Pfitzinger 2003). Figure 1 shows (inter alia) the outer vowels of British English, and their respective F1s (see also section 2.1 for other details shown in this figure).
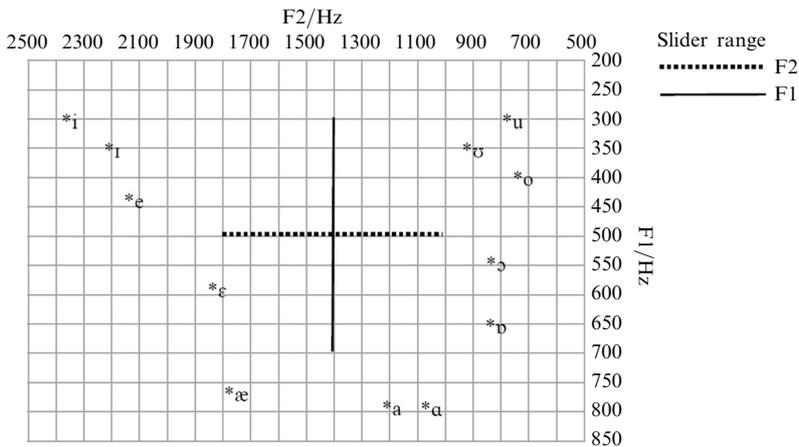

Figure 1. Vowel space in British English, with the ranges of our F1 slider and F2 slider shown in solid and dashed, respectively. Adapted from Hayes (2008).

Although vowels of different heights have formant differences *other* than F1 (ie slight differences in F3 and F4), a difference in F1 is most perceptually salient. As such, varying F1 alone results in a synthesised vowel that sounds to be changing in height (Pfitzinger 2003). F2 is the frequency of the second formant, and this corresponds (within the ranges we have chosen) to vowel front/backness. This measure is well established in other linguistic and cross-modal literature (eg Diffloth 1994; Maurer et al 2006; Newman 1933; Sapir 1929), and, again, this feature is salient and phonemic for English speakers. Figure 1 also shows the outer vowels of British English, and their respective F2s. As with F1 values, the formants of spoken front and back vowels also contain more than simply a change in F2 (ie small changes in F3 and F4).

However, changes in F2 correlate with the perceived notion of front/backness (Ladefoged 1993; Pfitzinger 2003). (F2 values can also affect perceived vowel roundedness, but this is not phonemic in English—Giegerich 2005; Traunmuller et al 2003.)

Our third speech-related acoustic continuum is that of voice discontinuity, and this was chosen to reflect the interruptions in vowel quality that occur in natural speech from interspersed voiceless consonants in words. Voicing is also found in sound symbolic and cross-modal literature (Kohler 1929, 1947; Maurer et al 2006; Ohala 1984, 1994; Ramachandran and Hubbard 2001, 2005). Our final sound continuum was a manipulation of the balance of overall spectral energy of noise emitted. Unlike the other sliders, which produce a synthesised vowel output, the output of this slider was filtered white noise. In this continuum, the value of all four formants moves together as the slider is manipulated from low to high. Perceptually, the low value of the slider sounds like lower-pitched white noise, and the high end of the slider sounds like higher-pitched white noise. This manipulation allows us to examine the association between taste stimuli and the spectral balance of acoustic energy in isolation.

Although the paucity of work on taste – sound associations makes some predictions difficult, there are certain specific hypotheses to propose. If Crisinel and Spence (2009) show patterns of data that do reflect taste – sound correspondences, we predict that participants will share systematic preferences for the sound qualities they associate with each taste quality. Crisinel and Spence (2009) suggest that high pitch corresponds to sour tastes while low pitch corresponds to bitter tastes, and so we may find that sour tastes are attributed higher values than bitter tastes in selections of F1, F2 (as vowel height and backness can affect perceived pitch of a vowel; Gonzales 2009; Pape and Mooshammer 2008), and in spectral balance (which reflects a shift in the frequency concentration of overall noise). We may also find that people engage in a type of 'intensity matching' (eg Smith and Sera 1992), in that higher concentrations of taste may be matched to higher values in the sound continua. Finally, we may find in our voice discontinuity slider a type of 'hedonic matching', in which two ends of a dimension subjectively correspond because both are pleasant (as shown in other cross-modal studies—eg Dematte et al 2006a, 2006b). On the assumption that the most pleasant taste is sweet (eg Moskowitz et al 1974), this taste may match with the smooth auditory quality of a continuous vowel, rather than the staccato interruptions of a discontinuous vowel (ie sweet tastants will be lowest in voice discontinuity).

2 Experimental methods

2.1 Participants

Sixty-five participants were recruited form the University of Edinburgh community, and were compensated £3.50 for the 25 min required to complete the task. All participants were monolingual English speakers between the ages of 18 and 42 years. All were non-smokers and all reported normal taste and hearing function. Smokers were excluded to avoid possible taste function impairment (Vennemann et al 2008). Informed consent was obtained from all participants, and ethical approval was obtained from the University of Edinburgh prior to testing.

2.2 Materials

2.2.1 *Taste materials.* Our materials comprised four tastants (sweet, salty, bitter, sour), each at two concentrations (medium, high). Our stimuli were taken from the Accusens Taste Kit© (Henkin 2005) which instantiates tastants as sweet (glucose), salty (NaCl), bitter (urea), and sour (HCl). The kit contains three different concentrations of each taste, hereafter referred to as low, medium, and high (see table 1). The low concentration of each taste is slightly above the average recognition threshold for individuals with normal taste function (Brosovic and McLaughlin 1989). Each concentration level

Table 1. Tastant materials and their concentrations. All four tastes (sweet, salty, bitter, sour) were presented at two different concentrations (medium, high). Note: low concentration is not examined in the current study.

Flavour	Tastant	Concentration/mM		
		low	medium	high
Sweet	sucrose	60	90	150
Salty	NaCl	60	90	150
Bitter	urea	150	300	500
Sour	HCl	6	15	30

thereafter demonstrates at least a 30% increase in intensity (Ganong 2005), making the difference in concentration discernible at equal intervals for each taste (Henkin 2005). From these levels of concentration, we selected both medium and high, but excluded low as our participants were not able to consistently detect taste at the lowest level.

