THE PSYCHOLINGUISTICS OF SYNAESTHESIA

by

Jennifer Lauren Mankin

Thesis submitted for the degree of Doctor of Philosophy

School of Psychology

University of Sussex

September 2017

Final corrections: May 2018

Declaration

The thesis conforms to an 'article format' in which the middle chapters consist of discrete articles written in a style that is appropriate for publication in peer-reviewed journals in the field. The first and final chapters present synthetic overviews and discussions of the field and the research undertaken.

Chapter 2 is published in *Cognition* as:

Mankin, J. L., Thompson, C., Branigan, H. P., & Simner, J. (2016). Processing compound words: Evidence from synaesthesia. *Cognition, 150*, 1-9.

The author contributions are as follows: Jennifer Mankin was responsible for all aspects of designing and preparing study materials, collecting and analysing the data, and writing, revising, and finalizing the manuscript; Christopher Thompson was responsible for creating the online data collection apparatus; Julia Simner and Holly Branigan were responsible for providing feedback on study design and corrections to the manuscript. Jennifer Mankin and Julia Simner were collectively responsible for initial conception of the research and preparing the manuscript for publication.

Chapters 3 and 4 have both been submitted for publication.

The author contributions are as follows: Jennifer Mankin was responsible for all aspects of data collection, data analysis, and writing, revising, and finalizing the manuscript; Julia Simner was responsible for providing feedback on study design and corrections to the manuscript. Jennifer Mankin and Julia Simner were collectively responsible for initial conception of the research.

Chapter 5 is published in *Multisensory Research* as:

Mankin, J. L., & Simner, J. (2017). A is for *apple*: The role of letter–word associations in the development of grapheme–colour synaesthesia. *Multisensory Research, 30*(3–5), 409–446.

The author contributions are as follows: Jennifer Mankin was responsible for all aspects of data collection and analysis, and writing the manuscript; Julia Simner was responsible for providing feedback on study design and corrections to the manuscript. Jennifer Mankin and Julia Simner were collectively responsible for initial conception of the research and preparing the manuscript for publication.

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree. However, Chapter 2 of the thesis incorporates material already submitted for the degree of Master of Science in Psychology of Language, which was awarded by the University of Edinburgh. This chapter contains analyses and ideas in addition to the material previously submitted.

____

Jennifer L Mankin
September 2017

Acknowledgements

I owe heartfelt thanks to so many people:

To my supervisor, Professor Julia Simner, for her tireless dedication to excellence and clarity; for sharing her infectious enthusiasm and expertise in both synaesthesia and language; and for bringing me to Sussex, which turned out for a multitude of reasons to be one of the best things that has ever happened to me. Thank you for working with me, fighting for me, and believing in me throughout this long process.

To those who have been instrumental in the completion of this thesis, in particular:
Professors Jamie Ward and Holly Branigan, who guided and supported me with their expertise, their compassion, and their wisdom.
Dr Christopher Thompson, who programmed my online tests and was unfailingly patient and helpful in preparing, running, and obtaining data from the tests.
Dr James Alvarez, for invaluable assistance with statistical analyses and colourspace calculations, putting up with incessant questions, and not killing us all with a dragon that one time.
Dr Duncan Carmichael, for a sympathetic ear and steady, grounded encouragement and perspective when I needed it most.
Drs Clare Jonas and Alan Nielsen, for constructive and insightful advice.

To all the people who have come with me on this odyssey of my research, particularly the members of the Synaesthesia and Sensory Integration Lab in both its Sussex and Edinburgh incarnations, who have listened patiently and constructively to excited theories, puzzled through difficult analyses, and put up with endless diagrams of the word *rainbow*.

To the people who have battled monsters with me – sometimes with dice, sometimes with words – and kept me well-armed for the fight with chocolate, laughter and hugs: especially Fraser, Louisa, Vlad, Becks, Góa, Julie, James, Molly, Clare, and Tory.

Very especially, to the synaesthetes who have not only shared their colours with me, but also sent me letters, written me detailed descriptions of their synaesthesia, and given their time and effort to represent their unique experiences of language and colour. Thank you for your help, your dedication, and your trust.

Finally, and by no means least, to my mother. This is for you, Grandpa, and Dad.
I owe you everything. Thank you.

Explanations exist; they have existed for all time;
there is always a well-known solution to every human problem —
neat, plausible, and wrong.
H. L. Mencken (1917)

Words are magic.
Prinzmetal, Hoffman, and Vest (1991, p. 1)

UNIVERSITY OF SUSSEX

JENNIFER LAUREN MANKIN

PHD PSYCHOLOGY

THE PSYCHOLINGUISTICS OF SYNAESTHESIA

Summary

To most people, a question like "What colour is the letter *A*?" may seem nonsensical, but to a grapheme-colour synaesthete, each letter and word has an automatically evoked colour sensation associated with it. This thesis asks whether the synaesthetic colours for letters and words are shaped by the same influences that inform the typical use of language – that is, if grapheme-colour synaesthesia is fundamentally psycholinguistic in nature. If this is the case, the colour experiences of synaesthetes for letters and words can also be used to investigate long-standing questions about how language acquisition and processing work for everyone.

This thesis addresses two aspects of the psycholinguistic roots of synaesthesia: structure/morphology and meaning/semantics. The first two studies on word structure collected colour responses from synaesthetes for compound words (e.g. *rainbow*), the constituent morphemes of those words separately (e.g. *rain* and *bow*), and the letters that in turn form those words (e.g. *R, A, B,* etc.). These studies showed that synaesthetic word colouring does indeed encode linguistic properties such as word frequency and morphological structure. Furthermore, both linguistic and colour elements of words were important in determining their synaesthetic colour. The second two studies turned to the semantic aspect of language, asking how the meanings associated with words (e.g. *red, fire*) and even individual letters (e.g. *A, Q*) can influence the colours that a synaesthete experiences for them. The first of these studies indicated that the synaesthetic colour for a word like *red* or *fire* was measurably influenced by the colour that word typically evokes (e.g. the red of *red* and the orange of *fire*). The second showed that trends in letter-colour associations in large-scale studies (e.g. *A* is typically red) may be rooted in connections to particular words (e.g. *A* is red because *A* is for apple and apples are red). Overall, this thesis shows that both word structure and meaning have a systematic, measureable effect on synaesthetic colour, which allows these colours to then be used as a new tool to investigate psycholinguistic questions.

Table of Contents

Chapter 1

Introduction:
Language and grapheme-colour synaesthesia

Why might *A* be blood red, or *7* leaf green? People with *grapheme-colour synaesthesia* experience highly specific impressions of colour associated with graphemes such as letters, numbers, and words (Rich, Bradshaw, & Mattingley, 2005; Simner, 2012; Ward, Simner, & Auyeung, 2005). Synaesthesia in general is present in approximately 4.4% of the general population (Carmichael, Down, Shillcock, Eagleman, & Simner, 2015; Simner, Mulvenna, et al., 2006), and is characterised by unusual and automatically evoked experiences (termed *concurrents*, such as a specific colour like blood red) in response to everyday stimuli (termed *inducers*, such as a letter like *A*; Simner, 2012). Grapheme-colour synaesthesia in particular represents more than a quarter of all cases of synaesthesia, or about 1% of the general population (Carmichael et al., 2015; Novich, Cheng, & Eagleman, 2011; Simner, Mulvenna, et al., 2006; Simner & Carmichael, 2015). Although still uncommon, this puts the number of grapheme-colour synaesthetes worldwide in the tens of millions. Therefore, grapheme-colour synaesthesia represents a striking opportunity to study human perception and cognition using these extraordinary colour experiences. The main goal of this thesis is to investigate grapheme-colour synaesthesia through a psycholinguistic lens, both to better understand synaesthesia itself and to address enduring questions about normal language processing in everyone.

Before delving into the linguistic aspects of grapheme-colour synaesthesia, it is first important to establish why studying these unusual experiences can yield such important insights into cognition in general, and language in particular. I will first briefly summarise the defining behavioural and neurological characteristics of synaesthesia, including how synaesthetes can be distinguished from non-synaesthetes. I will then turn to the main question of this thesis: how and why do particular words and letters evoke particular synaesthetic colours? For instance, why might *A* be red rather than blue for any given synaesthete, and what might influence the colours of everyday words like *rain* or *red*? This thesis will suggest that the colours that synaesthetes experience for these items are fundamentally rooted in their linguistic characteristics. To establish this, I explore the ways that research thus far has addressed the roots of grapheme-colour synaesthesia. I will characterise one main approach as primarily *perceptual*, such that synaesthetic colours are thought to be determined by low-level properties of letters and numbers, such as their visual shape. The second approach I will characterise as *conceptual*, which suggests that synaesthetic colours are also based on abstract ideas and conceptual features of language. This thesis is founded on the central proposition, currently gaining interest and support, that the influences that guide normal cognition – and language in particular – may be reflected in synaesthetic experiences as well. Therefore, I will consider both the perceptual and conceptual influences on synaesthesia within the domain of language. Finally, I will introduce the experiments that form the body of this thesis, which are designed to investigate both structural and semantic aspects of language to better understand the underlying systematicity of synaesthesia.

Identifying grapheme-colour synaesthesia

I begin by asking how synaesthetes can reliably be identified from non-synaesthetes. That is, how can synaesthetic experiences be verified as genuinely automatic and involuntary, and not invented or imagined? Implicit in this question is defining what uniquely characterises synaesthetic experiences as separate from mental imagery, hallucinations, or an "overactive imagination." What does and does not count as synaesthesia is an ongoing debate in the field (Simner, 2012a, 2012b), but both behavioural and neurological evidence have identified synaesthetes as quantitatively distinguishable from non-synaesthetes in a number of ways. The first objective verification of synaesthesia was a case study of grapheme-colour synaesthete EP (Baron-Cohen, Wyke, & Binnie, 1987). EP was asked to report the colours she experienced for letters and words, and was then given a surprise retest 10 weeks later on the same items. EP was 100% consistent with her colours at the surprise retest, while a control participant was only able to accurately recall 17% of the colour associations she had been asked to invent and memorise with only a two-week delay before retest. This suggested that EP's colours were not invented or remembered, but rather perceptually *experienced*; she could report them with perfect accuracy because she was seeing them in her mind's eye in response to the inducing letters and words.

Baron-Cohen et al.'s (1987) study also established consistency over time as a defining characteristic of the genuineness of synaesthetic experiences. Subsequent studies have pervasively employed this surprise test-retest method, consistently finding that synaesthetes performed at or near ceiling accuracy, far above the performance of non-synaesthete controls (e.g. Baron-Cohen, Harrison, Goldstein, & Wyke, 1993; Mills et al., 2002; Rich et al., 2005; Simner et al., 2005; Simner, Harrold, Creed, Monro, & Foulkes, 2009; Simner, Mulvenna, et al., 2006; Ward & Simner, 2005; Ward et al., 2005). The obvious downside of this method is the long delay to verify a synaesthete as genuine. In 2007, Eagleman, Kagan, Nelson, Sagaram, and Sarma introduced an online, single-session synaesthesia test available at www.synesthete.org, built on the same principle of consistency. This test presents letters and numbers three times each in random order, and synaesthetes use an online palette with 16.8 million possible colours to indicate their colour experience for each. The distance in colourspace between the three colours for each grapheme is calculated to produce a consistency score. A lower score indicates a smaller distance between the colours for any given grapheme and therefore higher consistency. Eagleman et al. (2007) suggested that a score lower than 1 indicated the level of consistency characteristic of genuine synaesthesia. This cutoff was later revised to 1.43 by Rothen, Seth, Witzel, and Ward (2013), who showed that this maximised sensitivity in distinguishing synaesthetes from non-synaesthetes. This test means that synaesthetes can now be identified quickly and easily, and anyone can take the test online and opt into a database for research participation, as in many of the experiments reported in the following chapters.

However, consistency-based tests of genuineness are not the only way that synaesthetes are distinguished from non-synaesthetes, and likewise consistency is not the only hallmark of synaesthesia. Another well-researched effect is the so-called synaesthetic Stroop test (cf. Jensen & Rohwer, 1966; Stroop, 1935). For non-synaesthetes, it might not make a difference whether the letter *M* is coloured orange or blue; for a synaesthete, however, an orange *M* might be as natural as the word *red* coloured red, while a blue *M* will have the same disruptive effect as the word *red* coloured blue. Indeed, synaesthetes are consistently slower to name the colours of graphemes printed in an incongruent colour than in a colour congruent with their synaesthesia (Dixon, Smilek, Cudahy, & Merikle, 2000; Dixon, Smilek, & Merikle, 2004; Mattingley, 2009; Mattingley, Rich, Yelland, & Bradshaw, 2001; Mills, 1999; Rouw, van Driel, Knip, & Ridderinkhof, 2013). This too is taken as evidence that synaesthetic colour experiences are automatically evoked and perceived rather than remembered.

