



ORIGINAL ARTICLE

Crossmodal interference between language and flavour[☆]



Eugenia Razumiejczyk^{a,b,*}, Carolina Pereyra Girardi^{a,c},
María del Carme Crivello^{a,b}, Mauro Fioramonti^{a,b}, Guillermo Macbeth^{a,b},
Fernando Marmolejo-Ramos^d

^a Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

^b Facultad de Ciencias de la Educación, Universidad Nacional de Entre Ríos, Argentina

^c Instituto de Investigaciones Cardiológicas ININCA, Universidad de Buenos Aires, Argentina

^d Gösta Ekman Laboratory, Department of Psychology, Stockholm University, Sweden

Received 5 March 2015; accepted 23 January 2016

Available online 15 December 2016

KEYWORDS

Stroop;
Flavour;
Vision;
Hearing;
Crossmodality

Abstract This work describes crossmodal Stroop interference in flavoural, visual, and auditory representations. A mixed design was used with two randomized groups. As a between-subjects factor, words were presented in visual (group 1) or auditory (group 2) forms. Stimulus congruency (congruent, incongruent, and control) was defined as a within-subjects factor. Reaction times and the number of correct answers were recorded. The results showed a minor crossmodal Stroop interference in conditions of congruency. In contrast, with incongruent and control stimuli, reaction times increased and accuracy rates diminished in both experimental groups. Data from the two groups were compared, and it was concluded that the interference was greater when the distractor was written than when it was spoken. These results are discussed in terms of the difficulty of visual linguistic representation and in relation to previous studies. © 2016 Fundación Universitaria Konrad Lorenz. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

PALABRAS CLAVE

Stroop;
Sabor;
Visión;
Audición;
Crossmodal

Interferencia intermodal entre lenguaje y sabor

Resumen El objetivo del presente trabajo es describir la interferencia del Stroop crossmodal en representaciones del sabor, visuales y auditivas. Es un diseño mixto con dos grupos seleccionados al azar. Los estímulos fueron palabras presentadas de forma visual (grupo 1) o auditiva (grupo 2). La congruencia de estímulo (congruentes, incongruentes y control) se definió como

[☆] Best Article of the Issue Award.

* Corresponding author.

E-mail address: eugeniaraz@hotmail.com (E. Razumiejczyk).

un factor intrasujetos. Se registraron los tiempos de reacción y el número de respuestas correctas. Los resultados mostraron una interferencia Stroop crossmodal menor en condiciones de congruencia. Mientras con estímulos incongruentes y de control los tiempos de reacción se incrementaron y las tasas de precisión disminuyeron en ambos grupos experimentales. Los datos de los dos grupos se compararon y se concluyó que cuando el distractor era escrito la interferencia fue mayor que cuando era hablado. Estos resultados se discuten en términos de la dificultad de representación lingüística visual y en relación con estudios anteriores.

© 2016 Fundación Universitaria Konrad Lorenz. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Eating is an activity that integrates all the sensory modalities. Taste is activated when the chemical stimuli reach the taste buds in the mouth. Such an environment involves saliva (Matlin & Foley, 1996), retronasal smell (Sun & Halpern, 2005) and tactile information (Purves et al., 2004). Flavour results from the combination of these three modalities: taste, smell, and tactile processes. Eating is concerned with flavour, but also with vision and audition. Orthonasal olfaction and vision provide information prior to the eating activity; that is, they are concerned with distal stimuli. The other perceptual modalities provide proximal information (Shankar, Levitan, & Spence, 2010); that is, audition, retronasal olfaction, taste, and tactile processing operate when the food is inside the mouth. In this sense, all the modalities contribute to the eating activity and can also anticipate information about what will be eaten (Yamada, Sasaki, Kunieda, & Wada, 2014). Before perceiving the taste of a food, we can see it and predict its flavour (Spence, Levitan, Shankar, & Zampini, 2010). In a similar manner, we can hear what we are eating. For instance, we can hear the sound of a crispy food and infer how moist that food is (Spence & Shankar, 2010). In sum, all this information is integrated during the eating behaviour. Thus sensory information available at the time we are eating enables us to form expectations about what we are about to eat or are eating (Piqueras-Friszman & Spence, 2015). The more familiar we are with a food product, the more likely it is that our eating expectations are going to be correct (Ludden, Schifferstein, & Hekkert, 2009).

White and Prescott (2007) studied interference of cross-modal perception between taste and smell. The integration of both perceptions appears to arise from their co-exposure (Prescott, 2015). In their experiment, they asked participants to identify a gustatory stimulus while simultaneously administering an olfactory stimulus. The results showed that the identification of gustatory stimuli was facilitated in the condition of congruence (White & Prescott, 2007). Neuroimaging studies have shown that retronasal perceptions inside the mouth activate the primary somatosensory cortex where touch information is processed, while orthonasal perceptions do not (Small, Gerber, & Hummel, 2005). Consequently, these researchers emphasized the adaptive significance of the results due to the biological importance of rapid and accurate discrimination of nutritive and toxic compounds. These data are consistent with those of Morrot, Brochet and Dubourdieu (2001), who studied the link between visual and olfactory representations. They found

that experienced enologists committed numerous errors when tasting coloured white wine. Similarly, Stevenson and Boakes (2004) showed the influence of odour on the perception of how sweet a drink is. Similar results were obtained by Velasco, Jones, King and Spence (2013) in relation to whiskey and the odour in the environment in which it was drunk.

Crossmodal researchers have identified many interference phenomena concerned with such integration through the manipulation of the information provided by the different modalities. It has been extensively concluded that the working memory plays a critical role in these complex processes. Working memory is the memory system that processes information temporarily (Baddeley, 1992) and makes it possible to integrate information that is received at different times and from different modalities (Baddeley, 1992). Working memory must be highly adaptive. This is because the diverse sources of sensory modality data need to be combined to accurately provide information about the external environment (Driver & Spence, 2000; Orhan, Sims, Jacobs, & Knill, 2014). Cases of crossmodal integration include the McGurk effect (McGurk & MacDonald, 1976), which is concerned with the combination of incongruent information from visual and hearing modalities (Massaro, 1999). Cross-modal correspondences between colour and flavour seem to fulfil an evolutionary purpose. For example, the red colour and sweet taste of apples are signals of high nutritional value (Hoegg & Alba, 2007; Spence et al., 2010). However, the relation of hearing to taste has yet to receive a satisfactory explanation (Knöferle & Spence, 2012).