2.2.2 *Sound materials.* Four sound sliders were created to present sound continua of F1, F2, voice discontinuity, and spectral balance, respectively. Each slider was programmed with Tcl/Tk and the Snack$^{©}$ real-time sound synthesis package, modelled after Beskow's (2001) formant synthesiser. With the exception of the voice discontinuity slider, the amplitude of all sounds was set at 70 dB within the programme. In all sliders, the band widths for F1, F2, F3, and F4 were held constant at 50, 75, 100, and 150 Hz, respectively. The individual settings for each slider are described below.

F1 slider. To examine vowel height, we created a slider that varied F1 on a continuum from low to high frequency while holding all other formants constant. Low and high settings of F1 were 300 and 700 Hz, respectively, which is roughly analogous to spanning the centre of the vowel space in British English, from high to low vowels. The corresponding spectrograms for these lower and upper settings are shown in the panels of figure 2, and figure 1 shows (with a solid line) the range of the F1 slider with respect to vowel space and frequency. F3 and F4 were held at 2500 and 3500 Hz, respectively, which are the default values given by Beskow (2001). F2 frequency was held at 1400 Hz, which represents the mid-point of our F2 slider (see below).

F2 slider. To examine vowel front/backness, we created a slider that varied F2 on a continuum from low to high frequency while holding all other formants constant. Figure 1 shows the range of the slider, indicated by the dashed line. F3 and F4 frequencies were held at 2500 and 3500 Hz, respectively (following defaults given in Beskow 2001). F1 was held at 500 Hz which represents the mid-point of the F1 slider. Low and high settings of F2 were 1000 and 1800 Hz, respectively—approximately spanning the centre of the English vowel space from back to front vowels (and remaining firmly within the vowel space, as dictated by the fixed value of F1).

Voice discontinuity. To examine voice discontinuity, we created a slider that interrupted an otherwise continuous vowel, to an increasing extent on a continuum, according to the value of the slider. We selected a vowel roughly equivalent to /ə/ (schwa), with the following formant values: F1 = 500 Hz, F2 = 1500 Hz, F3 = 2500 Hz, and F4 = 3500 Hz. To manipulate perceived discontinuity, we varied the volume of the sound every 360 ms, back and forth between two amplitude levels. These amplitude levels became further apart as the slider moved from low to high. When the slider was set at zero, there was no variation in volume, and the result was a smooth, fully continuous vowel. As the slider increased, the perceived effect was a vowel sound with a throbbing pulse of volume change, resulting eventually in a fully discontinuous

vowel alternating between sound and silence. The slider's scale runs from 0 to 3000, and its units are our own 'discontinuity quotient' (DQ).

We effected our manipulation by connecting the slider to a counter which varied the volume of the vowel around a baseline setting of 75 dB. Snack$^©$ denotes this level as an arbitrary value of 2000, and this value forms the basis of our DQ scale, as follows. Every unit on the slider represents a fluctuation of 1 either side of the baseline 2000. For example, at a setting of 0 DQ, the slider plays a vowel with no fluctuation, which therefore stays at the constant amplitude of 2000, and sounds fully continuous. At a setting of 1 DQ, the slider plays a vowel which fluctuates in amplitude from 2001 to 1999 (ie 2000 ± 1). At a setting of 1000 DQ, the slider plays a vowel which fluctuates in amplitude from 3000 to 1000 (ie 2000 ± 1000) and this sounds somewhat discontinuous, with throbbing pulses of volume change. At 2000 DQ, the slider plays a vowel that fluctuates even more, from 4000 to 0 (ie 2000 ± 2000), and this is now a fully discontinuous vowel interspersed with silence. The maximum setting of 3000 DQ emphasises that discontinuity by alternating between silence and the highest volume of 5000 (ie $2000 + 3000$). Note that because the volume cannot fall below zero, our lowest setting of volume (ie $2000 - 3000 = -1000$) is realised simply as zero or silence.

Spectral balance. To examine overall acoustic energy, we created a slider that varied spectral balance on a continuum. This was accomplished by using a white-noise source and manipulating the value of all four formants simultaneously from 0 to 5000 Hz, resulting in a shift in the spectral balance of the sound from low to high.

2.3 Procedure
The experiment was run on a desktop PC attached to Sennheiser PXC 250 head-phones. The PC volume was set to 50% for each participant to ensure consistent delivery of stimuli. Participants were given brief written instructions about overall aims, and were then shown the PC interface. This interface showed four horizontal sliders placed one above the other, in the centre of the screen. Each slider could be manipulated with a mouse by clicking on a virtual grip, and dragging it back and forth across the length of the slider. This changed the quality of the related sound (F1, F2, voice discontinuity, spectral balance, respectively). Instructions were then given verbally with the sliders on screen to facilitate understanding. Participants were told they would receive drops of 12 tastes. These comprised 8 target items [four tastants (sweet, salty, bitter sour); each at two concentrations (medium and high)] and 4 fillers (the same tastants at low concentration—not examined here). Participants were told that after each taste, they would be required to choose a sound from each of the four on-screen sliders. They were told to select one setting from each slider (four sounds in total) that 'best matched' the taste, and that there was no right or wrong answer. Participants then put on headphones, and the experimenter prepared to administer one drop of each tastant directly onto the tongue. The drop was delivered from the pipette of the standard 4 ml bottles provided with the taste kit. Drops were administered directly onto the centre of the tongue, and participants were encouraged to roll the solution around in the mouth. To ensure participants understood the procedure they first received a drop of water, and were encouraged to interact with the sound sliders before beginning the test items. Participants received distilled water between each tastant to cleanse the palate. Tastes were administered to each subject in one of ten pseudo-randomised orders, with the restriction that no two concentrations of one taste quality could be administered directly in succession. Sound sliders were also presented in one of four different random orders on screen, such that each slider occupied each possible position. Slider values were also counterbalanced, such that approximately half the participants had low values on the left and high on the right, while the remaining participants had the reverse.