If this is the case, the connections between graphemes and colours unique to synaesthesia should also be evident in brain activity when synaesthetes are exposed to synaesthesia-inducing stimuli. Indeed, fMRI studies have found that grapheme-colour synaesthetes display increased activation in colour-selective regions of the brain, such as the V4 colour processing area, compared to non-synaesthetes when viewing monochrome graphemes (Gray et al., 2006; Hubbard & Ramachandran, 2005; Nunn et al., 2002; Ramachandran & Hubbard, 2001b; Specht & Laeng, 2011). Laeng, Hugdahl, and Specht (2011) also showed that the distance in colourspace between a visually presented colour and a synaesthetically experienced colour correlated with the degree of activation in these colour-processing areas of the brain as well. Using ERP, Brang, Hubbard, Coulson, Huang, and Ramachandran (2010) similarly identified increased activation in V4 in synaesthetes, but not non-synaesthetes, which occurred nearly simultaneously with grapheme recognition. Synaesthetes also display differences in brain structure, including more coherent white matter and increased grey matter volume in regions of the fusiform gyrus implicated in colour processing (Banissy et al., 2012; Rouw & Scholte, 2007, 2010; Weiss & Fink, 2009). These studies provide clear evidence that grapheme-colour synaesthesia is an objectively verifiable neuropsychological condition.

Apart from establishing that synaesthesia is genuine and automatic, it is also important to distinguish synaesthesia from the cross-modal correspondences between perceptual qualities found in non-synaesthetes as well (Deroy & Spence, 2013; Spence, 2011). An example of these correspondences in the general population is the tendency to pair objects larger in size (visual/tactile modalities) with lower-pitched sound (aural modality). Some researchers have suggested that the automatic associations in synaesthesia are conscious manifestations of cross-modal correspondences underlying general cognition and perception in all people (Simner, 2013; see Deroy & Spence, 2015, for further discussion). This idea has been useful in understanding systematic synaesthetic associations, such as colours for music (Ward, Huckstep, & Tsakanikos, 2006). The colour associations that synaesthetes report for music often mirror the correspondences between basic qualities such as pitch, lightness, and volume also found in the general, non-synaesthetic population (e.g. lighter colours for higher pitches; Deroy & Spence, 2015; Lacey, Martinez, McCormick, & Sathian, 2016; Wan et al., 2014; Ward et al., 2006). These similarities in associations between synaesthetes and non-synaesthetes can be found for grapheme colours as well. Large-scale studies have found that both grapheme-colour synaesthetes and non-synaesthetes exhibit preferences for particular grapheme-colour pairs, such as *A* with red or *X* with black (in English, Jonas, 2010; Rich, Bradshaw, & Mattingley, 2005; Simner et al., 2005; Witthoft, Winawer, & Eagleman, 2015; in other languages, Lavrynenko, 2014; Nagai, Yokosawa, & Asano, 2015; Rouw, Case, Gosavi, & Ramachandran, 2014; Van Leeuwen, Dingemanse, Todil, Agameya, & Majid, 2016). Some of these pairings are shared by both synaesthetes and non-synaesthetes (e.g. *D* with brown) whereas others are particular to each group (e.g. synaesthetes pair *O* with white while non-synaesthetes typically choose orange). This surprising systematicity in seemingly random associations (e.g. why should large groups of people agree that *A* is red, and not blue, or yellow?) suggests that some fundamental influences must underlie the choice of letter-colour pairs in both synaesthetes and non-synaesthetes. To investigate the origin of these population-wide preferences, Spector & Maurer (2008, 2011) examined grapheme-colour associations in preliterate infants. They found evidence that these infants, with no knowledge of or formal exposure to literacy, still systematically associated *X* with black and *O* with white, but showed no sign of the common trend among literate adults to associate *A* with red and *G* with green. They concluded that at least some of these population-wide trends could therefore be based in innate biases linking particular shapes and colours, while others are based in literacy. These common perceptual biases may indeed help determine the particular pairings of grapheme and colour, but unlike synaesthetes, non-synaesthetes are often not conscious of these letter-colour pairings. In sum, grapheme-colour synaesthesia differs from the unconscious impressions of non-synaesthetes in that it is automatically and consciously experienced, and the reality of these synaesthetic experiences is objectively verifiable by both behavioural and neurological methods. Given this, I next ask what characterises these letter- and word-colour associations. That is, what qualities of letters and words influence the colours that a grapheme-colour synaesthete experiences?

What determines the synaesthetic colours of letters and words?

The main goal of this thesis is to investigate *why* synaesthetes experience particular colours for letters and words. Previous research tackling this question has most frequently taken one of two approaches. The first focuses on lower-level, perceptual influences, based on the idea that the particular synaesthetic colours evoked by graphemes are determined by the perceptual features of those graphemes, such as their visual shape. One study examined the basic visual shapes that make up English graphemes, and found that graphemes with similar shapes (e.g. *b* and *d*) tended to be more similarly coloured than graphemes with dissimilar shapes (e.g. *b* and *m*; Brang, Rouw, Ramachandran, & Coulson, 2011; Hubbard, Ambrosio, Azoulai, & Ramachandran, 2005). This is further evidenced by multilingual synaesthetes, who report that similarly shaped graphemes in different orthographies often have similar colours despite different pronunciations (e.g. Greek *ν, ρ* have the same colours as English *v, p* respectively; Mills et al., 2002; Rich, Bradshaw, & Mattingley, 2005; Witthoft & Winawer, 2006). Other work has shown that pronunciation also matters, with similar-sounding orthographic symbols having similar colours across languages (e.g. Korean 나 and Japanese な both pronounced "na" and both synaesthetically yellow; Shin & Kim, 2014). These colour associations can also transfer nearly instantaneously to completely unfamiliar graphemes. These novel graphemes often adopt colours similar to those of familiar graphemes (for example, Russian *И* /i/ coloured like English *N*) and become more refined in colour and consistency with increased exposure (Blair & Berryhill, 2013; Mroczko & Metzinger, 2009). These studies all support the idea that both the visual shape of the inducing grapheme, as well as its pronunciation, are associated with the concurrent synaesthetic colour. This would indicate that low-level perceptual information (i.e. shape and sound) informs synaesthetic associations.

However, the simple shape of the grapheme is not enough to determine its colour – there is a wealth of evidence to indicate that the synaesthetic colour is also critically linked to the identity, or the concept, of the grapheme itself. A study by Dixon, Smilek, Duffy, Zanna, and Merikle (2006) presented the same symbol, for example 5, biased by either a digit or letter context (e.g. *5* in 3 4 5 6 7 but *S* in MUSIC). When asked to indicate the colour of the ambiguous grapheme, the synaesthete participant was slower to respond when there was a mismatch between the biasing context and the colour (e.g. when it was coloured like *S* but in the digit context). This indicates that the trigger for the synaesthetic experience is accessing the meaning of the grapheme itself – i.e. recognising 5 as either an *S* or a *5*, not just perceiving its shape. This conceptually-based, "higher" synaesthesia is also evident when synaesthetes have the same concurrent colour response for different representations of the same concept (e.g. yellow for "4", "IV", "⚃", all representing the conceptual quantity *four*) or for typographical variations of a letter (e.g. red for "A", "a", "𝒜", and

"**A**"; Ramachandran & Hubbard, 2001; Ward, Li, Salih, & Sagiv, 2007). In a similar vein, the concept of a number alone is sufficient to induce a colour response, even without viewing any physical representation of that number; a synaesthete participant was slower to name a colour patch when given an equation such as "3 + 4 = □" if the patch was incongruent to their synaesthetic colour for the number *7* (Dixon et al., 2000). Altogether, this research suggests that the synaesthetic colour is frequently tied to the concept of the grapheme itself, not only its shape or sound.

This central role of meaning characterises a second, conceptual approach to the roots of synaesthetic word colouring. This approach focuses on inducers of synaesthetic colour experiences – such as numbers, letters of the alphabet, or months of the year – as concepts or ordered sequences rather than as perceptual units. Nikolić (2009) argued that the phenomenon of *synaesthesia* (from Greek *syn* "together" + *aesthesis* "sensing") should instead be renamed *ideasthesia*, the sensing of concepts, due to the critical role of semantic and conceptual information in synaesthetic experiences. In the most extreme formulation of synaesthesia having abstract, conceptual influences, a study of over 19,000 synaesthetes by Novich, Cheng, and Eagleman (2011) showed that synaesthesia for letters, numbers, weekdays, and months often co-occurred in the same individual, which led the researchers to re-categorise these synaesthetes into a single subtype they termed *coloured sequence* synaesthesia. They argued that synaesthetic colours for these items were not based on perception or language, but on their membership in overlearned sequences. That is, the synaesthetic colour would be tied to concepts such as "the first day of the week", rather than any of the linguistic features of the word representing that concept (e.g. the frequency, spelling, pronunciation, etc. of the word *Monday*).

The main piece of evidence to test this claim comes from the colours that synaesthetes experience for whole words. For synaesthetes, most words tend to be coloured by their first letter (e.g. *mother* synaesthetically orange due to an orange *M*; Mills et al., 2002; Rich et al., 2005; Ward et al., 2005). However, words belonging to overlearned sequences often have a colour independent of the colours of their letters, e.g., red for *Monday* despite an orange *M,* white *O*, yellow *N*, etc. I will denote these special synaesthetic colours for particular words like *Monday,* which derive from the word's membership in an overlearned sequence, as *idiosyncratic* colours, as red for *Monday* above. This contrasts with the colour the word would typically have based on its spelling, i.e. its *letter-based* colour, such as orange for *Monday* due to the orange *M*. In a large-scale study, Rich et al. (2005) found that words in sequences (i.e. weekdays and months) were indeed significantly less likely to match the synaesthetic colour of their first letter than non-sequential words such as names and occupations.[1] This supports Novich et al.'s (2011) claim that a word in an overlearned sequence, such as *Monday* in the sequence of weekdays, is conceptualised primarily as a unitised, discrete element of that sequence rather than a typical word, and thus has an idiosyncratic colour. Accordingly, colours for sequences such as weekdays and months are considered different subtypes than grapheme-colour synaesthesia, and can occur independently (e.g. a synaesthete having colours for months or days but not for letters or other words; Rich et al., 2005; Simner, Mulvenna, et al., 2006). However, leaving aside the special case of words fossilized in an overlearned sequence, I will argue that grapheme-colour synaesthesia is fundamentally and essentially rooted in language, combining both perceptual and conceptual elements.

The linguistic roots of grapheme-colour synaesthesia

While vital and substantive contributions have been made by both perceptual and conceptual approaches to uncovering the roots of grapheme-colour associations, the linguistic aspect of grapheme-colour synaesthesia has thus far received relatively little attention. In order to understand the synaesthetic colours of letters and words, I will show that it is imperative to treat them primarily as *linguistic* items, and therefore subject to the same influences that govern the normal processing of language. To investigate this, I next ask: how, in everyday life, do synaesthetes experience the colouring of letters and words, and what aspects of those words influence the final colour? This approach, looking at the synaesthetic colours for words based on their linguistic features, is now gathering interest, and it is this aspect of synaesthetic colouring that I will focus on for this thesis.

In 2007, Simner suggested that language has a "special" status in synaesthesia, based on the predominance of linguistic items as inducers, and called for more research into the essentially psycholinguistic basis of grapheme-colour synaesthesia. However, such an investigation would be fruitless if so-called "linguistic" or "lexical" synaesthesia were actually based solely on overlearned sequences, as Novich et al.

---

[1] This analysis apparently did not account for synaesthetes who colour whole words by their first vowel rather than their first letter, which may have influenced these results (see Chapter 3).