The cognitive system in general and the working memory in particular have certain limitations that have been demonstrated in the laboratory. For instance, subjects are not always capable of processing two sources simultaneously (Baddeley, 1992; Stroop, 1935). Stroop conducted a study in which participants were presented with words that referred to colours. Those words were written in colours that either did or did not match the referent. Participants were instructed to name the colour in which the word was written in the shortest time possible. Results showed higher accuracy rates and shorter reaction times when the stimuli were congruent (MacLeod, 1991). In contrast, when the stimuli were incongruent, participants evidenced significantly longer reaction times in their responses. Congruence refers to the coherence between the colour of the stimulus and the meaning of the word. The results are explained by the automaticity hypothesis, which states that reading a word is a more automated process than naming the colour

of the ink in which it is written (Posner & Dehaene, 1994). When the stimuli were incongruent, the processing load for the working memory was more arduous and favoured the automatic process of reading the word over processing the colour of the word.

Nonetheless, several authors interpreted the Stroop phenomenon as a selective attention process (Lamers & Roelofs, 2007). Thus it was interpreted that the effect of the task was to produce attentional competition between stimuli (Kahneman & Chajczyk, 1983; Kim, Cho, Yamaguchi & Proctor, 2008), i.e., between the colour of the word and the linguistic distractor. The Stroop test requires participants to inhibit distractors in order to process the information requested by the task (Kirn, Kirn & Chun, 2005). Stroop interference is observed when the executive function of attention fails due to distractors. Thereby generating a longer reaction time and/or more errors.

Various modifications have been introduced to the original Stroop task. McCown and Arnoult (1981) modified that task (1935) by presenting subjects with the word in two experimental conditions: (a) vertical versus horizontal, and (b) all the letters coloured versus only the first three letters coloured. Interference equivalents were found in all cases. Regan (1978) administered words with the first letter in a congruent or incongruent colour. In order to examine the integration of the stimulus, Dyer (1973) placed the word written in black on one side and a spot of colour on the other side of a fixation point. In a similar manner, Kahneman and Chajczyk (1983) placed the word above or below a colour spot. These studies found significant interference.

The Stroop effect has been studied in crossmodal situations (Rolls, 2004). Razumiejczyk, Macbeth and Adrover (2011) studied interference between flavour, visual pictorial, and linguistic perceptions. In their experiment, participants were asked to identify a gustatory stimulus while simultaneously being administered a visual stimulus (spoken word or picture). In accordance with the presentation of this pair of stimuli, three conditions were generated: (a) congruence: the visual stimulus and the gustatory stimulus belonged to the same object; (b) incongruence: the visual stimulus and the gustatory stimulus did not belong to the same edible object; and (c) control: the visual stimulus was an inedible object. Their results showed that the interference pattern was similar in pictorial and in linguistic distractors, i.e. in both cases greater levels of crossmodal Stroop interference were observed with incongruent and control stimuli, while the level of congruent stimuli worked as a facilitator in identifying gustative stimuli. Additionally, interference between pictorial and linguistic representations were compared. The results showed that words functioned as more of a distractor than pictures, and that words needed more processing time for the identification of the flavour stimulus. It was concluded that crossmodal Stroop interference between flavour and vision is higher with stimuli such as words than with pictures because the attentional demand of the former is higher. This suggests that the habitual processing of visual linguistic representations is part of the individual's adjustment to the environment and that words are stimuli present in the environment and in everyday human life.

We argue that linguistic representations contribute to flavour perception. Food is often associated with a written

label or a verbal expression. Therefore, the processing of such linguistic information might be relevant for the production of expectations before eating. Language may play a critical role in the crossmodal integration of flavour, visual, and auditory modalities. For example, Shankar, Levitan, Prescott and Spence (2009) found that the colour and label of chocolate packaging modulate the rating of flavour, as they generate expectations. Razumiejczyk, Macbeth and Leibovich de Figueroa (2013) found a facilitation effect in the identification of flavour stimuli when incomplete words were presented for the congruent condition. They concluded that congruent flavour stimuli performed a completion of the incomplete flavour linguistic representations through crossmodal integration. The effect was observed in the congruent condition when compared to the incongruent and control conditions. Razumiejczyk, Macbeth, Marmolejo-Ramos and Noguchi (2015) obtained similar results when comparing the Stroop interference between flavour and incomplete words and anagrams.

In sum, the understanding of crossmodal phenomena concerned with flavour, vision, and language processing has evolutionary implications. The study of such phenomena may contribute to the comprehension of adaptive behaviour in our actual complex environments. In the context of these previous findings, we argue that the study of crossmodal Stroop interference between flavour and visual and auditory linguistic representations is relevant for the comprehension of sensory modality integration in working memory. The present study aims to describe crossmodal Stroop interference between flavour, visual linguistic, and auditory linguistic representations and to compare crossmodal Stroop interference between these modalities. Our working hypothesis states that the crossmodal integration of flavour and written words is harder to process than the crossmodal integration of flavour and spoken words. This phenomenon might be observed through different degrees of interference between visual and auditive modalities.

Method

Participants

The total sample consisted of sixty-nine Argentine university students, 51 females and 18 males, whose mean age was 24.14 years ($SD=7.407$). As in previous studies (Razumiejczyk et al., 2011), the criteria for inclusion in the sample were that participants were between 20 and 40 years old (West, 2004), were non-smokers, and had not ingested any food or drink other than water during the three hours prior to the experiment. Participants were volunteers, who gave their written, informed consent. Each participant was randomly assigned to the experimental group's visual condition, G1 ($n=26$), or auditory condition, G2 ($n=43$). The study was approved by the ethics committee of The National University of Entre Ríos.

Materials

Flavour stimuli were administered in the form of peach, plum, strawberry, and orange puree liquefied at room temperature. The characteristics of these stimuli were as found

Table 1 Stimuli used in the experiment.

Condition	Flavour stimuli	Words
Congruent	Frutilla	Frutilla
	Durazno	Durazno
	Ciruela	Ciruela
	Naranja	Naranja
Incongruent	Frutilla	Banana
	Durazno	Ananá
	Ciruela	Sandía
	Naranja	Manzana
Control	Frutilla	Moneda
	Durazno	Anteojos
	Ciruela	Bicicleta
	Naranja	Reloj

Note. Congruent condition words: durazno = peach, ciruela = plum, naranja = orange, frutilla = strawberry. Incongruent condition words: banana = banana, ananá = pineapple, sandía = watermelon, manzana = apple. Control condition words: moneda = coin; anteojos = spectacles or glasses, bicicleta = bicycle, reloj = clock or watch.

in their natural state, so that the study would respect the ecological relationship between the individual and their environment.