3 Results

Data from each slider were analysed separately with a 4×2 within-subjects ANOVA crossing taste quality (sweet, salty, sour, bitter) and taste concentration (medium, high). All family-wise comparisons are Bonferroni corrected. We did not perform cross-slider analyses because each slider presented qualitatively different sounds, with different scale lengths that would make any interactions difficult to interpret. In doing this, we also kept the total number of tests small, and so minimised the required correction of the significance level (following Curtin and Schulz 1998). Because of equipment failure, we lost the data from one participant on the spectral balance slider, and from another participant on both spectral balance and voice discontinuity. The results from each slider are given below.

3.1 *F1 slider (vowel height)*

Sixty-six participants contributed data to this analysis. The mean F1 values selected for each taste at each concentration (where low F1 values correspond to a high vowel) are shown in figure 2. There was a significant main effect of taste concentration ($F_1 = 6.7$, $p = 0.01$) showing that higher concentrations of all tastes corresponded to sounds with higher F1 (ie a lower vowel). There was also a main effect of taste quality ($F_3 = 7.6$, $p < 0.001$). Planned corrected tests revealed that the sweet taste had a significantly lower F1 (463 Hz) than the bitter taste (508 Hz; $t_{65} = -2.8$, $p = 0.007$), the salty taste (511 Hz; $t_{65} = -2.9$, $p = 0.005$), and the sour taste (538 Hz; $t_{65} = -5.2$, $p < 0.001$). All other planned statistics were non-significant (all $ts < |1.8|$, all $ps > 0.05$) and there was no interaction of factors ($F_3 < 1$).
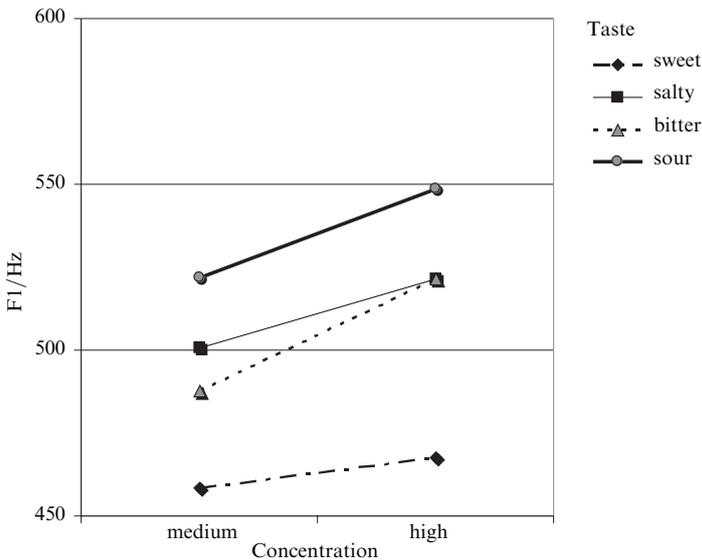

Figure 2. Mean F1 values selected for four tastes (sweet, salt, bitter, sour) at two concentrations (medium, high). High F1 corresponds to low vowels.

3.2 *F2 slider (vowel backness)*

Sixty-six participants contributed data to this analysis. The mean F2 values selected for each taste at each concentration (where low F2 values correspond to a back vowel) are shown in figure 3. As before, participants were again systematic in their choice of taste–sound association. There was a significant main effect of taste concentration on the F2 quality of sound ($F_1 = 7.7$, $p = 0.007$), since higher concentrations of tastes were assigned to significantly higher F2. There was no interaction of effects ($F_3 < 1$), but there was a main effect of taste quality ($F_3 = 2.6$, $p = 0.05$). In numerical terms,
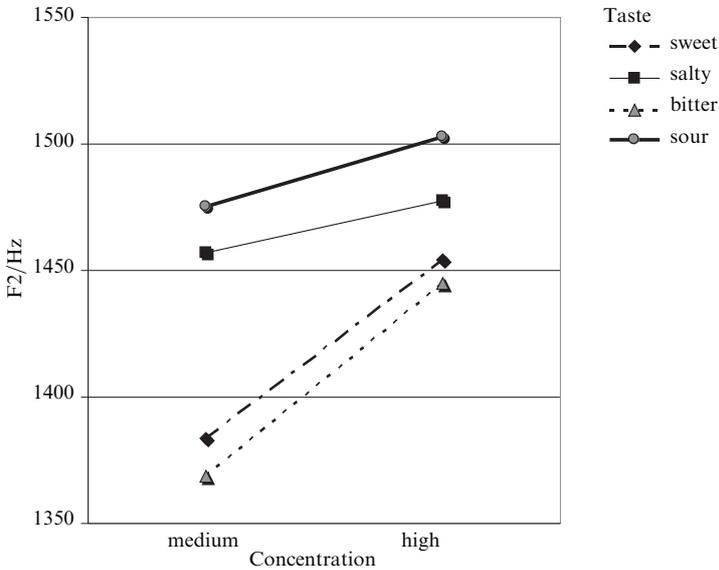
Figure 3. Mean F2 values selected for four tastes (sweet, salt, bitter, sour) at two concentrations (medium, high). High F2 corresponds to front vowels.

the sweet taste and the bitter taste each produced lower F2s (1422 and 1408 Hz, respectively) than the sour taste (1496 Hz), but these effects failed to survive Bonferroni correction (respective ts were -2.4 and -2.2 for sweet – sour, bitter – sour; all ps > 0.05; all dfs = 64).