(2011) suggested. That is, if colours for words are based entirely on the word's membership in a sequence, then the linguistic characteristics of that word are irrelevant. There are several convincing lines of evidence to suggest that this is not the case, and that a psycholinguistic investigation of synaesthetic word colouring is both warranted and potentially highly rewarding for the fields of both synaesthesia and psycholinguistic research. The first, and most obvious, is that the vast majority of words are *not* part of overlearned sequences, yet still induce synaesthetic colour experiences for these synaesthetes (Baron-Cohen et al., 1987; Blazej & Cohen-Goldberg, 2015; Goodhew & Kidd, 2017; Mankin, Thompson, Branigan, & Simner, 2016 [Chapter 2]; Mills et al., 2002; Rich et al., 2005; Simner, Mulvenna, et al., 2006). In addition, it is not the case that even words in overlearned sequences are always idiosyncratically coloured. A cross-linguistic study by Barnett, Feeney, Gormley, and Newell (2009) found that multilingual synaesthetes tended to experience colours for months based on their spelling (i.e. letter-based colours), rather than idiosyncratic colours based on their shared concept in a sequence. For example, an English-German multilingual synaesthete was more likely to have similar colours for the month *May* in English and in German (*Mai*), while an English-Irish synaesthete was more likely to have different colours for *May* and *Baeltaine*. Furthermore, both idiosyncratic and letter-based colours can co-occur for items in a sequence: Rich et al. (2005) reported that weekdays were more likely to be idiosyncratically coloured than were months. They suggested that this was because weekdays are learned earlier than months in the local school curriculum, so weekdays were learned during childhood as items in a sequence rather than as words. This distinction is also reported in case studies; for example, the synaesthete JW reported idiosyncratic colours for weekdays, but letter-based colours for months (Simner, Glover, & Mowat, 2006). That is, while some words in may be unitised into elements of overlearned sequences with idiosyncratic colours, this by no means explains the colours for all linguistic items.

One of the main topics that this thesis will investigate is how and why whole words are coloured synaesthetically. As mentioned briefly above, previous studies have shown that many synaesthetes experience a whole-word colour derived from a particular letter in that word. For many synaesthetes, this is the first letter in the word (e.g. *mother* coloured like *M*), but can also be the first vowel (e.g. *mother* coloured like *O*) or the first consonant (e.g. *uncle* coloured like *N;* Mills et al., 2002; Ward, Simner, & Auyeung, 2005). Even so, how do letters themselves obtain their particular colours? One synaesthetic colour influence is the frequency of the letter in the language; for example, higher-frequency letters tend to be more saturated, i.e. vividly coloured (Beeli, Esslen, & Jäncke, 2007; Smilek, Carriere, Dixon, & Merikle, 2007; Watson, Akins, & Enns, 2012). More importantly, the synaesthetic colours for words are also influenced by characteristics of the words themselves, beyond the colour influences of letters. For example, Simner, Glover, et al. (2006) showed that for a synaesthete who experienced word colours based on vowels, the prosodic stress of a word changed its synaesthetic colour: ***'con***-*vict* was coloured like *O*, whereas *con*-***'vict*** was coloured like *I*. This is clear evidence that the pronunciation of the word as a whole was critically important in determining its synaesthetic colour. Beyond pronunciation and prosody, morphological structure may also influence synaesthetic colours. In a case study, Blazej and Cohen-Goldberg (2015) reported that their synaesthete participant was more likely to give compound words (e.g. *chessboard*) one colour rather than two depending on the frequency of the first morpheme (e.g. *chess*) and of the compound as a whole, as well as on the meaning of the compound. Meaning also has an effect on the exact colour that a synaesthete experiences for a particular word: Rich et al. (2005) gave an example of a synaesthete reporting yellow for *banana* (i.e. the canonical colour of a banana), which contrasted with the colour the word "should" be based on its letters, and EP similarly reported to Baron-Cohen et al. (1987) that the word *elephant* was grey. It is clear from this evidence that linguistic information in a word, including its frequency, pronunciation, morphological structure, and meaning, can be represented in its synaesthetic colour.

There is even evidence that a synaesthete's knowledge of their language as a whole contributes to their synaesthetic associations. Goodhew and Kidd (2017) asked synaesthetes to give colours for words with conceptual valence, e.g. positive *sun, happy* and negative *mud, doom*. They found that the colour the synaesthetes gave (e.g. yellow for *happy* or black for *doom*) was predicted by language use statistics, particularly the frequency with which the word and colour term co-occurred in English (e.g. how often *yellow* and *happy* occurred together). This means that, above and beyond their constituent letters, the synaesthetic colours for words depend on their context and past exposure in the language. It is clear that although synaesthetic word colours may frequently be derived from their letters, the information that the word carries as a linguistic unit – its pronunciation, prosody, internal structure, meaning, and statistical distribution in the language – all contribute to its colour as well. Aside from the well-evidenced cases of idiosyncratic colours for elements of overlearned sequences, synaesthetic colours for words are indeed just that – colours for *words*, not for rote-learned elements in a sequence. Therefore, this thesis will investigate how linguistic elements of these words may be systematically represented in synaesthetic colours. Exploring this interaction between language and synaesthesia, using the normal words that synaesthetes read, use, write, and hear in everyday language, may open new possibilities in both synaesthesia and psycholinguistic research.

As a final point, I must address a crucial concern, specifically: if synaesthetes experience the world in general, and language in particular, through the perceptual and cognitive lens of synaesthesia, is it reasonable to use the special experiences of these synaesthetes to investigate normal language processing in everyone? That is, are these synaesthetic experiences, and indeed the language faculties of synaesthetes, generalisable to language processes in non-synaesthetes? Synaesthetes do indeed show group differences from non-synaesthetes in particular areas, such as personality (Banissy et al., 2012, 2013; Janik-McErlean & Banissy, 2016; Rouw & Scholte, 2016), mental imagery ability (Barnett & Newell, 2008; Chun & Hupé, 2016; Janik-McErlean & Banissy, 2016; Price, 2009), creativity (Chun & Hupé, 2016; Mulvenna & Hubbard, 2003; Ward, Thompson-Lake, Ely, & Kaminski, 2008), memory (Rothen & Meier, 2010; Rothen, Meier, & Ward, 2012; Teichmann, Nieuwenstein, & Rich, 2017), and cognitive styles (Mealor, Simner, Rothen, Carmichael, & Ward, 2015; Meier & Rothen, 2013). However, these differences appear to be of degree, not of kind, and generally within the normal range (with a few extraordinary exceptions; Bor, Billington, & Baron-Cohen, 2007; Simner, Mayo, & Spiller, 2009; Simner, Treffert, Hughes, Baron-Cohen, & Ward, 2017). For language specifically, as argued by Mankin (2017), claiming that synaesthesia could not inform typical language processing would involve postulating a fundamentally different language faculty from non-synaesthetes. This is not only extremely unlikely, but would be a find of staggering consequence in its own right. Thus far, experimental results have indicated that synaesthesia is an additional experience mapped on top of typical, pre-existing cognitive structures, rather than an essentially different system. For instance, Brang, Hubbard, Coulson, Huang, and Ramachandran (2010) pointed out that in their ERP study of grapheme-colour synaesthesia, the activation patterns for grapheme recognition were essentially the same between synaesthetes and non-synaesthetes. The difference rather emerged in the subsequent activation, or lack thereof, of colour processing areas.

This also aligns with the cognitive development of language in childhood. Children are usually competent, fluent, and fully grammatical speakers of their first language by the age of 5 years (Siegler, Eisenberg, DeLoache, & Saffran, 2014). On the other hand, grapheme-colour synaesthesia does not emerge until graphemes have been acquired, i.e. until the child attains basic literacy, typically during primary school and approaching competency by age 7 (Worden & Boettcher, 1990). Accordingly, synaesthesia is still developing during this period, with children increasing in both the number of graphemes that have colours, and the consistency of those colours, as they grow older (Simner & Bain, 2013, 2017; Simner, Harrold, et al., 2009). This indicates that synaesthetic experiences are formed concurrently with literacy acquisition, which in turn is founded on typical first language acquisition. Therefore, synaesthetes acquire and process language like non-synaesthetes, but with the addition of synaesthetic experiences systematically corresponding to their language's underlying structure. This means synaesthetic colours can be used to better discern the shape and function of that structure, and the linguistic patterns identified in synaesthetes will apply to non-synaesthetes as well.

Thus far, I have argued that grapheme-colour synaesthesia is a genuine and quantifiable condition that is fundamentally and meaningfully rooted in language. Therefore, studying linguistically-induced synaesthetic experiences can provide valuable insights into both synaesthesia and language. First, this will elucidate the underlying conceptual influences on synaesthesia that result in the colour a synaesthete actually experiences when using language. Furthermore, if synaesthetic colours systematically reflect the mechanisms at work in reading, understanding, and using language, synaesthesia can then be used as a tool to investigate typical language processing in everyone. In other words, I argue that grapheme-colour synaesthesia is mapped onto the underlying systematicity of language, and studying these associations will also allow us to study the corresponding psycholinguistic processes underneath. The main purpose of this thesis is to investigate the synaesthetic experience of colour for the letters and words that synaesthetes, and indeed any typical speaker of English, may encounter or use daily. In particular, this thesis will investigate the special status of language in synaesthesia by focusing on two features of language: structure and meaning. By "structure", I refer specifically to the structure of graphemes within a word and morphemes within words. I will explore the current understanding of the role that structure and meaning play in synaesthesia, and thence how these characteristics may reflect or affect synaesthetic experiences.

Structure: The processing of simplex and complex words

In this section, I will focus on how the structural features of language, defined in detail below, may influence synaesthesia. I argue that the structural aspect of linguistic items is worthy of specific interest because the complex internal structure that words have is unique among the stimuli that induce synaesthetic experiences (e.g. shapes, music, etc.). In the first two experimental chapters of this thesis, I will focus on two types of words: complex words, specifically compound words (e.g. *rainbow*; Chapter 2) and the simplex words that function as morphemes in making up those compounds (e.g. *rain* and *bow*; Chapter 3).

First, I investigate *complex* words, which comprise more than one *morpheme* – that is, they contain more than one meaning-bearing element. Complex words can be composed by adding affixes to simplex words (e.g., adding a plural –*s* to *house* to form *houses*), or by combining monomorphemic simplex words to form compound words (e.g., *house* + *boat* → *houseboat*). Compound words such as *houseboat* are of particular interest because they have internal structure between their constituent morphemes (Libben, 2006); that is, a *houseboat* is different than a *boathouse*, for example. I will specifically focus on compound words, as they allow us to compare the colours that synaesthetes experience for each element individually (e.g. the colours of *house* and *boat* separately) to the colour of the compound as a whole. In Chapter 2, I present a study that asks how synaesthetes colour compound words. This study elaborates on an unpublished case study in German, now over a decade old, that investigated whether synaesthetes experience one or two colours for compound words. Concurrently to the publication of Chapter 2, Blazej and Cohen-Goldberg (2015) conducted a similar case study with an English-speaking synaesthete for compound words. In both these studies, if the synaesthete reported two colours, this suggested that the two independent morphemes in the compound (e.g. *house* and *boat* in *houseboat*) are both activated during processing. On the other hand, if the synaesthete only experienced a single colour for *houseboat*, this would suggest that the whole compound has been lexicalised as a single unit. In other words, this is not only interesting in terms of synaesthetic mapping onto linguistic features, but can also provide evidence for the processes by which compound words are stored and accessed in the mental lexicon.

Chapter 2 therefore uses synaesthetic colours to test theories of compound word processing, which fall broadly into three categories. The *supralexical* position posits that compound words are first recognised as wholes, after which the representations of their constituent morphemes are activated (e.g. *rainbow* accessed first, which then activates *rain* and *bow* separately; Butterworth, 1983). Conversely, the *prelexical* theory claims that the constituent morphemes are accessed separately first, followed by the compound as a whole (Pinker, 1991; Stockall & Marantz, 2006; Taft, 1979, 1988, 2004). Combining both, *dual-* or *multiple-route* theories suggest that both of these processes operate simultaneously, with different routes being more or less efficient depending on the features of the compound, such as its frequency (e.g. Kuperman et al., 2008). To investigate this, Chapter 2 describes the first experimental investigation of these compound colour effects with a group of grapheme-colour synaesthetes. We asked our synaesthetes for the colours that they experienced for two lists of compound words, allowing them to choose up to two separate colours for each word. The first list of compounds varied by frequency (e.g. high-frequency *rainbow* and low-frequency *seahorse*). Here we hypothesised that the high-frequency compounds would be more likely than low-frequency compounds to have a single colour rather than two. This would constitute evidence that the higher-frequency compounds were lexicalised, using a direct-lookup processing route and therefore only activating a single colour for the whole word. However, the lower-frequency compounds would first be decomposed into their constituent morphemes, which would evoke the synaesthetic colours associated with both of those morphemes and leading to a two-colour response (Baayen, Dijkstra, & Schreuder, 1997; Baayen, Kuperman, & Bertram, 2010; Kuperman et al., 2008). That is, the variation in synaesthetic colour due to word frequency would indicate that these colours meaningfully reflect the underlying processing of words.