A preliminary study (Razumiejczyk et al., 2010) had been conducted in order to select these four stimuli were selected from a group of seven. Cronbach's α for the identification of these stimuli was .536, suggesting adequate psychometric homogeneity (Fleiss, Levin, & Paik, 2004). Regarding linguistic stimuli, for G1 (visual), the words were printed in black on the white background of a computer screen. In G2 (auditory), the words were presented aurally off-line. The words used in the experiment are presented in Table 1.

All the utensils (spoons, spectacles, and hygiene items) were discarded after being used by the participants.

Procedure

The procedure was carried out by two volunteer experimenters, who were unaware of the purpose of the study. The paradigm of crossmodal Stroop-like task used was programmed in PsychoPy (Peirce, 2007). For participants in group G1 (visual), the words were presented on a computer screen. In G2 (auditory), the words were presented aurally and the relationship between the flavour and language stimuli determined the three levels of congruency (congruent, incongruent, and control stimuli; see Design and Analyses section). In order to respect the individual participant's ecological adjustment to the experiment's environment, the auditory stimuli were presented verbally by an experimenter, who said each word aloud in each trial. The trials followed a sequence. First, a sound signalled the start of each trial. Next, the participant tasted the flavour while simultaneously listening to or reading the word for that flavour. Then the participant orally informed the experimenter what flavour was being tasted. Finally, the experimenter recorded the participant's answer. An example of the pattern for G1 would be: commencing sound; FRUTILLA-strawberry_{flavour} ||

FRUTILLA_{written word} or aural word for that flavour. An example for G2 would be: commencing sound; FRUTILLA-strawberry_{flavour} || FRUTILLA_{aural word}.

In each trial, the experimenter recorded the participant's responses by pressing a computer key. Response times (RTs) and the accuracy rates (i.e. number of correct answers) were obtained from these records. This procedure was applied after analysing pilot results in which participants tended to press the response key several seconds before recognizing the flavour stimuli. A pair of spectacles that prevented the observation of the flavour stimulus was used in order to avoid the incidental processing of colour (Morrot et al., 2001). In this way, we tried to avoid visual processing prior to the consumption of the flavour stimulus as the eye would give an exteroceptive cue and thus produce expectations about that stimulus (Piqueras-Friszman & Spence, 2015; Spence, 2012).

The pairs of flavour and linguistic stimuli were administered to each participant in random order. Before each trial, participants were instructed to perform a dental cleaning with water. The experiment was conducted at the same time of the day for each participant and lasted seven minutes approximately (see Fig. 1).

Design and analyses

A mixed experimental design with two randomized groups (Kuehl, 1999) was used to perform within-subject and between-subject comparisons. The between-subjects factor was the presentation of words in visual or auditory forms, i.e. visual stimuli were presented to group 1 (G1 – visual) and auditory stimuli to group 2 (G2 – auditory). The within-subjects factor was the congruence between linguistic and flavour stimuli. For this independent variable, three levels were set: (a) congruent stimuli: the flavour and linguistic stimuli are congruent when the word presented in visual or auditory form matches the name of the flavour stimulus; (b) incongruent stimuli: the flavour and language stimuli are inconsistent when the word presented in visual or auditory form does not match the name of the flavour stimulus but it refers to a fruit; and (c) control stimuli: flavour and language stimuli do not coincide and the word presented visually or aurally does not represent any food-like item.

The dependent variables were the RT and the accuracy rates. The response time was estimated as the time lapse, in seconds to the nearest hundredth, between the presentation of the stimuli and the verbal response of the volunteer. The accuracy rate was operationalized as the proportion of the participant's correct responses in identifying the flavour stimulus. The experimental design is presented in Fig. 2.

Normality and homogeneity tests were applied in order to decide the adequacy of using parametric or non-parametric tests. The Stouffer method was used to combine the p -values of the Anderson–Darling, Shapiro–Wilk and Shapiro–Francia normality tests. The homogeneity of variances between independent samples was estimated via the modified robust Brown-Forsythe Levene-type test based on the absolute deviations from the median. The homogeneity of variances between dependent samples was estimated via robust

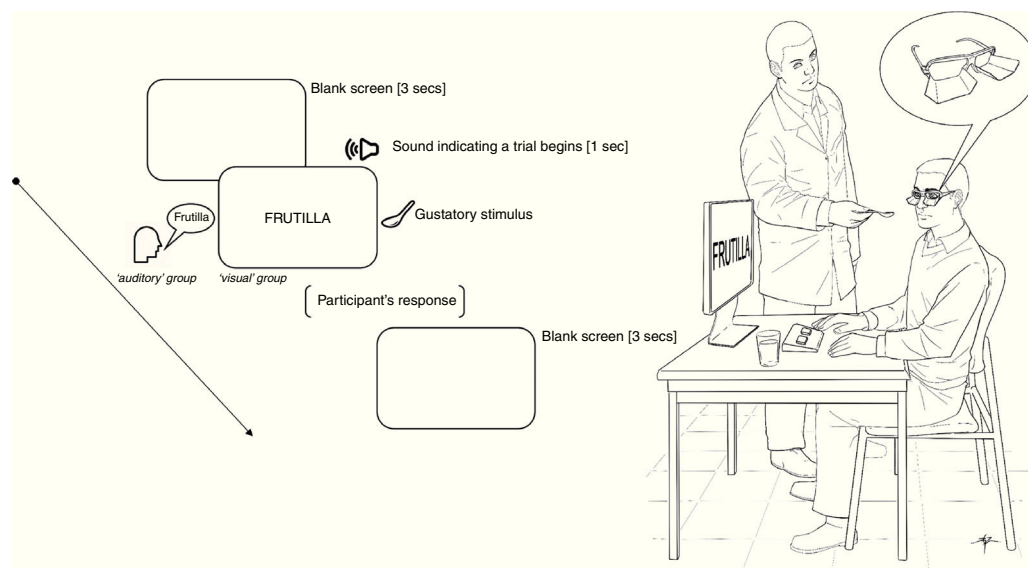


Figure 1 Illustration of the experiment's sequence of events and set-up.