3.3 *Voice discontinuity*
Sixty-five participants contributed data to this analysis. The mean voice discontinuity values selected for each taste at each concentration (where low values correspond to an uninterrupted vowel) are shown in figure 4. In voice discontinuity there was again a main effect of concentration ($F_1 = 10.5$, $p = 0.01$), with higher concentrations
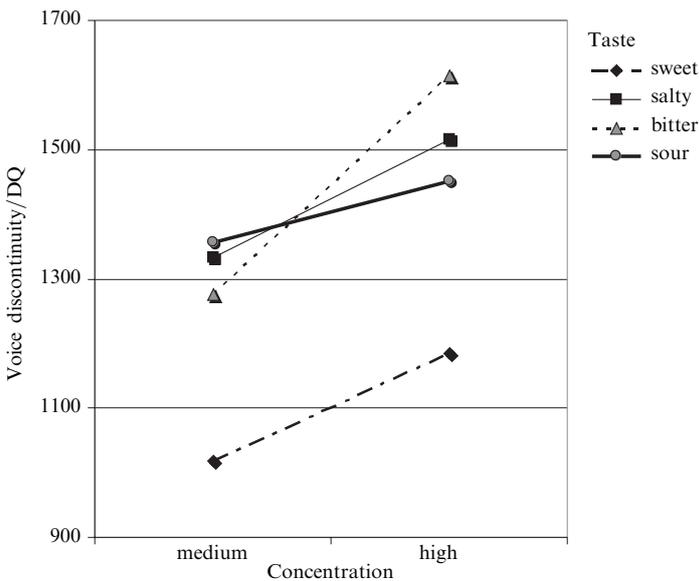

Figure 4. Mean values of voice discontinuity selected for four tastes (sweet, salt, bitter, sour) at two concentrations (medium, high). High DQ corresponds to discontinuous vowels.

rated as more discontinuous. There was no interaction of factors ($F_3 < 1$), but there was again a main effect of taste quality ($F_3 = 3.9$, $p = 0.01$). The sweet taste produced more continuous vowels (ie had lower discontinuity: 1098 DQ) than the bitter taste (1461 DQ; $t_{65} = -2.8$, $p = 0.04$) and the sour taste (1421 DQ; $t_{65} = -2.7$, $p = 0.05$). All other planned statistics were non-significant (all ts $< |2.4|$, all ps > 0.05).

3.4 Spectral balance

Sixty-four participants contributed to this analysis. The mean spectral balance, or overall acoustic energy, selected for each taste at each concentration (where low values correspond perceptually to low-pitch white noise) is shown in figure 5. In spectral balance there was again a main effect of concentration ($F_1 = 19.2$, $p < 0.001$) since more highly concentrated tastes were assigned sounds with higher-frequency spectral balance. There was also a main effect of taste quality ($F_3 = 3.2$, $p = 0.03$), since the sweet taste had a lower-frequency spectral balance overall (2050 Hz) than the sour taste (2652 Hz; $t_{63} = -3.1$, $p = 0.02$, corrected). All other comparisons were non-significant (all ts $< |2.2|$, all dfs $= 63$, all corrected ps > 0.05) and there was no interaction of effects ($F_3 = 1.0$, $p > 0.05$).
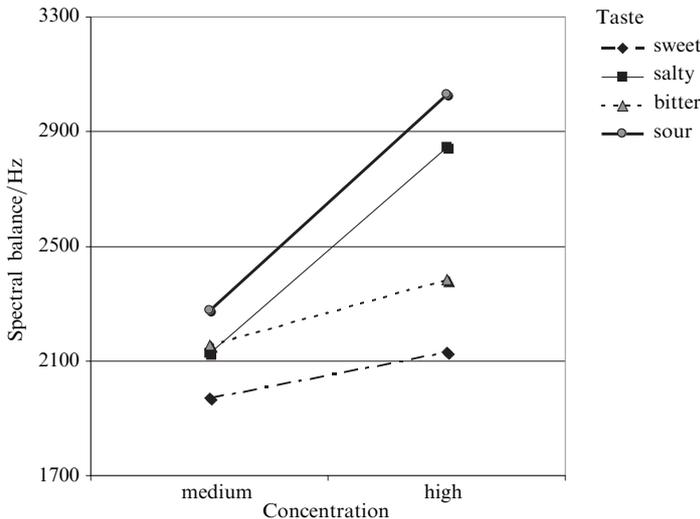

Figure 5. Mean values of spectral balance selected for four tastes (sweet, salt, bitter, sour) at two concentrations (medium, high). High spectral balance corresponds to high F1–F4 values.

4 Discussion

In our study we have examined the cross-modal mapping of gustation and audition. We manipulated qualities of taste and sound, and asked participants to make intuitive matches across these two sensory domains. Participants tasted drops of sweet, salty, bitter, and sour, with each tastant presented at medium and high concentrations. Participants selected their preferred sound for each taste from each of four auditory sliders that represented continua of F1, F2, voice discontinuity, and spectral balance, respectively. The first two continua present frequency changes in components of vowel quality, and the last in overall balance of the acoustic energy. The third represented a change in vowel continuity, from a smooth uninterrupted vowel to a staccato discontinuous one. We found that participants were systematic in their choice of taste–sound mapping, and that they showed the following patterns of response. Increasing concentrations of taste corresponded to increasing values in F1, F2, and spectral balance, and they also corresponded to more staccato vowel sounds (ie higher voice discontinuity). We also found differences across taste qualities. Sweet tastes were judged to be

low in frequency in F1, F2, and spectral balance. In the latter two sliders, they were judged lower than sour tastes, and in the first slider, they were judged lower than bitter, salty, and sour tastes. An additional tendency in F2 was for bitter to be rated lower than sour. In other words, these sliders revealed a sequence from sweet to bitter to sour (with salty numerically between bitter and sour). Finally, the sweet taste was judged to match smoother, more continuous vowel sounds than the bitter and sour tastes, which were judged to match more staccato sounds. The implications of these findings are discussed below.

Our findings provide support for Crisinel and Spence (2009) in showing systematic cross-modal associations between taste and sound. In their study Crisinel and Spence had used food names to represent tastes, and we aimed to test their claims using tastants in the mouth, without the mediation of target words and their associated linguistic qualities. Crisinel and Spence's food-name stimuli had varied in other, unplanned, ways, for example, in grapheme rounding and lexical markedness. Nonetheless, even when taste – sound mapping is directly tested, as we did here, cross-sensory correspondences again emerge. In presenting our findings, we have provided positive evidence of taste – sound mapping, but cannot rule out the possibility that word – sound mappings may also have contributed to the Crisinel and Spence pattern of data. In other words, it is yet possible that rounded graphemes do independently pair to lower pitch sounds, for example, and we find this highly plausible given similar sound – shape mappings reported elsewhere (eg Karwoski et al 1942; see also Cuskley et al 2009).