We also tested a second list of compounds that were balanced for frequency but varied on semantic transparency. This quality captures how clearly the meaning of the whole compound is related to the meanings of its constituent morphemes (e.g. transparent *boathouse* vs opaque *hogwash*) and has also been shown to modulate the processing of compounds (e.g. Heyer & Kornishova, 2017; Juhasz, 2007; Libben, 1998, 2010; Marelli & Luzzatti, 2012). Here, we expected that since the morphemes in a transparent compound like *boathouse* are both relevant to its meaning, the colours of those constituent morphemes would both be activated during processing, so the compound would be more likely to have two colours. However, the opaque compound *hogwash* has no semantic relationship to either *hog* or *wash*, so we hypothesised that the representation of the whole word *hogwash* would be accessed directly, and therefore opaque compounds would be more likely to have a single colour. In summary, Chapter 2 investigates how the internal structure and processing of compound words maps onto their synaesthetic colours.

In Chapter 3, we next broke down these colouring effects further at the morpheme and letter level. In other words, we asked what aspects of word structure might be pertinent for synaesthetic colouring. For simplex words (e.g. what colour is *rain*?), we might expect that the synaesthetic colours will follow the same rules as visual word recognition. As noted above, synaesthetes often report that whole words take the colour of their first letter (e.g. *rain* coloured like *R*; Mills et al., 2002; Rich et al., 2005; Ward et al., 2005). This is supported by studies of reading which find that the first letter in a word is disproportionately influential in the recognition of the whole word compared to the other letters (Aschenbrenner, Balota, Weigand, Scaltritti, & Besner, 2017; Grainger, Bertrand, Lété, Beyersmann, & Ziegler, 2016; Johnson & Eisler, 2012; Scaltritti & Balota, 2013). This is somewhat complicated by the fact that for a minority of synaesthetes, the first *vowel* is the main contributor of whole-word colour, even when it is not the first letter, while other synaesthetes use both the first letter and the first vowel (Simner, Glover, et al., 2006; Ward et al., 2005). This suggests that at least for some synaesthetes, the vowel/consonant distinction may be highly salient in word recognition. In fact, the distinction between consonants and vowels is typically processed during word recognition in everyone (Carreiras, Duñabeitia, & Molinaro, 2009; Carreiras, Gillon-Dowens, Vergara, & Perea, 2008; Taft, Xu, & Li, 2017). Therefore, our experiments in Chapter 3 examine synaesthetic colours for words at two levels: letters within morphemes, and morphemes within compound words. We compare each synaesthete's colours for simplex words to the constituent letters of those words (e.g. the colours of *R, A, I,* and *N* compared to the colour of *rain*) to identify the source of the word's colour. We then compare these letter and word colours to the compound colours we obtained from the same synaesthetes in the experiment described in Chapter 2. This allows us to trace the propagation of synaesthetic colour from letter to simplex word, and from simplex word to complex word, to better understand linguistic processing as well as the nature of synaesthetic word colouring.

To summarise, the two chapters on synaesthetic colours for simplex and complex words look at two questions. First, do synaesthetic colours map onto the underlying orthographic and morphological structure of words? And second, how can we then use those systematic correspondences to better understand normal language processing? In each chapter, we summarise the relevant psycholinguistic theories relating to word recognition, simplex words, and compound words, and then apply them to the colours that our synaesthetes experience to show that synaesthesia is indeed a useful tool for studying the structure of language.

Meaning: The influence of imagery and canonical colour

The latter two experimental chapters in this thesis focus on the second unique feature of language: its capacity to convey concepts and meaning. The primary function of language is communication, and the extraordinary facility of language is purpose-built by a community of speakers to maximally convey complex and novel ideas and information with minimal ambiguity (Pinker & Jackendoff, 2005). No other synaesthetic inducer approaches this combination of specificity and variety of expression, and this may also be why words are such prominent inducers in synaesthesia. If this is the case, the meanings of individual words might influence the synaesthetic colours associated with them. The first systematic investigation of semantics in synaesthesia is presented in Chapter 4 using two types of words: colour terms (e.g. *red*) and words whose meanings evoke a clear mental image associated with them (e.g. *fire*). In Chapter 5, we take this a step further, and suggest that even colours for individual letters may be based on the meanings of associated words: that is, *A* is synaesthetically red because *A* is for *apple* and apples are red. Overall, these two studies investigate how the colour of a word based on real-world meanings (e.g. the orange of *fire* or the red of *apple*) can influence the colours that a synaesthete experiences for those words.

In Chapter 4, we explore how the colour typically associated with a word (i.e. its *canonical* colour, for example, blue for *sky* or orange for *fire*) may influence its synaesthetic colour. We begin with the clearest example of meaning-linked colour, namely the colours that are evoked by colour terms like *red, purple, beige*, etc. Somewhat surprisingly, the synaesthetic colours associated with colour terms have barely been touched on by research, and the two main studies that do exist did not take synaesthetes' individual colour associations into account. The first of these was a modified Stroop task conducted by Gray et al. (2002) with grapheme-colour synaesthetes investigating the *alien colour effect* (ACE). This effect describes the conflicting colour responses to a colour term that a synaesthete may experience: one based on the meaning (e.g. red for *red*) and the "alien" one based on the letters (e.g. purple for *red*, if the synaesthete has a purple *R*). Gray et al. (2002) asked synaesthetes to self-report whether they experienced an alien colour for four colour terms, *red, yellow, green,* and *blue*, with a higher degree of ACE denoted by more colour terms with an alien colour. They found that synaesthetes with higher degrees of ACE were slower to name colours in the Stroop task, even though the ink colours in the Stroop task were not matched individually to each synaesthete's colour associations. Subsequently, Gray et al. (2006) conducted an fMRI study contrasting synaesthetes who did report experiencing the ACE with synaesthetes who did not, as well as with non-synaesthetes. Compared to non-ACE synaesthetes and controls, ACE synaesthetes showed increased activity in the hippocampus and supplementary motor areas, which Gray et al. interpreted as indexing increased need for goal conflict resolution and inhibition of incorrect colour responses. This gives neurobiological evidence for the experience of a synaesthetic alien colour when processing colour terms. These two studies are the only systematic investigation of the synaesthetic colours for colour terms thus far. A few other studies have collected synaesthetic colour responses for colour terms, for example, Asano and Yokosawa (2012) in Japanese (e.g. red for 赤 "red", blue for 青 "blue", etc.), but these colour terms were not systematically contrasted with visually or phonologically similar words. Without this comparison, it is impossible to tell whether the synaesthetic red of 赤 or the word *red* is due to its meaning or to some other influence, such as its spelling.

This comparison is precisely what we did in the first experiment of Chapter 4. We collected the synaesthetic colours for twenty colour terms (e.g. *black, maroon*) as well as a list of control words matched on spelling, length, frequency, and imageability (e.g. *blade, matron*). We intended these control words to be "typically" coloured for synaesthetes – that is, they would evoke the usual, letter-based colour that synaesthetes would experience in the absence of any other semantic colour influences. We also included colour terms outside of the well-known eleven "basic" terms (Berlin & Kay, 1969), hypothesizing that the frequency of the colour terms might modulate the strength of their canonical colours. For example, a very high-frequency, basic colour term like *red* might have a stronger influence on the synaesthetic colour for that word than an uncommon colour term like *azure*. This experiment allowed us to directly measure how synaesthetes resolved the inherent conflict between the canonical colour (e.g. red for *red*) and the "alien", synaesthetic colour of the colour term (e.g. purple for *red*), and what aspects of these inherently coloured words affected their synaesthetic associations.

In the second half of Chapter 4, we expanded this idea to another list of words with a strong associated colour: those whose meanings make them high in *imageability*. Imageability is defined as the ease with which the meaning of a word evokes a clear mental image (e.g. high-imageability *fire* vs low-imageability *freedom*; Bird, Franklin, & Howard, 2001; Stadthagen-Gonzale & Davis, 2006). Experiment 2 looked at the synaesthetic colours of pairs of high vs low-imageability words, e.g. *fire* vs *fine*. These "colour-adjacent" words like *fire*, which strongly evoke an impression of colour but are not colour terms, create more interference in Stroop tasks than words unrelated to colour at all, but less interference than colour terms like *red* (Risko, Schmidt, & Besner, 2006; Schmidt & Cheesman, 2005). The few studies researching whether canonical colour is evoked directly during reading overall indicate that the canonical colour, e.g. the orange of *fire*, is indeed simulated to a sufficient degree to interfere with subsequent colour judgments (Connell, 2007; Connell & Lynott, 2009; Zwaan & Pecher, 2012). We accordingly wanted to find out whether the automatically evoked orange of *fire* or green of *mint* would change the synaesthetic colour for these words relative to the typical, letter-based colours of control words *fine* and *mind*. We expected that the high-imageability words would have a similar effect on synaesthetic colour as true colour terms, although perhaps to a more limited extent, as their colour association is not as explicit as those of colour terms. As a whole, this chapter provides the first experimental investigation of conceptual, canonical colour in synaesthesia, controlling for the individual colours that each synaesthete typically experiences.

Finally, Chapter 5 investigates whether word meaning can influence the development of synaesthetic colour associations even for letters of the alphabet. Large-scale studies of letter-colour associations, both in synaesthetes and non-synaesthetes, consistently find that particular letters and colours tend to be paired together above chance level, for example *A* with red, *D* with brown, and *Q* with purple (Jonas, 2010; Rich et al., 2005; Rouw, Case, Gosavi, & Ramachandran, 2014; Simner et al., 2005; Witthoft et al., 2015). Various studies have suggested these commonalities derive from innate shape-colour biases (Spector & Maurer, 2008, 2011), learned sequences from coloured childhood toys (Witthoft & Winawer, 2006, 2013; Witthoft et al., 2015), or linguistic factors such as frequency (Beeli et al., 2007; Smilek, Carriere, et al., 2007). All of these accounts are described in Chapter 5, but are limited in scope, and do not explain *why* particular letter-colour pairs, like *D* and brown (rather than, say, blue), occur so widely in both synaesthete and non-synaesthete populations. In Chapter 5, we tested an explanation for these consistent pairings that had been frequently suggested (e.g. Hancock, 2013; Spector & Maurer, 2011) but never systematically investigated: simply put, that *A* is red because *A* is for *apple*, and apples are red. In other words, these particular letter-colour pairs may be rooted in literacy acquisition, and mediated by words beginning with that letter that are frequently used in learning aids such as alphabet books. This means that, for example, the brown of *D* would come from the canonical brown of *dog*, and the purple of *Q* from the royal colour of *queen*.

The experiments in Chapter 5 explore how these words that are associated with letters of the alphabet, which we term *index words* (e.g. *apple, dog, queen*), may explain the connection to particular colours. Alphabet books and posters using index words are an extremely common feature of reading pedagogy (Nodelman, 2001) and have been shown to improve reading ability (e.g. Brabham, Murray, & Bowden, 2006; DiLorenzo, Rody, Bucholz, & Brady, 2011; Evans, Saint-Aubin, & Landry, 2009; Nowak & Evans, 2012). The question then becomes whether the association between letter and index word is fossilised into adulthood, and can therefore explain the association between letter and colour. To test this, we obtained associations from a very large cohort of participants for letters to words (e.g. "*A* is for…?"), for words to colours (e.g. "What colour is an apple?") and from letters to colours directly (e.g. "What colour is *A*?"). We compared the colours obtained via the index word route (*A* → *apple* → red) to the colours associated directly (*A* → red) to establish whether index words could predict the directly associated colour. This experiment demonstrated that not only may canonical colour affect synaesthetic colour associations in real time, as we showed in Chapter 4, but may even be fundamental to the formation of those associations in childhood (see Simner & Bain, 2013; Simner, Harrold, et al., 2009).