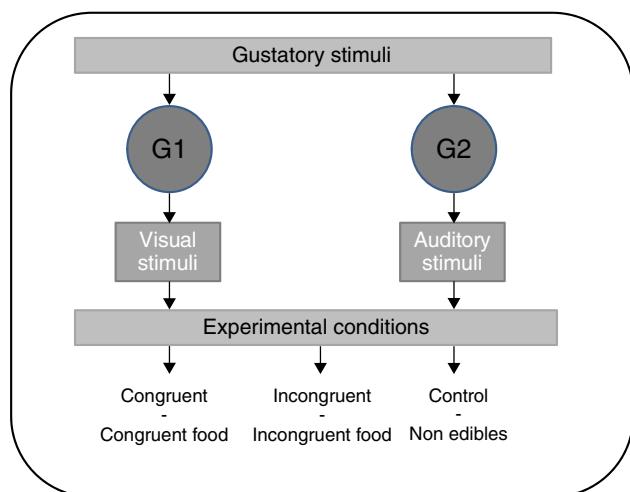


Figure 2 Experimental design.

one-way repeated measures ANOVA on means. This test was applied on the Levene-transformed scores for the dependent variable of the within-subjects variables being compared. The Levene transformation for each data set was performed by estimating the median for the data set and then calculating the absolute difference between each observation and the estimated median (i.e. $|x_i - \text{Mdn}(X)|$; such that x_i represents an observation in the data set X). The results of the normality and homogeneity tests are shown in Table 2. These results suggested the use of robust, nonparametric tests.

Average ratings and accuracy rates were submitted to a robust analysis of variance, ANOVA-type statistic (F_{ATS}) with the between-subjects factors of gender and experimental group (vision and audition), and the within-subjects factor of congruence level (congruent, incongruent, and control). Gender differences were evaluated in order to treat the sample as representative of the population. The R function 'f2.ld.f1' implements this test and can be found in the

package 'nparLD'. (Details of this analysis can be found in Noguchi, Gel, Brunner, & Konietschke, 2012. See Marmolejo-Ramos, Elosúa, Yamada, Hamm, & Noguchi, 2013, for an application of this test.)

One-way comparisons of $2 \leq$ independent samples were performed via the modified Cucconi permutation test (Marozzi, 2012, 2014), MC . One-way comparisons of $2 \leq$ dependent samples were performed via the Agresti-Pendergast test, F_{AP} , (Wilcox, 2005; via the R function 'apanova'). Multiple, pairwise comparisons of independent and dependent samples were submitted to the sequentially rejective (Ψ_{SR}) and the percentile bootstrap methods (Ψ_{PB}), respectively (Wilcox, 2005; via the R functions 'pbmcp' and 'rmmcppb'). Both methods adjust the p -values by using the sequentially rejective method (Wilcox, 2005, p. 313). This is achieved by setting the option 'bhop = FALSE' in 'pbmcp' and 'hoch = FALSE' in 'rmmcppb'.

The distributions of the results were displayed using beanplots (Kampstra, 2008). In this graphical display, the data's probability density function is shown along with the actual observations. The observations are represented via thin horizontal ticks; the data's average is represented via a thick horizontal line.

Results

The Harrell-Davis estimator of the median, a robust measure of central tendency resistant to outliers (Vélez & Correa, 2014), was used to estimate average response times for each participant in each condition. Accuracy rates for each participant were estimated as described above. The distributions of the results are shown in Fig. 3.

Response times

The results of the ANOVA-type statistic showed a main effect of experimental group ($F_{ATS}(1, 41.11) = 94.11, p < .001$)

Table 2 Results of the normality and homogeneity tests for the response time and accuracy rate data.

Flavour Word pairing condition	Normality test		Homogeneity test (independent samples)
	Vision group	Audition group	
Congruent _{RT}	$p_{Stouffer} = 0.401$	$p_{Stouffer} = 0.365$	$L = 13.00, p < .001$
Incongruent _{RT}	$p_{Stouffer} = 0.907$	$p_{Stouffer} = 0.652$	$L = 9.87, p < .003$
Control _{RT}	$p_{Stouffer} = 0.927$	$p_{Stouffer} = 0.015$	$L = 8.09, p < .006$
<i>Homogeneity test</i> (dependent samples) _{RT}	$F(2, 50) = 0.76, p = 0.47$	$F(1.81, 76.02) = 2.30, p = 0.11$	
Congruent _{AR}	$p_{Stouffer} = 3.58e^{-12}$	$p_{Stouffer} = 2.13e^{-10}$	$L = 8.73, p < .005$
Incongruent _{AR}	$p_{Stouffer} < .001$	$p_{Stouffer} = 2.36e^{-6}$	$L = 0.18, p = 0.67$
Control _{AR}	$p_{Stouffer} = 2.29e^{-9}$	$p_{Stouffer} = 1.52e^{-9}$	$L = 1.16, p = 0.28$
<i>Homogeneity test</i> (dependent samples) _{AR}	$F(1.93, 48.48) = 1.57, p = 0.21$	$F(2, 84) = 0.78, p = 0.45$	

Note. RT = response times, AR = accuracy rates, $p_{Stouffer}$ = combined p -value of three normality tests via the Stouffer method, and L = results of the robust Brown-Forsythe Levene-type test. Significant results are shaded.

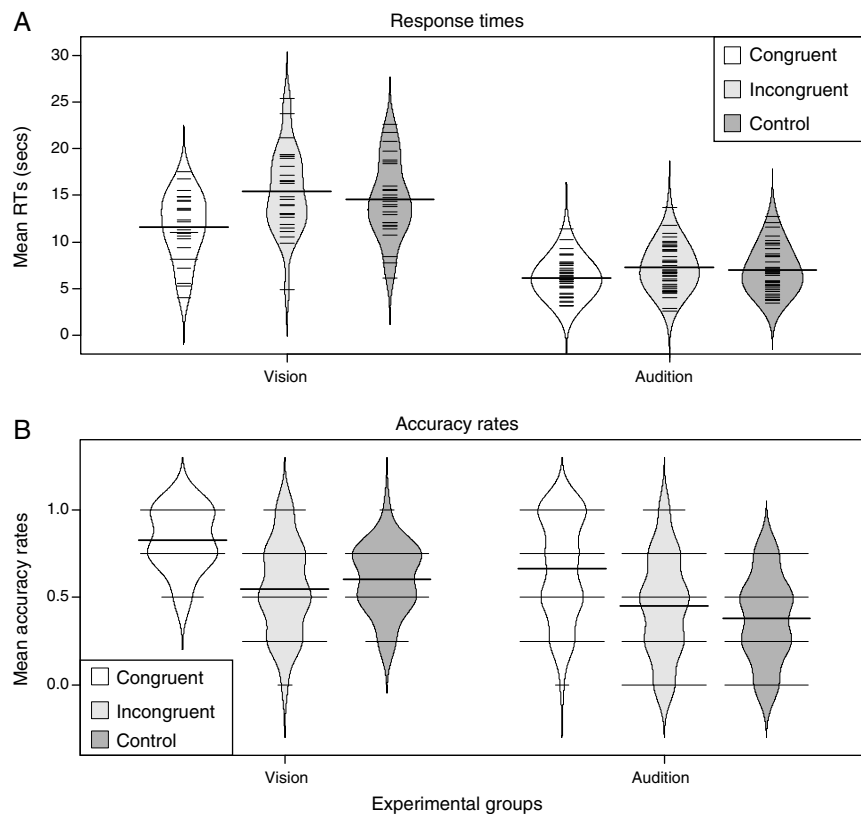


Figure 3 Beanplots representing the distribution of response times (A) and accuracy rates (B).