Our findings also support Crisinel and Spence's (2009) in other ways. These authors concluded that sour taste pairs with high-frequency sounds, and bitter taste pairs with low-frequency sound, and, indeed, our figures 2, 3, and 5 show a trend for sour to map to values higher than bitter in all manipulations where the unit was hertz. Furthermore, sour achieved consistently the highest hertz value across all four tastes and, for this reason, it seems well supported that sour tastes may map to sound characteristics with higher frequency. Nonetheless, our four-taste comparison allows us to modify slightly the suggestion of Crisinel and Spence that bitter may correspond to low frequency, per se. More correctly, Crisinel and Spence point out that their two-taste comparison can only show that bitter is mapped to a *lower* frequency than sour (as our findings in F2, for example, also suggest). Nonetheless, our inclusion of sweet and salty shows that bitter is not, in fact, judged to be low, per se. For example, bitter was judged to be significantly *higher* in frequency than sweet in F1 (and significantly higher in discontinuity). In other words, bitter is itself a relatively mid-to-high-mapped taste, and the truly low-mapped taste might best be considered as sweet.

Of particular interest is what may lie at the root of these unusual preferences. As a clue to their roots, we hypothesised, and subsequently found, a type of 'intensity matching', found also in other cross-modality studies (eg Smith and Sera 1992). Higher concentrations of taste were matched to higher values in all four continua, and this may represent an instantiation of what Walsh (2003) has described in his "A theory of magnitude" (ATOM). Walsh proposed a general mechanism responsible for quantitative (and spatial) processing which matches high magnitude across dimensions. This is a particularly appealing account for our spectral balance slider, for example, when over-all spectral energy was changing in magnitude from low to high (as a result of the low-pass filtering of the spectrum), and participants mapped this change to increasing concentration (another magnitude dimension). On other sliders, however, there was no such prothetic (magnitude) dimension, but, instead, a type of metathetic (qualitative) dimension. For example, our slider of vowel discontinuity had no clear end of high magnitude, and indeed the high end of this scale could be described as either more discontinuous or less continuous. However, this scale does offer the opportunity for

'hedonic matching'—a process in which two ends of a dimension subjectively correspond because both are pleasant or unpleasant (eg Dematte et al 2006a, 2006b). The sound produced at the highest end of this scale was a staccato pulsing, which participants spontaneously described as unpleasant. In contrast, the lowest end of this scale represents a more pleasant, smooth, continuous vowel, and it is therefore no surprise that the most pleasant taste—sweet (see Moskowitz et al 1974)—was placed lowest on this scale.

Evidence of systematic pairings across the senses in the general population can provide useful information, in and of itself, to theories of sensory integration and sensory processing. One interesting issue regarding sensory integration comes from a consideration of the population of individuals that experience cross-sensation to an extraordinary degree. Cross-modality of taste and sound can be felt explicitly in a condition known as synaesthesia. Synaesthetes experience two (or more) sensations when only one modality is stimulated. For example, in music–colour synaesthesia, listening to music triggers both an auditory experience of sound and a visual experience of colour. Synaesthete JW, for example, experiences a 370 Hz single piano note as a darkish off-yellow colour in the visual field (Ward et al 2006). These experiences have a neurological basis and affect at least 4% of the general population (Simner et al 2006). Of particular interest here is that some variants of synaesthesia involve the pairing of sound and taste. Synaesthete ES, for example, experiences taste and flavour in the mouth in response to music (Beeli et al 2005). ES experiences a bitter taste when hearing a major second interval, and experiences the flavour of cream when hearing a minor sixth interval. Of interest to the current study is that the experiences of synaesthetes are elsewhere known to reflect the implicit associations made by all people (eg see Simner 2009 for review). In this way, our evidence of cross-sensory mapping between taste and sound (and the similar findings of Crisinel and Spence 2009) provides yet another area in which synaestheses and non-synaesthetes may correspond. Put differently, the phenomenon of taste–sound synaesthesia makes it no surprise to find inherent taste–sound correspondences in all people.

Our findings might also have implications for traditional models of taste perception. Such models often rest on the idea that all tastes are reducible to a few basic categories, and that these are "distinctly different sensations" (Erickson 2008, page 63). Given that basic tastes are theoretically separate and assumed to be distinctly different, we might expect to find no similarities at all between them in the way they manifest themselves in sound sliders. Our significant main effects of taste show that our slider method can indeed capture differences between tastes, although the four basic tastes did not pattern entirely distinctly, each from the other, as the traditional model might predict. Instead, salt and sour generally patterned with each other (no significant differences in planned comparisons) and also tended to pattern together with bitter, while sweet patterned away from all three other tastes. This reflects other behavioural tasks, where participants often confuse sour/salt, bitter/sour, and occasionally salty/bitter, but where sweet is rarely involved in discrimination errors (O'Mahony et al 1979). Such findings suggest a hierarchical organisation in the four basic tastes, and is more compatible with Erickson's (2008) across-fibre patterning theory (see also Alba 2007;

[1] An anonymous reviewer has asked whether the cross-modal preferences we found for tastants might instead be attributed to their subjective intensities (ie in some way similar to the effect of objective intensity we found by manipulating concentration). We believe not. First, our tastants were taken from the Accusens Taste Kit© (Henkin 2005), which provides increments for each taste with reference to a baseline slightly above the average recognition threshold for each taste (Brosovic and McLaughlin 1989). In other words, the relative differences in perceptibility (and hence intensity) for each taste were controlled. Furthermore, participants did not report any subjective differences in intensity across tastes, and were often unable to detect any subjective change in

Jones et al 2006). In any event, our methods might provide an innovative way to critically approach measurements of taste distinctiveness.[1]