In summary, Chapters 4 and 5 of this thesis investigate how the meaning of a word, and particularly its canonical colour (e.g. the red of *red* and *apple*) may influence the synaesthetic colours of both words and letters. In each chapter, we review the relevant psycholinguistic and synaesthetic evidence in detail and elaborate on the theoretical and practical implications of our results. These two studies represent the first systematic accounts of canonical colour effects in synaesthesia and demonstrate the critical role of meaning in the development and experience of synaesthetic colour.

Summary

This thesis explores two fundamental aspects of language – its internal morphological and orthographic structure, and its ability to evoke and convey meaning – through the multi-coloured lens of grapheme-colour synaesthesia. I show that this type of synaesthesia is essentially and inextricably tied to language: it is indeed a fundamentally psycholinguistic phenomenon. Each chapter reviews the relevant literature in detail and presents empirical data to address the hypotheses. By demonstrating the systematic influences that linguistic features exert on synaesthetic colour, I also show how these synaesthetic experiences can be used to test questions about language processing in everyone. At the conclusion of this thesis, the general discussion elaborates on the implications of these experiments for both synaesthesia and psycholinguistics and calls for further expansion of the psycholinguistic investigation of synaesthesia.

A note on typographical conventions

It is common practice in the academic literature on language to use italics to identify a referential use of a word in text (e.g. my frequent examples *fire* and *red* above). However, at some points it has been necessary to explicitly distinguish the meaning of a word from its orthographic representation. In this introduction and in most of the subsequent chapters, I have made this distinction typographically between meaning and orthography using *italics* and "quote marks" respectively. The exception to this is Chapter 4, which is concerned with colour-evoking words such as *fire* and *red*. There is great potential for confusion in many uses of the same word with different intended meanings (e.g. *fire* could mean the concept of fire, the word spelled *F-I-R-E*, or the colour of fire in the real world), so I have attempted to clarify this as follows. I employ quote marks in Chapter 4 for the referential uses of words ("red", "fire") to emphasize that this specifically refers to their orthographic representation. That is, the combination of letters *R-E-D* is a word, "red", that refers to the visual experience of a red colour. I hope that this notation will assist the reader in differentiating the orthographic representation of a word versus its meaning.

Chapter 2

Processing compound words: Evidence from synaesthesia

Abstract

This study used grapheme-colour synaesthesia, a neurological condition where letters evoke a strong and consistent impression of colour, as a tool to investigate normal language processing. For two sets of compound words varying by lexical frequency (e.g., *football* vs *lifevest*) or semantic transparency (e.g., *flagpole* vs *magpie*), we asked 19 grapheme-colour synaesthetes to choose their dominant synaesthetic colour using an online colour palette. Synaesthetes could then select a second synaesthetic colour for each word if they experienced one. For each word, we measured the number of elicited synaesthetic colours (zero, one, or two) and the nature of those colours (in terms of their saturation and luminance values). In the first analysis, we found that the number of colours was significantly influenced by compound frequency, such that the probability of a one-colour response increased with frequency. However, semantic transparency did not influence the number of synaesthetic colours. In the second analysis, we found that the luminance of the dominant colour was predicted by the frequency of the first constituent (e.g. *rain* in *rainbow*). We also found that the dominant colour was significantly more luminant than the secondary colour. Our results show the influence of implicit linguistic measures on synaesthetic colours, and support multiple/dual-route models of compound processing.

Introduction

Synaesthesia is a familial condition (e.g., Ward & Simner, 2005) where the perception of a stimulus in one modality triggers an automatic secondary sensation in another (e.g., Simner, 2012b). Our study seeks to investigate natural language processing using grapheme-colour synaesthesia, where letters and numerals are perceived to have unique and consistent colours (e.g. *A* might be scarlet red or *7* might be leaf green; Rich, Bradshaw, & Mattingley, 2005; Simner et al., 2005; Ward, Simner, & Auyeung, 2005). Grapheme-colour synaesthetes also experience colours for whole words, and these colours are often systematically related to their synaesthetic colours for the component graphemes (Mills et al., 2002; Simner, Glover, et al., 2006; Ward et al., 2005). For example, a synaesthete with a red *M* may also experience the whole word *man* as red as well (Mills et al., 2002). It is this linguistic aspect of whole-word colouring in grapheme-colour synaesthesia we explore in the present study, especially as it relates to the colouring of compound words (described further below).

Grapheme-colour synaesthesia is estimated to have a prevalence of about 1% in the general population (Simner, Mulvenna, et al., 2006) and to account for 35 - 45% of all cases of synaesthesia reported (Novich et al., 2011). Many aspects of the condition have been investigated in recent years, including its behavioural characteristics (e.g., Hubbard & Ramachandran, 2005; Ward et al., 2005), neurological roots (e.g. Rouw & Scholte, 2010; Sperling, Prvulovic, Linden, Singer, & Stirn, 2006) and associated advantages for cognition (e.g., Pfeifer, Rothen, Ward, Chan, & Sigala, 2014; Price, 2009; Ward, Thompson-Lake, Ely, & Kaminski, 2008). Of particular interest to the current paper, Simner (2007) suggested that there may be a special role for language as a synaesthetic inducer, since linguistic stimuli like words and graphemes are the triggers in 88% of the total reported cases of synaesthesia (Simner, Mulvenna, et al., 2006). This study seeks to use grapheme-colour synaesthesia to answer psycholinguistic questions about compound words and to provide a tool for exploring the mutual influences of synaesthesia on language and vice versa (for a review of this approach, see Cohen Kadosh & Henik, 2007; Simner, 2007). In particular, we ask what the synaesthetic colours for compound words can tell us about how such words might be stored in the mind for all people. Below, we first review previous evidence for linguistic influences in grapheme-colour synaesthesia, and then we provide a brief overview of the psycholinguistic evidence to date for how compound words are processed in English.

Interaction of grapheme-colour synaesthesia and language

Many studies have already shown the close mutual influence that synaesthetic colour and language have on each other. By putting a symbol such as 5, ambiguous between S and 5, in different linguistic contexts – that is, with bias towards a letter reading (M U 5 I C) or a number reading (3 4 5 6 7) – case studies showed that the synaesthetic colour experienced depends on the grapheme's linguistic meaning in context, and not simply its shape (Dixon et al., 2006; Myles, Dixon, Smilek, & Merikle, 2003). Synaesthetes also show significant trends in the colouring of certain graphemes – for example, *A* is red more often than chance would predict (Rich et al., 2005; Simner, Glover, et al., 2006) and these trends are influenced by linguistic qualities like grapheme frequency. For example, high-frequency graphemes like *A* are likely to elicit higher frequency colour terms in English like *red* (Simner, Glover, et al., 2006; see also Emrich et al., 2002). Later, a study in German showed that when the elicited synaesthetic colour was broken down into hue, saturation, and luminance (HSL), grapheme frequency was positively correlated with synaesthetic colour luminance and saturation (Beeli et al., 2007). The luminance effect was also replicated in English (Smilek, Carriere, et al., 2007). These studies clearly show that synaesthetic colour associations are not haphazard but systematic, and often based on linguistic qualities of the trigger.

The linguistic influences on grapheme-colour pairings are also seen in the way synaesthetes perceive colour for whole words. Grapheme-colour synaesthetes tend to report that words can have a combination of different colours (Mills et al., 2002), but as mentioned above, the colour of the first grapheme generally dominates the word in some way. For instance, having a blue *F* would mean a blue emphasis to the word *fan* (Baron-Cohen et al., 1993), even though the colours for other letters in the word may also be perceived by that synaesthete. From synaesthete to synaesthete, this primary emphasis on the colour of words can come either by their first consonant (e.g., *fan* is the colour of *F*) or first vowel (e.g., *fan* is the colour of *A*) with the former being the most common (Simner, Glover, et al., 2006). Simner, Glover, et al. (2006) found that letters downstream in the word could influence colouring too; for example, in the word *ether*, the synaesthetically dominant colour of *E* was reinforced by a second *E* downstream in the word, evoking that colour more quickly and strongly than in a word like *ethos*, where the colour of *E* conflicted with the downstream *O* (Simner, Glover, et al., 2006).

These studies together reveal a complex but rule-based system of word colouring influenced by linguistic factors such as grapheme frequency, serial letter position, vowel/consonant status, grapheme repetition, and also by individual differences among synaesthetes. These linguistic influences in synaesthetic colouring also extend to non-alphabetic orthographies as well. We describe this here because it is possible to draw parallels with English compounding, the focus of our current paper. Hung, Simner, Shillcock, and Eagleman (2014) studied Chinese synaesthetes who experience synaesthetic colours for characters (i.e., the logographic writing units of Chinese). Hung et al. found that certain components of these characters, called *radicals*, influenced the colour of the character as a whole. For example, the character 櫻, meaning "cherry blossom", is a compound made up of the radicals 木, meaning "tree" (and providing semantic information for the whole compound), and 嬰, a character pronounced ying1 (providing the whole compound's phonetic pronunciation). In Hung et al.'s study, radicals on the right side of the compound (like 嬰 in 櫻) predicted the compound's overall luminance, whereas radicals on the left side (like 木 in 櫻) were marginally better for predicting its hue. Furthermore, semantic radicals on the left side of a compound, such as 木, marginally predicted saturation. This complex picture of how logographic radicals influence overall compound colouring may lead us to anticipate a similarly detailed situation in English compound colouring as well. We do note, however, that Chinese characters contrast in their storage and processing with English compound words in important ways, including their decomposability and mental representations; for instance, the radicals that form Chinese characters may be processed more analogously to letters within an English word rather than as words within a compound, and compound characters predominate in (written) Chinese, unlike in English (see e.g. Taft, Zhu, & Peng, 1999; Zhou, Marslen-Wilson, Taft, & Shu, 1999). Therefore, we may expect that synaesthetic colours for English speakers may reflect the characteristics of English compounds in particular, which we explain below.

Characteristics of compound words

In the current study, we look at synaesthesia in English compound words. Compound words in English are made up of two independent constituent words combined to make a new word, as in *rainbow* (i.e., *rain* + *bow*). These compounds are of special interest in lexical access research because their combined meanings and structure can be used to study how words are composed and represented in the mind (e.g., Taft & Forster, 1976). Several different types of theories have been proposed for how compounds are processed, which we test in our current study and so briefly review here.

Full-listing models of word processing propose that all words are stored in the mental lexicon as wholes, regardless of complexity. Lexical processing of compounds therefore consists of direct lookup of whole words in the lexicon (Butterworth, 1983). At the other extreme, full-parsing models claim that all complex words are decomposed into their constituents prior to lookup (Pinker, 1991; Stockall & Marantz, 2006; Taft, 1979, 1988, 2004). For example, a full-listing model would posit separate lexical entries for *rain*, *bow*, and *rainbow*, and the input *rainbow* would access that entry directly. A full-parsing model would posit that *rainbow* would first be obligatorily broken down into *rain* and *bow*, and those constituents would then be used to access the whole-word entry for *rainbow*. Combining the two are dual-route models, which suggest that both direct lookup and parsing routes work to process a word's representation. In particular, parallel dual-route models (Bertram & Hyönä, 2003; Schreuder & Baayen, 1995) propose that the two strategies race to the correct representation. More recently, dual-route models have been extended to probabilistic multiple-route models to account for information integrated from many sources during processing, including full forms, constituent words, morphological family size, and contextual and semantic cues (Kuperman et al., 2008; Kuperman, Schreuder, Bertram, & Baayen, 2009).

We aim to provide data to test these models using the synaesthetic colours of compound words. We present our synaesthetes with compounds that vary on two linguistic features that are often used to test models of compound processing – word frequency and semantic transparency. Word frequency expresses how often a word occurs in a given language, and studies show that reading times decrease as a word's frequency increases (e.g., Oldfield & Wingfield, 1965; Ellis, 2002). Compounds can be quantified in terms of their overall compound frequency (e.g., the frequency of the word *rainbow* itself), but also by the frequencies of their constituents – for example, the frequencies of *rain* and *bow* independently. Frequency effects have been used often in compound-word research, the rationale being that if constituent frequencies influence how quickly a compound is processed, we would conclude that the compound has been decomposed in some fashion. For example, studies have shown that compound processing is facilitated if the first or second constituent is high frequency (Bien, Levelt, & Baayen, 2005) and in particular, if the high frequency element is the second constituent/head (Juhasz, Starr, Inhoff, & Placke, 2003; also Andrews, Miller, & Rayner, 2004; Inhoff, Starr, Solomon, and Placke, 2008). These studies point to a model where constituents within a compound are activated during processing and therefore suggest that compounds are decomposed into their constituents, as per the full-parsing or dual/multiple-route models. The constituents that are higher frequency are recognised more easily and, by extension, facilitate access to the compound as a whole.