in that participants in G2 (audition) responded faster than participants in G1 (vision): $M_{G2} = 6.79$, $SD = 2.30$ and $M_{G1} = 13.83$, $SD = 4.41$, respectively ($MC = 59.36$, $p < .001$). There was also a main effect of congruence level ($F_{ATS} (1.80, \infty) = 19.52$, $p < .001$) in that participants responded fastest to congruent items ($M_{cong} = 8.18$, $SD = 3.75$), followed by RTs to control items ($M_{cont} = 9.84$, $SD = 4.87$), and RTs to incongruent items ($M_{incong} = 10.30$, $SD = 5.20$) irrespective of the experimental group. The Agresti-Pendergast one-way test corroborated the differences among the levels of congruence ($F_{AP} (2, 136) = 15.84$, $p < .001$). Multiple pairwise comparisons showed that while there were differences between RTs to congruent and incongruent items ($\psi_{PB} = -2.13$, $p < .001$) and congruent and control items ($\psi_{PB} = -1.66$, $p < .001$), differences between RTs to incongruent and control items were not significant ($\psi_{PB} = 0.46$, $p = .27$). Finally, there was no main effect of gender ($F_{ATS} (1, 41.11) = 3.01$, $p = .082$); and the three 2-way and the one 3-way interactions did not reach significance (all $p > .2$).

Accuracy rates

The results of the ANOVA-type statistic showed a main effect of experimental group ($F_{ATS} (1, 25.42) = 8.48$, $p = .007$) in that G1 (vision) participants were more accurate in their responses than G2 (audition) participants: $M_{G1} = 0.66$, $SD = 0.24$ and $M_{G2} = 0.49$, $SD = 0.29$, respectively ($MC = 8.54$, $p < .001$). There was also a main effect of congruence level ($F_{ATS} (1.92, \infty) = 23.58$, $p < .001$). Participants were more accurate in responding to congruent items ($M_{cong} = 0.72$, $SD = 0.26$), followed by incongruent ($M_{cont} = 0.48$, $SD = 0.27$), and control items ($M_{incong} = 0.46$, $SD = 0.24$) irrespective of the experimental group. A one-way test corroborated the differences among the levels of congruence ($F_{AP} (2, 136) = 28.36$, $p < .001$). Multiple pairwise comparisons showed that while there were differences between accuracy rates to congruent and incongruent items ($\psi_{PB} = 0.24$, $p < .001$) and congruent and control items ($\psi_{PB} = 0.26$, $p < .001$), differences between accuracy rates to incongruent and control items were not significant ($\psi_{PB} = 0.02$, $p = .43$). Finally, there was no main effect of gender ($F_{ATS} (1, 25.42) = 0.08$, $p = .76$); and the three 2-way and the one 3-way interactions did not reach significance (all $p > .4$).

Other descriptive statistics and multi-wise comparisons of interest are shown in Table 3.

Discussion

This study was concerned with the crossmodal Stroop interference between flavour perception and vision (G1) and flavour perception and audition (G2). The main findings showed faster RTs for spoken than for written words and lower accuracy rates for spoken than for written words. Also, faster RTs and fewer mistakes were made in the congruent condition than in the incongruent and control conditions. Taken together, these results suggest that visual representation of words demands higher attentional resources than auditory representations of words.

The results from G1 (visual) showed lower crossmodal Stroop interference in congruent conditions, where accuracy rates were higher and RTs shorter when the visually

presented word referred to the same fruit the participant tasted. In this case, attentional competition (Cho, Lien, & Proctor, 2006; Kahneman & Chajczyk, 1983; Kim et al., 2008) between flavour and visual linguistic stimuli was minimal because the information converged on the same object. These results are consistent with those of Razumiejczyk et al. (2011, 2015). By contrast, with incongruent and control stimuli, the data showed longer RTs and lower accuracy rates. In these conditions, crossmodal Stroop interference was greater and participants failed to inhibit visual linguistic distractors in order to identify taste stimuli quickly and accurately.

G2 (auditory) showed the same response pattern as G1 (visual) in regards to the type of stimuli used. That is, crossmodal Stroop interference in conditions of congruency was lower than with control and incongruent stimuli. With congruent stimuli, RTs were shorter and accuracy rates were higher than in control or incongruent conditions. This means that the Stroop test activates the processes of selective attention (Cho et al., 2006; Kahneman & Chajczyk, 1983; Kim et al., 2008) so that the processing of flavour stimuli competes for attentional resources with the processing of auditory stimuli. In the case of incongruent and control stimuli, the results suggest that the participants were unable to inhibit the auditory linguistic distractors to identify the flavour stimulus. This is in accordance with previous studies (Razumiejczyk et al., 2011, 2015). Consequently, the data showed longer RTs and lower accuracy rates when participants were administered incongruent or control stimuli together with flavour stimuli.

It should be noted that both experimental groups showed the same pattern of results in the dependent variables of RTs and the number of correct responses. Crossmodal Stroop interference levels were greater with incongruent and control stimuli. Subjects were not able to avoid the perceptive analysis of attributes irrelevant to the task (Kahneman, 1997), such as processing words presented either visually or aurally. In contrast, congruent stimuli functioned as facilitators in the identification of flavour stimuli. Thus, the results obtained for the dependent variables of congruency in both experimental groups were coherent.

Comparison between vision and hearing

Regarding the RT variable, results showed that G1 (visual) scored significantly higher than G2 (auditory). These data show that, in the task of identifying flavour stimuli, words presented visually consistently functioned as a greater distractor than words presented aurally. Thus, although in both cases the distractors were linguistic representations, the sensory channel and the encoding mode of the representation influenced processing time. Consequently, when linguistic (visual or auditory) and flavour stimuli compete for attention (Kahneman & Chajczyk, 1983; Kim et al., 2008), visual representations of words demand more processing time and, therefore, higher attentional resources than auditory representations of words. This is consistent with Roelofs (2005).

These findings are in agreement with those of Razumiejczyk et al. (2011), who verified the same trend

Table 3 Descriptive statistics and results of multiple pairwise comparisons (Note. Descriptive statistics represent mean values; standard deviations are presented in brackets).