Finally, we consider how our results can be interpreted in light of the fact that we chose speech-related acoustic properties for our sound sliders. Our data, together with those of Crisinel and Spence (2009), show that people make cross-sensory association between taste and sound, and our own study has extended this to sounds associated with the qualities of speech. Our chosen methodology shows that these mappings are explicit, in the sense that they manifest themselves in tasks that rely on explicit judgments (even if participants may not be able to consciously verbalise their underlying 'rules'). Since humans map taste to sound within the acoustic boundaries of speech, it is possible that this process is exploited in object naming, and that it may even have played a role in the earliest stages of vocabulary construction. This in turn would have important implications for theories of language evolution. There is some evidence that cross-modal preferences may indeed underlie decisions in object naming. Klink (2000) presented participants with the fictional names of a food product (lemonade), and found that participants chose a name containing a front vowel /i/ (eg *bilad*) to be more bitter than the same name containing a back vowel (/o/, *bolad*). Our own findings suggest that these preferences may have arisen from inherent cross-modal mappings in F2 (the metric of vowel front/backness). Our participants found that increased bitterness, from higher concentrations compared to medium concentrations, was significantly tied to higher F2 values, and this corresponds to the association between bitterness and front vowels found by Klink (2000). Nonetheless, we found this same increase in F2 also in sweetness, saltiness, and sourness, suggesting that Klink's findings are better interpreted as showing that vowel frontness is generally associated with increases in *any* taste, rather than with bitterness per se.

We have suggested that taste – sound correspondences might have influenced the naming of words in language evolution, but the reverse is also possible. For example, perhaps when our participants experienced bitter tastants (for example) this triggered lexical access of the word *bitter*, and this in turn dictated the choices made in our sound-matching task. To examine this more closely, we compared the mean F1/F2 generated for each tastant (sweet, salty, bitter, sour) with the F1/F2 for each (stressed) vowel in the corresponding word (*sweet*, *salty*, *bitter*, *sour*; see figure 6). Overall the results were mixed: the F1 correlation is near-significant ($r = 0.93$, $n = 4$, $p = 0.07$) showing that the sounds attributed cross-modally to tastants indeed correlate with the sounds of taste names (ie the higher the F1 in the taste name, the higher the F1 mapped cross-modally). This initially suggests that implicit taste – sound mappings might mimic lexicalised terms for taste names (or vice versa). However, a similar analysis based on F2 presented contradictory results; again there was a near-significant correlation, although this time it was in the opposite direction to that predicted ($r = -0.92$, $n = 4$, $p = 0.08$). This weakens any suggestion that participants were influenced by taste names during our task.

Nonetheless, these findings are intriguing in suggesting at least some possible mediation of language, at least for F1. Taken at face value, this effect could have two roots: either participants were subconsciously vocalising taste names *during the task*,

[1] (continued)
concentration at all (ie they often believed that our targets were 8 different tastes, rather than a smaller subset of 4, across two concentrations). Finally, our findings across taste mimic those found elsewhere: sweet tended to pattern away from the other tastes in cross-modal matching, and this same hierarchy is found in behavioural tests of confusedness (O'Mahony et al 1979). Finally, our effects for taste reflect those of Crisinel and Spence (2009), who did not manipulate intensity at all. For these reasons we believe that we have correctly detected an effect of taste, although future studies might benefit from eliciting participants' subjective impressions of intensity for each tastant presented.
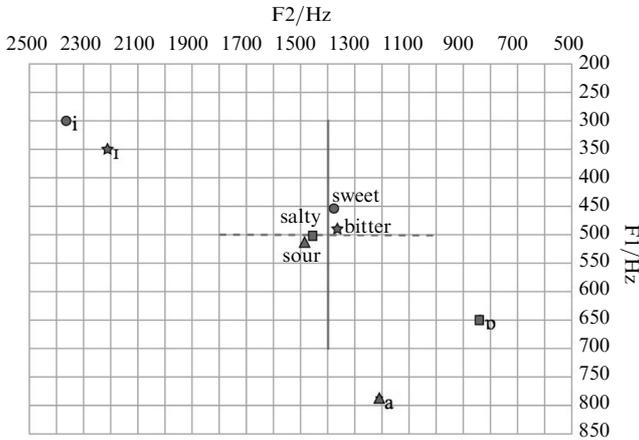
Figure 6. Mean F1 responses plotted against mean F2 responses for all four tastants (indicated by labelled shapes) at medium concentration. F1 and F2 values for the stressed vowels in the words *sweet*, *salty*, *bitter*, and *sour* (/i/, /ɒ/, /I/ and /a/, respectively) are also shown. The solid/dashed cross represents the possible search space (see figure 1).

or, instead, some pre-existing cross-modality may have played a role in the very naming of tastes during vocabulary construction. We tentatively propose the latter, and for the following reason: our participants were simply not able to consciously name the tastants during our task as *sweet*, *bitter*, *salty*, *sour*. Indeed, when lexical descriptions were provided, they showed a very strong tendency to be flavours not tastes (eg, respectively: *orangey*, *metallic*, *fishy*, *lemony*). For this reason, any mediation of language during our task would have focused on the highly activated flavour names, rather than the comparatively inactivated taste names.[2] In contrast, our findings are more compatible with an evolutionary account: taste – sound mappings might have influenced vocabulary construction for taste names, even if taste names were not activated during our task. An anonymous reviewer proposed an intriguing test of this hypothesis: if our cross-modal study were repeated with articulatory suppression (eg Stevenson and Oaten 2008), this would alter the pattern of our data only if they had been influenced by any self-generated verbal descriptors. If our results remained unchanged, our cross-modal mapping between test and sound might yet indicate pre-existing relationships that were brought to bear during vocabulary construction.

5 Conclusion

We have shown that people make systematic associations across taste and sound in terms of shared mapping between tastants in the mouth, and the acoustic qualities of accompanying auditory stimuli. We have shown that bitter does not map to low-frequency sounds, per se, but simply to lower-frequency sounds than sour, while bitter is in fact judged higher in frequency than sweet. These cross-sensory mappings are explicit in that they are found in explicit judgments of sensory pairings. Importantly, these associations can be seen in sound qualities related to speech, and this suggests that vocabulary construction in the evolution of language may have relied on these same cross-modal preferences, to pair names to foods in non-arbitrary ways.