However, there is also evidence that whole-word frequency influences response times independent of the frequency of the compound's constituents (Baayen, 2005). Eye-tracking reading studies have found significant reductions in gaze times for compounds with higher compound (i.e., whole-word) frequency, an effect that appears at least as early as the facilitation effects for the constituents (Andrews et al., 2004; also in Finnish: Pollatsek, Hyönä, & Bertram, 2000). Further eye-tracking investigations in compound reading have shown that whole-compound frequency has a significant effect on reading times, even before the second constituent has been fully identified (Kuperman et al., 2008; also in Dutch, Kuperman et al., 2009). Together with the facilitation effects of the constituents, this points to a combination of whole-word lookup and constituent decomposition, wherein the processing system makes use of all available strategies to arrive at the correct meaning (e.g., Libben, 2006). We might therefore expect to find a similar type of frequency effect at both constituent level and global level in the synaesthetic colouring of compound words.

The second linguistic characteristic this study considers is semantic transparency, or how clearly the meaning of a compound word is related to the meanings of its constituents. For example, *birdhouse* is a relatively transparent combination of the meanings of *bird* and *house*, but the compound *hogwash* is fully opaque in that it is related in meaning to neither *hog* nor *wash*. Does transparency affect whether compound words are mentally decomposed during language processing? Together, Sandra (1990) and Zwitzerlood (1994) found differences between transparent and opaque compounds in a priming task, using semantic associates of the compound's constituents, e.g. *moon* for *Sunday*. Zwitserlood (1994) found priming effects with fully and partially transparent words (e.g., *birdhouse* and *jailbird*, respectively) but not with fully opaque compounds (e.g., *hogwash*). This provided evidence that opaque words may have less decomposition than transparent words.

However, the evidence on transparency has been mixed. Frisson, Niswander-Klement and Pollatsek (2008) found no effect of transparency at all on eye movements in compound reading in English (see also Pollatsek & Hyönä, 2005). On the other hand, Juhasz (2007) reported a main effect of transparency in gaze durations (see also Marelli & Luzzatti, 2012, for Italian). In each case researchers were again seeking factors that might influence whether and when compounds are understood via decomposition. Of particular interest for the current study, MacGregor and Shtyrov (2013) found evidence that compounds may be decomposed differently dependent on their frequency. In an EEG study involving opaque compounds, they found that higher compound frequency elicited a stronger mismatch negativity component (MMN; known to index both lexical frequency and the congruence of semantic combinations) as opposed to low compound frequency. This indicates a high degree of lexicalisation (i.e., whole-word storage and processing) for high frequency compounds. Together, the studies above point to both whole-word access and decomposition strategies (the latter less so for opaque words) in the processing of compound words. Hence, Libben (1998; 2006) suggests that the language system may utilise all possible avenues of understanding a compound's meaning, including constituent processing and whole-word lookup.

Current Study

The current study examines how word frequency and semantic transparency influence synaesthetic colouring of compound words. We consider both the number of synaesthetic colours triggered by different compound words (do they trigger one colour, or more than one?) as well as the nature of these colours (what is their saturation and luminance?). Our aim is to not only understand how linguistic features influence synaesthetic colours, but also to use synaesthesia to better understand models of compound processing. For words with low whole-word frequency (e.g., *lifevest*), we expect that the primary strategy for processing will be decomposition (Bien, Levelt, & Baayen, 2005), which will therefore activate the constituents of compounds (e.g., *life* and *vest*). The activation of these two constituents may cause synaesthetes to be more likely to give low-frequency compounds two colours. High-frequency compounds, however (e.g., *football*), may be processed more directly via whole-word lookup (Andrews et al., 2004; Kuperman et al., 2008), so we expect a single synaesthetic colour will be more likely for these types of words. In summary, our predictions are that high-frequency compounds may be more likely to trigger one synaesthetic colour rather than two, whereas high-transparency compounds may be more likely to have two synaesthetic colours rather than one. This would provide support for a model of two different routes in compound processing, by which compounds are more likely to be decomposed in the mind if they are low (vs. high) frequency.

These predictions are partly inspired by an unpublished study by Kubitza (2006), who reported a case study with a single German synaesthete. Kubitza found that higher-frequency compounds in German were more likely to receive a single colour than low-frequency compounds. As this study did not ultimately appear in the literature, we attempt to first confirm this effect in a larger group of synaesthetes and, at the same time, see whether it extends to English. Furthermore, we also hypothesise that we may find an analogous effect regarding transparency. In transparent compounds (e.g., *birdhouse*), the meaning of the compound is directly related to the meanings of the constituents, and previous studies suggest this may lead to processing via decomposition (e.g. MacGregor & Shtyrov, 2013). If so, we predict that the activation of these constituents may lead to a higher likelihood of two synaesthetic colours. Conversely, the meanings of opaque compounds (e.g., *hogwash*) cannot be calculated from their constituents, and studies show that this discourages decomposition (e.g. Ji et al., 2011), so we predict a higher likelihood that synaesthetes will give these compounds only one colour. This would support theoretical models of compound processing that propose two routes for lexical access, dependent on the semantic content of the compound (e.g., Zwitserlood, 1994)[2].

In the second part of our experiment, we also examine more closely the precise nature of the colours that synaesthetes perceive for these words. We follow Beeli et al. (2007), Smilek et al. (2007), and Hung et al. (2014) in focusing on the saturation and luminance of synaesthetic colours, as this allows us to compare our results for compound words directly to previous findings in the synaesthesia literature (e.g., for

---

[2] This footnote did not appear in the published article, but acknowledges a study that was under review at the same time as this chapter. Blazej and Cohen-Goldberg (2015) reported a case study of a grapheme-colour synaesthete, from whom they obtained colour associations for compound words. They found that the number of colours (one versus two) that their participant reported for compound words was influenced by the frequency of the constituents and the whole compound, as well as by the compound's semantic transparency.

graphemes; see above). We test whether, and under what circumstances, compounds might produce quantitatively different types of colours (e.g., colours with higher or lower luminance or saturation). For example, if whole-word frequency influences the nature of colours (e.g., if high whole-word frequency produces more luminant colours), this would support full-listing models of lexical access by showing influences only at the level of the whole word. However, if colours are influenced by the frequencies of constituents (e.g., if high constituent-frequency produces more luminant colours), this would show the influence of constituents within compounds and therefore support full-parsing models (or dual-route models if both types of frequency play a role). Finally, we also test whether transparent and opaque compounds produce different types of colours (again in their luminance or saturation). If, for example, transparent compounds are more likely to be decomposed during lexical access, we might find, say, additive luminances from two different constituents; this would support full-parsing models of lexical access by again suggesting decomposition during processing.

Methods

Participants

Nineteen grapheme-colour synaesthetes (17 female, mean age 24.8, SD = 11.3) were recruited from the Sussex-Edinburgh Database of synaesthete participants and paid £12 for their participation. All participants were native English speakers and were confirmed to be genuine synaesthetes using the gold-standard behavioural test as follows. Synaesthetic colour associations are highly consistent, and synaesthetes can therefore be identified using a consistency test (Baron-Cohen et al., 1987; Cytowic, 1989; Ward & Simner, 2003). Our diagnostic test was the online Synesthesia Battery at synesthete.org (see Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007, for methods). This test presents all 26 English graphemes three times in random order. For each grapheme, participants must choose their associated colour from a 16.8 million colour palette. The mean distance in colour space between the three colours given for each grapheme is converted into a standardised consistency score, with a score less than 1 indicating the high level of consistency characteristic of genuine synaesthesia. All 19 of our participants were below this required threshold (mean score = 0.62, SD = 0.13) and were therefore confirmed to have grapheme-colour synaesthesia.

Materials and Procedure

Our core test materials were two sets of compound words. The first group of words varied incrementally by frequency (high to low), and the second varied by transparency (opaque to transparent). The first set (N = 59 compounds) were drawn from Janssen, Pajtas, and Caramazza (2011) and varied on lemma frequency[3] measured by the CELEX lexical database (Baayen, Piepenbrock, & van Rijn, 1993). These compounds were stress-initial, two-syllable, noun-noun compounds (e.g. *rainbow*). The second list (N = 51 compounds), which varied by semantic transparency, were taken from a study by Ji, Gagné, and Spalding (2011). In that study, transparency was rated on a scale from 1 ("totally opaque") to 7 ("totally transparent") by 36 raters, and the means of these ratings were the final transparency score for each compound. Our two wordlists contained no items with repeating consonants or vowels in the onset or nucleus of the constituent words (e.g. none such as *crossbow*). These words were removed because repeated letters have been shown to influence the synaesthetic colour of the whole word in a way not relevant to our current investigation (Simner et al, 2006, see above). Table 1 below lists the descriptive information about the wordlists and measures. In the frequency list, there was a marginal correlation between compound and second-constituent frequency [$r = 0.25$, $p = .054$]; all other correlations were nonsignificant [$p > .46$]. In the transparency list, there was no correlation between transparency rating and compound frequency [$p = .74$]. Moreover, the items from Ji et al. (2011) were categorised as high or low transparency by Ji et al. and balanced between transparency conditions for lemma frequency [with a median split by transparency rating: $t(49) = -.018$, $p = .99$].

Table 1. Means and standard deviations of the variables in each wordlist. Here "Constituent" is abbreviated as "Const."

Measure	Compound Frequency		1st Const. Frequency		2nd Const. Frequency		Transparency Rating	
	M	SD	M	SD	M	SD	M	SD
Janssen et al. (Frequency)	2.4	1.67	7.08	1.34	6.48	1.66	-	-
Ji et al. (Transparency)	1.21	0.68	-	-	-	-	4.89	1.53

[3] Lemma frequency is the frequency of a word as it appears in all its inflexional variants (e.g., *rainbow*, *rainbows* etc.) and this was the frequency measure available in the set of norms from which we drew our materials (Janssen et al., 2011).

An online test was developed for the purposes of this experiment and participants were sent a link to this test via email. The test presented our frequency and transparency compounds separately, with the items randomised for each participant within each block. The order of the blocks was counterbalanced across participants. All target words were presented midway down the screen in bold (see Figure 1). The participants were required to indicate whether each word had synaesthetic colour, and then chose that colour using a clickable colour palette. They were then asked whether the word had a second synaesthetic colour, and used a second colour palette to specify that colour (see Figure 1). Therefore, participants could provide zero, one, or two colours for each compound.

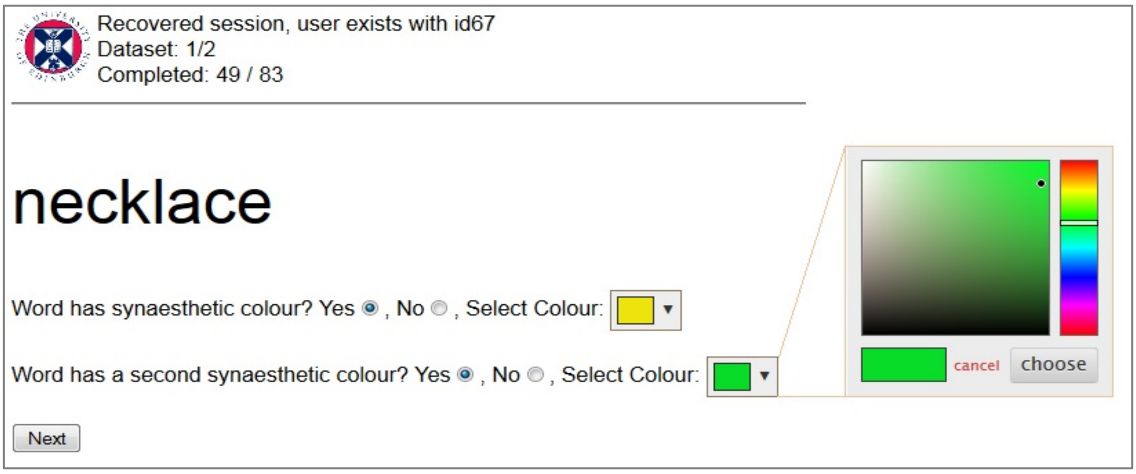


Figure 1. The online word colour test. The test item is presented in bold letters (here *necklace*). Participants indicate whether the word has synaesthetic colour(s), then select those colour(s) using the colour palette (shown in its expanded form to right).