Test	WS _f	BS _{is}		Test
		Vision group	Audition group	
PB method ($\hat{\Psi}_{PB}$)	C	11.57 (3.60) _{RT}	6.13 (1.92) _{RT}	$\hat{\Psi}_{VG-AG}=5.70, p<.001$ _{RT}
		0.82 (0.16) _{AR}	0.66 (0.29) _{AR}	$\hat{\Psi}_{VG-AG}=0.16, p=.03$ _{AR}
	I	15.39 (4.54) _{RT}	7.23 (2.43) _{RT}	$\hat{\Psi}_{VG-AG}=8.12, p<.001$ _{RT}
		0.54 (0.26) _{AR}	0.45 (0.28) _{AR}	$\hat{\Psi}_{VG-AG}=0.09, p=0.19$ _{AR}
	C'	14.52 (4.24) _{RT}	7.01 (2.42) _{RT}	$\hat{\Psi}_{VG-AG}=7.62, p<.001$ _{RT}
		0.60 (0.18) _{AR}	0.37 (0.24) _{AR}	$\hat{\Psi}_{VG-AG}=0.22, p<.002$ _{AR}
		$\hat{\Psi}_{C-I} = -3.56, p<.001$ _{RT}	$\hat{\Psi}_{C-I} = -1.13, p<.002$ _{RT}	
		$\hat{\Psi}_{C-C} = -2.74, p<.001$ _{RT}	$\hat{\Psi}_{C-C} = -0.82, p<.007$ _{RT}	
		$\hat{\Psi}_{I-C} = 0.81, p=0.29$ _{RT}	$\hat{\Psi}_{I-C} = 0.30, p=0.41$ _{RT}	
		$\hat{\Psi}_{C-I} = 0.27, p<.001$ _{AR}	$\hat{\Psi}_{C-I} = 0.21, p<.003$ _{AR}	
		$\hat{\Psi}_{C-C} = 0.21, p<.007$ _{AR}	$\hat{\Psi}_{C-C} = 0.29, p<.002$ _{AR}	
		$\hat{\Psi}_{I-C} = -0.06, p=0.39$ _{AR}	$\hat{\Psi}_{I-C} = 0.08, p=0.10$ _{AR}	

Note. SR method = sequentially rejective pairwise comparison method, PB method = percentile bootstrap pairwise comparison method, WS_f = within-subjects factor: CL = congruence level (C=congruent, I=incongruent, and C'=control), BS_{is} = between-subjects factors, VG = vision group, AG = audition group, RT = response time, and AR = accuracy rate. Significant results are shaded.

between linguistic and pictorial distractors. Thus, visually presented words constitute a strong distractor due to the high level of processing that competes with the task of identifying flavour stimuli. Here, the results suggest that the meaning of the picture can be accessed rapidly (Potter, 1975). This occurred because all of the features in the picture are perceived simultaneously (Chen & Spence, 2011). Deciphering a written word requires a sequence.

In relation to the accuracy rate results, there were significant differences between congruent and control conditions. However, there were no differences between G1 (visual) and G2 (auditory) in the case of incongruent stimuli. With regard to congruent and control stimuli, G1 (visual) obtained higher values than G2 (auditory). This means that, when the distractor word was presented visually, participants made fewer errors than when the distractor word was presented aurally. However, with incongruent stimuli, accuracy rates were similar for either visual or auditory word distractors. In congruent and control conditions, participants were better at identifying the flavour stimulus when the distractor word was written than when it was spoken. Participants in both groups had similar error rates with incongruent stimuli.

In brief, in congruent and control conditions, RT was longer but the number of correct responses higher when the distractor was presented written rather than spoken. That is, in G1 (visual), processing time was longer and accuracy rates were higher with congruent and control stimuli but showed longer RT and lower accuracy rates in incongruent conditions. However, in G2 (auditory), in the case of incongruent stimuli, the RT was shorter but accuracy rates were similar to G1 (visual). The results establish that,

in congruent conditions, attentional competition between visual linguistic stimulus and flavour stimulus was higher in G1 (visual) than in G2 (auditory), as regards both RT and accuracy rates. These data suggest that due to the demands of the act of reading, processing written words is deeper than processing spoken words. This is consistent with Chen and Spence (2011).

Conversely, at the level of incongruent stimuli, the data showed that even though G1 (visual) had longer RTs, accuracy rates were similar to G2 (auditory). Thus, when the distractor was a written word, subjects took longer to identify the flavour stimuli. Nevertheless, they responded similarly when the distractor was a spoken word. This can be explained by the fact that the visual and the auditory linguistic distractors referred to specific objects (fruits) which were different from the ones which the participants tasted. Thus, the results suggest that the distracting effect was the same in both experimental groups as regards number of correct responses. However, processing time was different. Written words generated a greater attentional competition than gustative stimuli and required more processing time than spoken words.

In this experiment, there were no significant differences in incongruent and control conditions within G1 and G2. Contrary to expectations, there were no significant differences in the dependent variables (RT and number of correct responses) between the three levels of the congruency. It could be entertained that incongruent stimuli do not differ semantically from control stimuli. At a computational processing level, incongruent stimuli were different from the fruit being tasted. Nevertheless, they were similar to the control since they were both words that did

not generate differences in computational expense in the flavour identification process. However, in order to provide more insight for this alternative explanation, it is necessary to carry out semantic gradient experiments over a scale of incongruent stimuli. Thus, it is suggested a study be conducted of semantic gradient with two levels of incongruent stimuli according to the semantic distance between words of incongruent stimuli and flavour stimuli. For example, for the peach flavour stimulus, the visual stimulus conflicting could be damask. Both of these stimuli, peach and damask, are similar in flavour and visual appearance. In contrast, peaches and apples are dissimilar in flavour and in visual appearance.

A limitation of the current study is the use of a specific selection of flavour stimuli. It is recommended to extend this selection to include stimuli that meet the same validation conditions used here. However, an advantage of the current study is the natural condition of the materials used. Pured fruits were administered at room temperature, unlike other studies which used artificial essences as stimuli. All stimuli used in the current study belonged to the usual diet of the participants. These results are important because they provide evidence of adaptation of the organism to the environment (Pinker, 1997).

Another limitation of this study was the omission of digitally controlled, audio stimuli. Instead, it was decided to use in vivo auditory stimuli to maintain the same ecological condition that was used to produce the gustatory materials. This ecological condition is favourable, because it adapted materials to participants (Dhimi, Hertwig, & Hoffrage, 2004). However, it is also possible that it introduced uncontrolled variations. For the same reason, earphones were not employed. Consequently, the laboratory was prepared in such a way as to minimize external noise. We recommend using computerized audio recordings and earphones in future studies to explore possible variations thereon.