---

[2] Even if taste names were activated only implicitly, dominant models of lexical access (eg Levelt 1989) make it unlikely they would have mediated our cross-modal finding. Words brought consciously to mind (here, flavour terms) are necessarily more highly activated than those that are not (here, taste terms), and they tend to suppress the latter. Again, then, any lexical mediation would almost certainly have focused on the highly activated flavour names.

References

Alba J W, 2007 "Taste perception: more than meets the tongue" *Journal of Consumer Research* **33** 490 – 498

Auvray M, Spence C, 2008 "The multisensory perception of flavor" *Consciousness and Cognition* **17** 1016 – 1031

Beeli G, Esslen M, Jänke L, 2005 "Synaesthesia: when coloured sounds taste sweet" *Nature* **434** 38

Bernstein I H, Edelstein B A, 1971 "Effect of some variations in auditory input upon visual choice reaction time" *Journal of Experimental Psychology* **87** 241 – 247

Beskow J, 2001 *Open-source Formant Synthesiser*, retrieved 20 April 2008, from http://www.speech.kth.se/wavesurfer/formant/. Associated with the Centre for Speech Technology KTH, Stockholm, Sweden

Brosovic G M, McLaughlin W, 1989 "Quality specific differences in human taste detection thresholds as a function of stimulus volume" *Physiology & Behavior* **45** 15 – 20

Bult J H F, Wijk R A de, Hummel T, 2007 "Investigations on multimodal sensory integration: Texture, taste, and ortho- and retro-nasal olfactory stimuli in concert" *Neuroscience Letters* **411** 6 – 10

Christensen C M, 1980 "Effects of taste quality and intensity on oral perception of viscosity" *Perception & Psychophysics* **28** 315 – 320

Christensen C M, Vickers Z M, 1981 "Relationships of chewing sounds to judgements of food crispness" *Journal of Food Science* **46** 574 – 578

Crisinel A, Spence C, 2009 "Implicit associations between basic tastes and pitch" *Neuroscience Letters* **464** 39 – 42

Curtin F, Schulz P, 1998 "Multiple correlations and Bonferroni's correction" *Biological Psychiatry* **44** 775 – 777

Cuskley C, Simner J, Kirby S, 2009 "Cross-modality bootstrapping protolanguage: reconsidering the evidence", poster presented at the European Human Behaviour and Evolution Conference, St Andrews, Scotland, UK, 6 – 9 April 2009

Davis R, 1961 "The fitness of names to drawings: A cross-cultural study in Tanganyika" *British Journal of Psychology* **52** 259 – 268

Dehaene S, Bossini A, Giraux P, 1993 "The mental representation of parity and numerical magnitude" *Journal of Experimental Psychology: General* **122** 371 – 396

Delwiche J, 2004 "The impact of perceptual interactions on perceived flavour" *Food Quality and Preference* **15** 137 – 146

Dematte M L, Sanabria D, Spence C, 2006a "Cross-modal associations between odors and colors" *Chemical Senses* **31** 531 – 538

Dematte M L, Sanabria D, Sugarman R, Spence C, 2006b "Cross-modal interactions between olfaction and touch" *Chemical Senses* **31** 291 – 300

Diffloth G, 1994 "i: big, a: small", in *Sound Symbolism* Eds L Hinton, J Nichols, J J Ohala (Cambridge: Cambridge University Press) pp 107 – 114

Djordjevic J, Zatorre R J, Jones-Gotman M, 2004 "Odor-induced changes in taste perception" *Experimental Brain Research* **159** 405 – 408

Doty R L, 2009 "The olfactory system and its disorders" *Seminars in Neurology* **29** 74 – 81

Erickson R, 2008 "A study of the science of taste: on the origins and influence of the core ideas" *Behavioral and Brain Sciences* **31** 59 – 75

Ganong W F, 2005 *Review of Medical Physiology* 22nd edition (Columbus, OH: McGraw-Hill)

Gevers W, Reynvoet B, Fias W, 2003 "The mental representation of ordinal sequences is spatially organized" *Cognition* **87** B87 – B95

Giegerich H, 2005 *English Phonology: An Introduction* (Cambridge: Cambridge University Press)

Gonzales A, 2009 "Intrinsic F0 in Shona vowels: A descriptive study", in *Selected Proceedings of the 39th Annual Congress on African Linguistics: Linguistic Research and Languages in Africa* Eds A Ojo, M Liobo (Somerville, MA: Cascadilla) pp 145 – 155

Hayes B, 2008 *Vowel Chart with Sound Files*, retrieved from http://www.linguistics.ulca.edu/people/hayes/103/charts/VChart/ (retrieved 25 September 2009)

Henkin I, 2005 *Accusens T Taste Function Kit*, obtained from the Taste and Smell Clinic, Washington, DC, USA, www.tasteandsmell.com. Distributed by Westport Pharmaceuticals Inc.

Hinton L, Nichols J, Ohala J J, 1994 "Introduction: sound-symbolic processes", in *Sound Symbolism* Eds L Hinton, J Nichols, J J Ohala (Cambridge: Cambridge University Press) pp 1 – 14

Holt-Hansen K, 1968 "Taste and pitch" *Perceptual and Motor Skills* **27** 59 – 68

Holt-Hansen K, 1976 "Extraordinary experiences during cross-modal perception" *Perceptual and Motor Skills* **43** 1023 – 1027

Jones L M, Fontanini A, Katz D B, 2006 "Gustatory processing: a dynamic systems approach" *Current Opinion in Neurobiology* **16** 420 – 428

Karwoski T F, Odbert H S, Osgood C E, 1942 "Studies in synesthetic thinking: II. The role of form in visual responses to music" *Journal of General Psychology* **26** 199 – 222

Klink R R, 2000 "Creating brand names with meaning: The use of sound symbolism" *Marketing Letters* **11** 5 – 20

Kohler W, 1929 *Gestalt Psychology* (New York: Liveright)

Kohler W, 1947 *Gestalt Psychology* 2nd edition (New York: Liveright)

Kucera H, Francis W N, 1967 *Computational Analysis of Present-day American English* (Providence, RI: Brown University Press)

Ladefoged P, 1993 *A Course in Phonetics* 3rd edition (New York: HBJ)