Due to an oversight, the first four participants were given an option to skip the colour of any given word, which led to the loss of 47 responses (2.2% of the data overall). In the analysis, items that had been skipped were coded as uncoloured.

Results

In our study, we recorded two different dependent measures: the number of synaesthetic colours (zero, one, or two) that our synaesthetic participants provided for each compound, and the nature of those colours as measured particularly by their saturation and luminance values. We will address the results of these two separate measures in different sections below, and within each section, we will take into account the manipulation of frequency and then transparency.

The number of synaesthetic colours for compound words

The majority of the words in both lists were given two colours: 827 out of 1,121 (74%) in the frequency list and 720 out of 969 (74%) in the transparency list. In our analyses below, frequency and transparency are treated as continuous variables. However, for illustrative purposes only, Figure 2 below also divides our data into categories based on median splits of the frequency and transparency ratings, respectively.

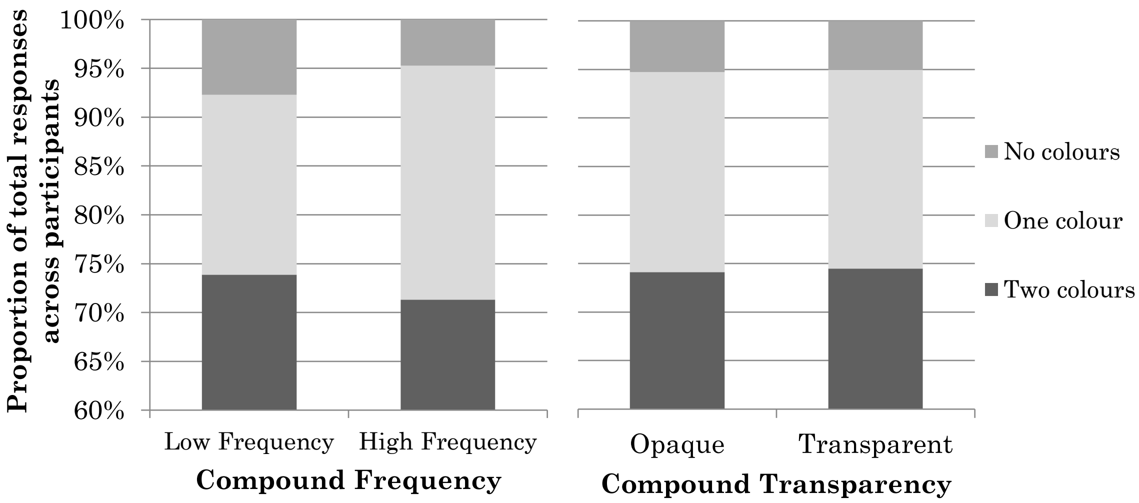


Figure 2. The proportion of zero, one, or two colours, collapsed across participants. In the left panel, the set of items varying by compound frequency are divided into groups of low (N = 30) and high (N = 29) frequency. In the right panel, the set of transparency items are divided into opaque (i.e. low transparency rating; N = 24) and transparent (i.e. high transparency rating; N = 25) groups. Both panels are on the same scale.

We hypothesized that in the set of items that varied by lexical frequency, synaesthetes would be more likely to experience one colour (instead of two) for words of higher (vs. lower) frequency. Hence, we first analysed the frequency wordlist to investigate the effect of frequency on number of colours. We constructed a binomial linear mixed effects model, which predicted the likelihood of a two-colour versus one-colour response, including compound frequency and first and second constituent frequency as predictors (e.g. the frequencies of *rainbow*, *rain*, and *bow*) and random slopes to account for the random variation in participants and items. Zero-colour responses were excluded from the model, as skipped and uncoloured items were both recorded as having zero colours, and it would therefore be difficult to draw any conclusions about this type of response. Overall compound frequency was a significant predictor of number of colours (see Table 2 below): as compound frequency increased, so did the likelihood of a one-colour response. First and second constituent frequency were not found to significantly influence the model. Table 2 below details the results of the linear mixed effects (LME) model, showing the influence of compound frequency on the likelihood of obtaining a two-colour response.

Table 2. LME model of frequency measures and number of colours. Here "Constituent" is abbreviated as "Const."

Predictor	Estimate	*z*	Random variance (item)	Random variance (participant)	*p*
Intercept	2.33147	4.141	0.02898	5.30781	**<.001**
Compound Frequency	-0.1606	-2.685			**.007**
1st Const. Frequency	-0.02898	-0.399			.69
2nd Const. Frequency	0.05815	0.974			.33

Our second hypothesis was that the number of colours would also be influenced by semantic transparency; higher transparency may make a two-colour response more likely because these compounds can easily be processed by splitting them into their constituents. However, this prediction was not supported by our data. There was no difference in the likelihood of one or two colours based on transparency ratings (see Figure 2, above). Table 3 summarises the linear mixed effects model showing the non-significant influence of semantic transparency on the likelihood of obtaining a two-colour response.

Table 3. LME model of semantic transparency and number of colours

Predictor	Estimate	*z*	Random variance (item)	Random variance (participant)	*p*
Intercept	3.3289	0.947	0.1565	10.6536	**<.001**
Transparency	-0.04041	0.814			.620

The linear mixed effects models above predict the binomial probability of a two-colour response. By transforming the log-odds into probabilities, we can represent the two models graphically as in Figure 3 below. It is clear from the left panel (frequency model) that the probability of a two-colour response drops as frequency increases, which means a one-colour response becomes more likely at higher compound frequencies. However, the horizontal line in the right panel (transparency model) shows that the increase of transparency rating has no meaningful effect on the probability of obtaining a two-colour response.

To summarise, our analyses did find a significant effect of overall word frequency on the number of reported synaesthetic colours, but no effect of constituent frequency or semantic transparency.

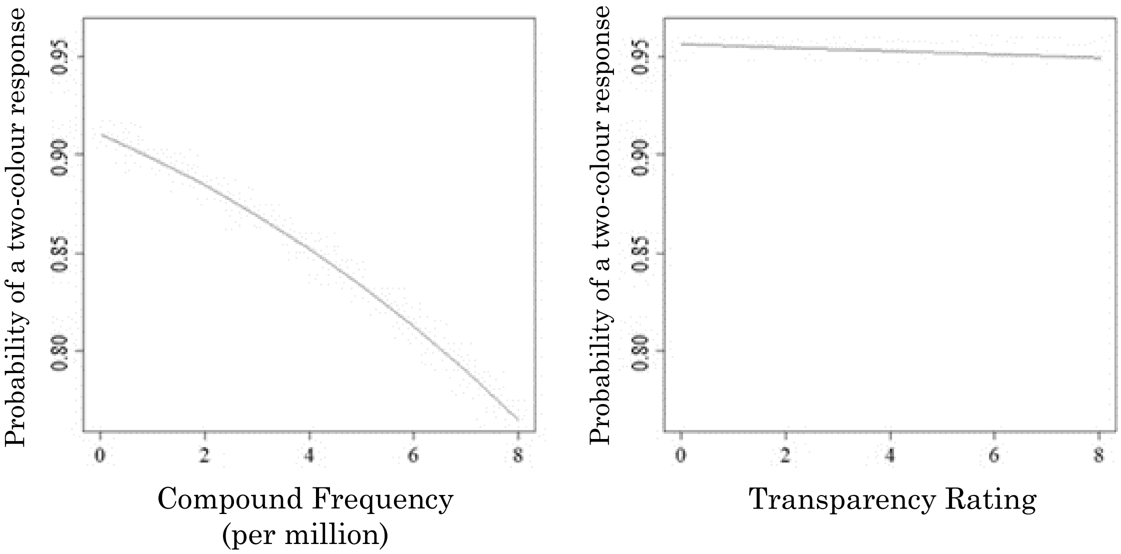


Figure 3. The probability of two-colour responses for sets of words varying by frequency (left) and transparency (right).

The nature of synaesthetic colours for compound words

The next analyses examine the nature of synaesthetic colours. This investigation focused on saturation and luminance values, as these were expected to show systematic variation with the variables of frequency and transparency (Beeli et al., 2007; Simner, Glover, et al., 2006; Smilek, Carriere, et al., 2007).

For each compound, there were up to four possible values: two possible colours per word, each having both saturation and luminance values. For example, participant 14 experienced the word *rainbow* with two synaesthetic colours, C1 and C2, with saturation and luminance values of 75, 61 (C1) and 79, 44 (C2) respectively. On the other hand, participant 1 experienced only one colour for this word, with saturation and luminance values of 33, 32 (C1). We calculated the average saturation and luminance values for the dominant and secondary colours elicited by each word across participants, which therefore gave us four mean values overall per compound (average saturation and luminance of both dominant and secondary colour). Looking first at the frequency manipulation, we constructed linear regression models predicting luminance and saturation values from (a) overall word frequency and the (b) first and (c) second constituent frequency. Of these models, only one showed a significant effect: we found a frequency effect on the luminance of the dominant colour of the word. Specifically, first constituent frequency significantly predicted dominant-colour luminance in a linear regression model (see Table 4 below). The relationship between dominant-colour luminance and first constituent frequency is depicted in Figure 4 below. No other predictors approached significance in regression models (all *b*s < 1.15, all *t*s < 1.67, all *p*s > .1).

Table 4. Linear regression model predicting dominant colour luminance in frequency word set

Predictor	*b*	*t*	*p*
Constant	33.98	5.61	**<.001**
Compound frequency	0.08	0.6	.552
First constituent frequency	0.38	3.05	**.004**
Second constituent frequency	-0.09	-0.7	.488

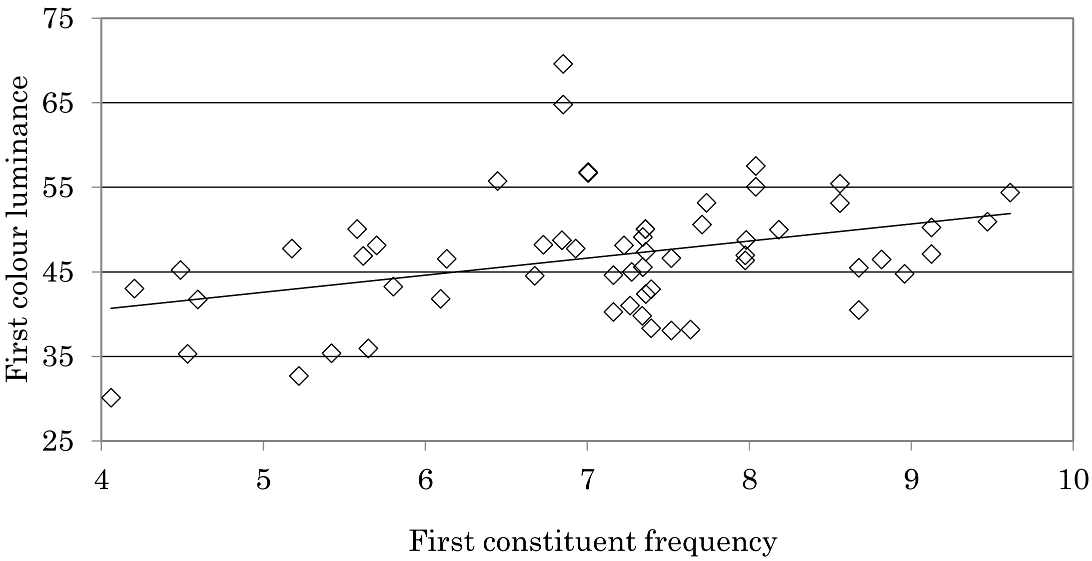


Figure 4. Scatterplot and regression line showing the relationship between mean first colour luminance and first constituent frequency

For the transparency wordlist, we chose a mixed repeated-measures ANOVA and divided the items into transparent and opaque groups taking the midpoint (4.0) on the rating scale as the point of division (transparent: N = 17, mean rating = 3.06, SD = 0.64; opaque: N = 34, mean rating = 5.80, SD = 0.88). We chose this simpler group design because it allows us to pursue an additional question: whether the average saturation and luminance values differed significantly between dominant and secondary colours. Since transparency was not found to influence number of colours in our first experiment, we suspected that we might not find an influence on the nature of those colours either. However, we could also investigate whether the two synaesthetic colours influenced each other, i.e. whether the saturation or luminance of the first, dominant colour was related to the saturation or luminance of the secondary colour in each word, aside from the effect of semantic transparency. Therefore, the mixed design ANOVA had two factors, each with two levels: transparency (transparent vs opaque) and colour dominance (dominant colour and secondary colour).