Finally, the experimental procedure itself had some limitations. The experimenter who collected data was naive concerning the specific aim of the study but was carefully trained to conduct the procedure. We decided to include an experimenter instead of a response device, because in previous studies we found that participants pressed the response key before giving their verbal response. We observed an irregular acceleration between the former and the latter that took 1–3 s approximately less than the procedure that included a naive experimenter. We think that both procedures are biased, but the inclusion of the same experimenter in all cases provided more control than the use of a response device. The latter introduced more lack of control, because each participant generated a different bias in each response. Hence, the response device procedure entailed both between-subjects and within-subjects biases.

Acknowledgments

This research was supported, in part, by grants from the National Scientific and Technical Research Council – Argentina (PIP11420100100139 granted to Eugenia Razumiejczyk). The authors thank Susan

Brunner for proofreading this manuscript, Alejandro Moreno González (alex.mgnz@gmail.com | <http://magnozz.blogspot.com.co/>) for illustrating the experiment set-up and Iryna Losyeva and La Patulya for illustrating the experiment's sequence of events.

References

- Baddeley, A. (1992). Working memory: The interface between memory and cognition. *Journal of Cognitive Neuroscience*, 4(3), 281–288. <http://dx.doi.org/10.1162/jocn.1992.4.3.281>
- Chen, Y. C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, 37(5), 1554–1568. <http://dx.doi.org/10.1037/a0024329>
- Cho, Y. S., Lien, M. C., & Proctor, R. W. (2006). Stroop dilution depends on the nature of the colour carrier but not on its location. *Journal of Experimental Psychology: Human Perception & Performance*, 32, 826–839. <http://dx.doi.org/10.1037/0096-1523.32.4.826>
- Dhimi, M. K., Hertwig, R., & Hoffrage, U. (2004). The role of representative design in an ecological approach to cognition. *Psychological Bulletin*, 6, 959–988. <http://dx.doi.org/10.1037/0033-2909.130.6.959>
- Driver, J., & Spence, C. (2000). Multisensory perception: Beyond modularity and convergence. *Current Biology*, 10, 731–735. [http://dx.doi.org/10.1016/S0960-9822\(00\)00740-5](http://dx.doi.org/10.1016/S0960-9822(00)00740-5)
- Dyer, F. N. (1973). Interference and facilitation for colour naming with separate bilateral presentations of the word and colour. *Journal of Experimental Psychology*, 99, 314–317. <http://dx.doi.org/10.1037/h0035245>
- Fleiss, J. L., Levin, B., & Paik, M. C. (2004). *Statistical methods for rates and proportions*. New York, NY: Wiley.
- Hoegg, J., & Alba, J. W. (2007). Taste perception: More than meets the tongue. *Journal of Consumer Research*, 33(4), 490–498. <http://dx.doi.org/10.1086/510222>
- Kahneman, D. (1997). *Atención y esfuerzo*. España: Biblioteca Nueva.
- Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by colour-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, 9, 497–509. <http://dx.doi.org/10.1037/0096-1523.9.4.497>
- Kampstra, P. (2008). Beanplot: A boxplot alternative for visual comparison of distributions. *Journal of Statistical Software*, 28(1), 1–9. Retrieved from <http://www.jstatsoft.org/v28/c01/>
- Kim, H., Cho, Y. S., Yamaguchi, M., & Proctor, R. W. (2008). Influence of colour availability on the Stroop colour-naming effect. *Perception & Psychophysics*, 70, 1540–1551. <http://dx.doi.org/10.3758/PP.70.8.1540>
- Kirn, S. Y., Kirn, M. S., & Chun, M. M. (2005). Concurrent working memory load can reduce distraction. *Proceedings of National Academy of Sciences of the United States of America*, 102, 16524–16529. <http://dx.doi.org/10.1073/pnas.0505454102>
- Knöferle, K., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic Bulletin & Review*, 1–15. <http://dx.doi.org/10.3758/s13423-012-0321-z>
- Kuehl, R. O. (1999). *Design of experiments. Statistical principles of research design and analysis*. Pacific Grove, CA: Duxbury Press.
- Lamers, M. J. M., & Roelofs, A. (2007). Role of gestalt grouping in selective attention: Evidence from the Stroop task. *Perception & Psychophysics*, 69, 1305–1314. <http://dx.doi.org/10.3758/BF03192947>