Levelt W J M, 1989 *Speaking: From Intention to Articulation* (Cambridge, MA: MIT Press)

Marks L E, Hammeal R J, Bornstein M H, 1987 "Perceiving similarity and comprehending metaphor" *Monographs of the Society for Research in Child Development* **215**(52) 1 – 100

Maurer D, Pathman T, Mondloch C J, 2006 "The shape of boubas: Sound – shape correspondences in toddlers and adults" *Developmental Science* **9** 316 – 322

Moskowitz H R, Kluter R A, Westerling J, Jacobs H L, 1974 "Sugar sweetness and pleasantness: evidence for different psychological laws" *Science* **184** 583 – 585

Newman S S, 1933 "Further experiments in phonetic symbolism" *American Journal of Psychology* **45** 53 – 75

Nicholls M E R, Bradshaw J L, Mattingley J B, 1999 "Free-viewing perceptual asymmetries for the judgement of brightness, numerosity and size" *Neuropsychologia* **37** 307 – 314

Ohala J J, 1984 "An ethological perspective on common cross-language utilization of F0 of voice" *Phonetica* **41** 1 – 16

Ohala J J, 1994 "The frequency code underlies the sound-symbolic use of voice pitch", in *Sound Symbolism* Eds L Hinton, J Nichols, J J Ohala (Cambridge: Cambridge University Press) pp 325 – 347

O'Mahony M, Goldenberg M, Stedmon J, Alford J, 1979 "Confusion in the use of the taste adjectives 'scan' and 'bitter'" *Chemical Senses* **4** 301 – 318

Pape D, Mooshammer C, 2008 "Intrinsic pitch is not a universal phenomenon: Evidence from Romance languages", in *Proceedings of the 11th Labphon (Laboratory Phonology)* University of Wellington, New Zealand, available at http://phonetics.danielpape.info/paper/pape_lap.pdf (retrieved 25 September 2009)

Parise C, Spence C, 2009 "When birds of a feather flock together: Synaesthetic correspondences modulate audiovisual integration in non-synaesthetes" *PLoS One* **4**(5) e5664

Pfitzinger H R, 2003 "Acoustic correlates of the IPA vowel diagram" *Proceedings of the International Congress of Phonetic Science* **2** 1441 – 1444

Ramachandran V S, Hubbard E M, 2001 "Synaesthesia: a window into perception, thought and language" *Journal of Consciousness Studies* **8** 3 – 34

Ramachandran V S, Hubbard E M, 2005 "Synaesthesia: a window into the hard problem of consciousness", in *Synesthesia: Perspectives from Cognitive Neuroscience* Eds L C Robertson, N Sagiv (Oxford: Oxford University Press) pp 127 – 189

Sapir E, 1929 "A study in phonetic symbolism" *Journal of Experimental Psychology* **12** 225 – 239

Simner J, 2009 "Synaesthetic visuo-spatial forms: Viewing sequences in space" *Cortex* **45** 1138 – 1147

Simner J, Ludwig V, 2009 "What colour does that feel? Cross-modal correspondences from touch to colour", paper presented at Third International Conference of Synaesthesia and Art, 26 – 29 April 2009 (Artecitta), Granada, Spain

Simner J, Mulvenna C, Sagiv N, Tsakanikos E, Witherby S, Fraser C, Scott K, Ward J, 2006 "Synaesthesia: The prevalence of atypical cross-modal experiences" *Perception* **35** 1024 – 1033

Simner J, Ward J, Lanz M, Jansari A, Noonan K, Glover L, Oakley D, 2005 "Non-random associations of graphemes to colours in synaesthetic and normal populations" *Cognitive Neuropsychology* **22** 1069 – 1085

Smilek D, Carriere J S A, Dixon M J, Merikle P M, 2007 "Grapheme frequency and color luminance in grapheme – color synaesthesia" *Psychological Science* **18** 793 – 795; doi:10.1111/j.1467-9280.2007.01981.x

Smith L B, Sera M D, 1992 "A developmental analysis of the polar structure of dimensions" *Cognitive Psychology* **24** 99 – 142

Spence C, Zampini M, 2006 "Auditory contributions to multisensory product perception" *Acta Acustica United With Acustica* **92** 1009 – 1025

Stevenson R J, Oaten M, 2008 "The effect of appropriate and inappropriate stimulus color on odor discrimination" *Perception & Psychophysics* **70** 640 – 646

Stevenson R J, Prescott J, Boakes R A, 1999 "Confusing tastes and smells: how odours can influence the perception of sweet and sour tastes" *Chemical Senses* **24** 627 – 635

Traunmuller H, Ericksson A, Menard L, 2003 "Perception of speaker age, sex and vowel quality investigated using stimuli produced with an articulatory model" *Proceedings of the International Congress of Phonetic Science* **2** 1739 – 1742

Vennemann M M, Hummel T, Berger K, 2008 "The association between smoking and smell and taste impairment in the general population" *Chemical Senses* **225** 1121 – 1126

Verhagen J V, Engelen L, 2006 "The neurocognitive bases of human multimodal food perception: Sensory integration" *Neuroscience and Biobehavioral Reviews* **30** 613 – 650

Walsh V, 2003 "A theory of magnitude: common cortical metrics of time, space and quantity" *Trends in Cognitive Sciences* **7** 483 – 488; doi:10.1016/j.tics.2003.09.002

Ward J, Huckstep B, Tsakanikos E, 2006 "Sound – colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all" *Cortex* **42** 264 – 280

Zampini M, Spence C, 2004 "The role of auditory cues in modulating the perceived crispness and staleness of potato chips" *Journal of Sensory Studies* **19** 347 – 363; doi:10.1111/j.1745-459x.2004.080403.x

Zampini M, Spence C, 2005 "Modifying the multisensory perception of a carbonated beverage using auditory cues" *Food Quality and Preference* **16** 632 – 641

Zampini M, Sanabria D, Phillips N, Spence C, 2007 "The multisensory perception of flavor: Assessing the influence of color cues on flavor discrimination responses" *Food Quality and Preference* **18** 975 – 984