- Ludden, G. D. S., Schifferstein, H. N. J., & Hekkert, P. (2009). Visual-tactual incongruities in products as sources of surprise. *Empirical Studies of the Arts*, 27, 63–89. <http://dx.doi.org/10.2190/EM.27.1.d>
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203. <http://dx.doi.org/10.1037/0033-2909.109.2.163>
- Marmolejo-Ramos, F., Elosúa, M. R., Yamada, Y., Hamm, N., & Noguchi, K. (2013). Appraisal of space words and allocation of emotion words in bodily space. *PLOS ONE*, 8(12), e81688. <http://dx.doi.org/10.1371/journal.pone.0081688>
- Marozzi, M. (2012). A modified Cucconi test for location and scale change alternatives. *Revista Colombiana de Estadística*, 35(3), 369–382. Retrieved from <http://www.emis.de/Journal/RCE/V35/v35n3a03.pdf>
- Marozzi, M. (2014). The multisample Cucconi test. *Statistical Methods and Applications*, 23(2), 209–227. <http://dx.doi.org/10.1007/s10260-014-0255-x>
- Massaro, D. W. (1999). Speechreading: Illusion or window into pattern recognition. *Trends in Cognitive Science*, 3, 310–317. [http://dx.doi.org/10.1016/S1364-6613\(99\)01360-1](http://dx.doi.org/10.1016/S1364-6613(99)01360-1)
- Matlin, M. W., & Foley, H. J. (1996). *Sensación y percepción* (3rd ed.). México: Prentice-Hall.
- McCown, D., & Arnoult, M. D. (1981). Interference produced by modified Stroop Stimuli. *Bulletin of the Psychonomic Society*, 17, 5–7. <http://dx.doi.org/10.3758/BF03333649>
- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264, 746–748. <http://dx.doi.org/10.1038/264746a0>
- Morrot, G., Brochet, F., & Dubourdieu, D. (2001). The colour of odors. *Brain Language*, 79, 309–320. <http://dx.doi.org/10.1006/brln.2001.2493>
- Noguchi, K., Gel, Y. R., Brunner, E., & Konietzke, F. (2012). nparLD: An R software package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal of Statistical Software*, 50(12), 1–23. Retrieved from <http://www.jstatsoft.org/v50/i12/>
- Orhan, A. E., Sims, C. R., Jacobs, R. A., & Knill, D. C. (2014). The adaptive nature of visual working memory. *Current Directions in Psychological Science*, 23(3), 164–170. <http://dx.doi.org/10.1177/0963721414529144>
- Pearce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162, 8–13. <http://dx.doi.org/10.1016/j.jneumeth.2006.11.017>
- Pinker, S. (1997). *How the mind works*. New York, NY: Norton.
- Piqueras-Friszman, B., & Spence, C. (2015). Sensory expectations based on product-extrinsic food cues: An interdisciplinary review of the empirical evidence and theoretical accounts. *Food Quality and Preference*, 40, 165–179. <http://dx.doi.org/10.1016/j.foodqual.2014.09.013>
- Posner, M. I., & Dehaene, S. (1994). Attentional networks. *Trends in Neurosciences*, 17, 75–79. [http://dx.doi.org/10.1016/0166-2236\(94\)90078-7](http://dx.doi.org/10.1016/0166-2236(94)90078-7)
- Potter, M. C. (1975). Meaning in visual search. *Science*, 187, 965–966.
- Prescott, J. (2015). Multisensory processes in flavour perception and their influence on food choice. *Current Opinion in Food Science*, 3, 47–52. <http://dx.doi.org/10.1016/j.cofs.2015.02.007>
- Purves, D., Augustine, G., Fitzpatrick, D., Katz, L. C., Lamantia, A.-S., & McNamara, J. O. (2004). *Invitación a la neurociencia*. Buenos Aires: Panamericana.
- Razumiejczyk, E., Bacci, C., Iriarte, M. P., Britos, P., Genovese, I., Grigera Monteaudo, D., & ... Bellucci, P. M. (2010). Selección de estímulos para el estudio de los procesos cognitivos relacionados con la modalidad gustativa. *Psicología y Psicopedagogía*, 23. Retrieved from <http://p3.usal.edu.ar/index.php/psico/article/view/146>
- Razumiejczyk, E., Macbeth, G., & Adrover, J. F. (2011). Comparación de la interferencia Stroop intermodal entre representaciones gustativas y visuales por imágenes y por palabras. *Boletín de Psicología*, 101, 7–20. Retrieved from <http://www.uv.es/seoane/boletin/previos/N101-1.pdf>
- Razumiejczyk, E., Macbeth, G., & Leibovich de Figueroa, N. (2013). Processing interference between gustative stimuli and incomplete linguistic visual stimuli in working memory. *Revista Latinoamericana de Psicología*, 45(2), 229–237. <http://dx.doi.org/10.14349/rlp.v45i2.1225>
- Razumiejczyk, E., Macbeth, G., Marmolejo-Ramos, F., & Noguchi, K. (2015). Crossmodal integration between visual linguistic information and flavour perception. *Appetite*, 91, 76–82. <http://dx.doi.org/10.1016/j.appet.2015.03.035>
- Regan, J. E. (1978). Involuntary automatic processing in colour-naming tasks. *Perception & Psychophysics*, 24, 130–136. <http://dx.doi.org/10.3758/BF03199539>
- Roelofs, A. (2005). The visual-auditory color-word Stroop asymmetry and its time course. *Memory & Cognition*, 33, 1325–1336. <http://dx.doi.org/10.3758/BF03193365>
- Rolls, E. T. (2004). Multisensory neuronal convergence of taste, somatosensory, visual, olfactory, and auditory inputs. In G. A. Calvert, C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 311–331). Cambridge, MA: MIT Press.
- Shankar, M. U., Levitan, C. A., Prescott, J., & Spence, C. (2009). The influence of color and label information on flavor perception. *Chemosensory Perception*, 2, 53–58. <http://dx.doi.org/10.1007/s12078-009-9046-4>
- Shankar, M. U., Levitan, C. A., & Spence, C. (2010). Grape expectations: The role of cognitive influences in colour-flavour interaction. *Consciousness and Cognition*, 19, 380–390. <http://dx.doi.org/10.1016/j.concog.2009.08.008>
- Small, D. M., Gerbar, J. C., & Hummel, T. (2005). Differential neural responses evoked by orthonasal versus retronasal odorant perception in humans. *Neuron*, 47, 593–605. <http://dx.doi.org/10.1016/j.neuron.2005.07.022>
- Spence, C., Levitan, C. A., Shankar, M. U., & Zampini, M. (2010). Does food colour influence taste and flavour perception in humans? *Chemosensory Perception*, 3(1), 68–84. <http://dx.doi.org/10.1007/s12078-010-9067-z>
- Spence, C., & Shankar, M. U. (2010). The influence of auditory cues on the perception of, and responses to, food and drink. *Journal of Sensory Studies*, 25, 406–430. <http://dx.doi.org/10.1111/j.1745-459X.2009.00267.x>
- Spence, C. (2012). Auditory contributions to flavour perception and feeding behaviour. *Physiology & Behaviour*, 107, 505–515. <http://dx.doi.org/10.1016/j.phybeh.2012.04.022>
- Stevenson, R. J., & Boakes, R. A. (2004). Sweet and sour smells: Learned synesthesia between the senses of taste and smell. In G. A. Calvert, S. C. Spence, & B. E. Stein (Eds.), *The handbook of multisensory processes* (pp. 69–83). Cambridge, MA: MIT Press.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 8, 643–666. <http://dx.doi.org/10.1037/h0054651>
- Sun, B. C., & Halpern, B. P. (2005). Identification of air phase retronasal and orthonasal odorant pairs. *Chemical Senses*, 30(8), 693–706. <http://dx.doi.org/10.1093/chemse/bji062>
- Velasco, C., Jones, R., King, S., & Spence, C. (2013). Assessing the influence of the multisensory environment on the whiskey drinking experience. *Flavour*, 2(23) <http://dx.doi.org/10.1186/2044-7248-2-23>
- Vélez, J. I., & Correa, J. C. (2014). Should we think of a different median estimator? *Comunicaciones en Estadística*, 7(1), 11–17. Retrieved from <http://revistas.usta.edu.co/index.php/estadistica/article/download/1138/1372>
- West, R. (2004). The effects of aging on controlled attention and conflict processing in the Stroop task. *Journal of*

- Cognitive Neuroscience*, 16, 103–113. <http://dx.doi.org/10.1162/089892904322755593>
- White, T. L., & Prescott, J. (2007). Chemosensory cross-modal Stroop effects: Congruent odors facilitate taste identification. *Chemical Senses*, 32, 337–341. <http://dx.doi.org/10.1093/chemse/bjm001>
- Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing*. San Diego: Elsevier.
- Yamada, Y., Sasaki, K., Kunieda, S., & Wada, Y. (2014). Scents boost preference for novel fruits. *Appetite*, 81, 102–107. <http://dx.doi.org/10.1016/j.appet.2014.06.006>