For the saturation values, the analysis showed no main effect of transparency [$F(1, 49) = 1.2$, $p = .27$] or of colour dominance [$F(1, 49) = 0.27$, $p = .61$] and no interaction between the two [$F(1, 49) = .10$, $p = .92$]. In other words, synaesthetic colours were equally saturated across transparent vs opaque compounds and across dominant vs secondary colours. For the luminance values, there was no main effect of transparency [$F(1, 49) = 0.06$, $p = .8$] and no interaction [$F(1, 49) = 2.01$, $p = .16$] but a marginally significant main effect of colour dominance [$F(1, 49) = 3.61$, $p = .06$]. Overall, the mean of the dominant colour luminance (M = 45.5, SD = 6.7) was higher than that of the secondary colour luminance (M = 42.5, SD = 5.5), with a mid-level effect size ($d = 0.49$;Cohen, 1988; Sawilowsky, 2009).

Given this marginal effect of colour dominance in the transparent compounds, we decided to look for this same effect collapsing over both frequency and transparency word lists, to allow greater power. This colour dominance effect is a result of the contrast between two synaesthetic colours and not specifically linked to any of the linguistic features above. Therefore, we conducted two repeated-measures ANOVAs using all of the compounds from our experiment in both the frequency and transparency sets combined (N = 110). For this combined wordset, we tested mean luminance and saturation values in dominant versus secondary colours. Although, as above, there was no effect in saturation values [$F(1, 109) = 1.596$, $p = .209$], there was a significant effect in luminance [$F(1, 109) = 9.708$, $p = .002$]. This indicates that the first-reported, dominant colour in a compound is significantly brighter (mean luminance = 46.2, SD = 7.0) than the secondary colour (mean luminance = 43.6, SD = 6.4; $d = 0.39$).

To summarise, we found further evidence of a frequency effect on the nature of the synaesthetic colours evoked by compound words. Specifically, we found that the luminance of the dominant word colour was significantly predicted by the frequency of the first constituent in that compound. As for transparency, we confirmed a lack of influence of semantic transparency on the colouring of compounds. However, we did find a significant effect of colour dominance on luminance when both compound lists were combined.

Discussion

This study explored how linguistic features can affect synaesthetic colours, and how synaesthesia can be used as a tool to test hypotheses about psycholinguistic phenomena (see Cohen Kadosh & Henik, 2007; Simner, 2007). We focused on frequency measures in compound words (how often a compound or its constituents occur in English) and semantic transparency (how transparently the meaning of a compound is connected to the meanings of its constituents) as a way to test how compound words are processed and stored in the brain. We collected the synaesthetic colours for two lists of compound words and found the following results: (1) the likelihood of experiencing one synaesthetic colour (vs. two) increases with compound word frequency; (2) semantic transparency has no effect on the number of synaesthetic colours; (3) increasing first constituent frequency significantly predicts higher luminance in the dominant synaesthetic colour; and (4) dominant synaesthetic colours are brighter (i.e. more luminant) than secondary colours. We will examine the implications of each of these findings in turn.

The finding that higher frequency compounds are more likely to be given a single colour than lower frequency compounds marries with the unpublished results of a single case-study reported by Kubitza (2006) in German, but here with a larger group of synaesthetes and in English. First, this result shows that psycholinguistic measures like word frequency do indeed influence synaesthetic colour responses, which can inform psycholinguistic theories. Specifically, our data suggest that high-frequency compounds would be more likely to be processed as wholes and would therefore be more likely to have a single synaesthetic colour. Although our frequency wordlist was not controlled for transparency, our results also showed that transparency has no appreciable effect on synaesthetic colour choice (see below), so this effect seems to stem entirely from the frequency of the whole compound. More importantly, however, we can extrapolate these results to provide evidence about how language processing occurs in non-synaesthetes as well as synaesthetes. The connection between single synaesthetic colours, high compound frequency, and lexicalisation of compounds suggests that even though compounds could be decomposed, those with high frequency are more quickly processed by direct lexical access (Kuperman et al., 2008). On the other hand, lower frequency compounds were more likely than high-frequency compounds to have two colours; this indicates that their constituents were more likely to be activated during processing, which in turn activated the two synaesthetic colours.

These results speak against a strictly full-parsing model in which all complex words, including compounds, would be broken down into their constituents preceding whole-word access (Stockall & Marantz, 2006; Taft, 1979, 1988, 2004). This would have predicted that the constituents would always be activated (regardless of compound frequency) and therefore would result in few to no one-colour compounds ever being found at all. Although we did find a preponderance of two-colour compounds overall, we also had the expected frequency effect for whole compounds (i.e., higher compound frequency predicts a greater likelihood of a single synaesthetic colour), which points to higher-frequency compounds, at least, being more likely to be parsed as whole words. On the other hand, in a full-listing model, positing a separate entry in the mental lexicon for all words (e.g. including *rainbow*; Butterworth, 1983), we would expect many more one-colour compounds and no constituent frequency effect. Our results are most compatible with dual/multiple-route models (e.g. Kuperman et al., 2009; Schreuder & Baayen, 1995), which would predict both one- and two-colour compounds with a frequency effect, reflecting the use of both lookup and decomposition strategies in processing. The compound frequency effect that we found for one- vs. two-colour compounds matches this model well, indicating that high-frequency compounds are indeed more likely to be processed via their whole form while low-frequency compounds may be processed via their constituents.

It may be noted as well that, although we might have expected to find a pervasive influence of the frequencies of the constituent words alongside the effect of whole-word frequency (following psycholinguistic studies by e.g., Bien et al., 2005; Inhoff et al., 2008), the frequency of neither constituent exerted a significant influence on the number of colours. Although both first and second constituent frequency have been shown to influence how quickly a compound is processed in English, it may be that the one/two colour binary outcome of our study was unable to capture this constituent effect. We did, however, find an effect of constituent frequency on the nature of the synaesthetic colours we measured (see below) and this may simply be a more fine-grained way to tap into this type of effect.

Our second finding was that compound transparency (cf. *birdhouse* vs *hogwash*) had no effect on the number of elicited synaesthetic colours. This was surprising, especially considering the analogous effect for word frequency, and may be a result of several different factors. First, since synaesthetes typically devote considerable care to reporting their colours precisely and accurately (e.g., Eagleman et al., 2007), the number of compounds in our item list was limited to minimise fatigue. Gathering synaesthetic colour data on a larger inventory of compounds varying by semantic transparency may provide enough data to reveal effects that our study could not capture. Alternatively, it is also possible that transparency is simply not a salient enough quality for it to influence synaesthetic colours. Little research has explored the effects of word meaning on synaesthetic colouring (see only Asano & Yokosawa, 2012; Gray et al., 2002). It is therefore possible that semantic transparency may not be strong enough to overcome the attested influences of grapheme frequency, colour term frequency, serial letter position, consonant/vowel status, and stress (Beeli et al., 2007; Simner, Glover, et al., 2006; Smilek, Carriere, et al., 2007; Ward et al., 2005) in addition to the frequency effect described in this current study.

In summary, our data did not show differences in the number of synaesthetic colours for compounds that were transparent (e.g., *keyhole*) versus opaque (e.g., *hogwash*) and so we have no evidence that only the former are fully parsed into constituents. However, a more refined interpretation of full-parsing models has been proposed by other authors (e.g. Juhasz, 2007; Taft & Ardasinski, 2006; Taft, 2004), in which every type of compound is necessarily decomposed regardless of transparency. In such models the constituents are always activated, and these constituents are equally activated regardless of whether they activate the whole word directly (as in the case of *keyhole*) or not (as with *hogwash*, which must receive direct activation from the form-level representation). Hence the lack of a transparency effect found here is compatible with this family of models also. Finally, frequency is captured in these models by the speed at which activation spreads from transparent constituents to whole-word representations, with this being faster for high (vs. low) frequency compounds (Taft & Ardasinski, 2006; Taft & Nguyen-Hoan, 2010). Our own frequency finding (i.e., high frequency compounds tend to take a single colour, rather than two) would therefore be interpreted in these models as single colours arising when whole words are activated quickly.

We further found that linguistic factors influenced the luminance of synaesthetic colours. An increase in first constituent frequency predicted an increase in dominant colour luminance. In other words, compounds with first constituents that are encountered more often in English (e.g. high frequency *hand* in *handcuffs*) have a brighter dominant synaesthetic colour than those seen less often (e.g. *cork* in *corkscrew*). Given the relationship between frequency and luminance in other areas of synaesthesia (e.g. in grapheme colours; Smilek et al., 2007), it is not surprising to find an analogous effect in word colour. We also speculate that the more dominant colour within any compound word may be related directly to its first constituent. This is because first constituents are important in compound processing in English (Andrews et al., 2004) and also in other languages (e.g., Finnish; Pollatsek et al., 2000). That is, we speculate that synaesthetic brightness reflects the first constituent's important role in processing. We are now comparing the synaesthetic colours of compounds and constituents in follow-up studies in our lab [see Chapter 3].

Finally, we found a significant effect across all compounds for the dominant colour to be brighter than the secondary colour. This brighter dominant colour in a compound appears to reflect the prominent role that the "dominant colour" has by definition. In other words, the fact that synaesthetes are able to identify which colour in a word is more dominant at all may well stem from that colour being overall brighter. This also marries well with our finding described above, that higher frequency first constituents are more luminant. These parallel findings involving luminance point to a complex but systematic picture of synaesthetic word colouring, influenced both by linguistic factors (such as word frequency) but also by the contrast between the colours within a single word. Future studies should address the source of these luminance effects and how constituent frequency and ordinal position interact to influence the luminance of the dominant colour.

In conclusion, our study has shown that synaesthesia can be used to investigate questions about how words are stored in the brain and to evaluate existing theories of word processing. This initial confirmation of the influence of frequency on synaesthetic colour is a starting point to consider in more depth the questions of how meaning and language influence synaesthetes, and therefore how these things affect cognition in all of us.

Acknowledgements

This research was supported in part by a University of Sussex PhD research support grant. HB was supported by a British Academy/Leverhulme Senior Research Fellowship. The authors would like to thank Dr. Martin Corley for his help with our analyses.

Chapter 3

Synaesthetic colours for simplex and complex words

Abstract

Research on grapheme-colour synaesthesia (i.e. automatic colour experiences for linguistic symbols, e.g. *A* is red, *4* is yellow) has successfully established the genuineness of the phenomenon (Hubbard & Ramachandran, 2003), but has thus far focused primarily on the colours of individual letters. This study investigates for the first time how the colours of letters propagate to simple whole words (e.g. how the colours of *R, A, I,* and *N* influence the colour for *rain* as a whole), and from simple to complex words (e.g. how *rain* and *bow* colour the compound *rainbow*). We obtained the synaesthetes' colours for letters, simple words, and compound words (e.g. *R, A, I, N, B, O, W, rain*, *bow*, and *rainbow*). In the first study, we use colourspace measures to show how the colours of words like *rain* systematically derive from the colours of particular letters in the word. We also show that word frequency influences the extent to which words are coloured by their letter colours. We then expand on Mankin, Thompson, Branigan, and Simner (2016) by showing that the strongest, most dominant colour of a compound word (e.g. *rainbow*) typically comes from its first constituent morpheme (*rain*), and the secondary colour of the compound from the second constituent morpheme (*bow*), indicating that synaesthetic colours reflect underlying morphological structure. Finally, we find that the frequency of both the first constituent morpheme and the compound itself (e.g. the frequencies of both *rain* and *rainbow*) influence whether the compound is coloured like its first constituent (*rain*) or like its letters (e.g. *R*). These results point to a system of synaesthetic word colouring based on multiple hierarchical linguistic factors, and this strengthens the case that synaesthetic colours are a useful and informative tool for investigating linguistic processing more widely.

Introduction

Grapheme-colour synaesthesia is a neuropsychological condition where graphemes such as letters and numbers are automatically and consistently associated with highly specific colour experiences. For example, *H* may be perceived as intrinsically yellow, which can lead to the word *house* also being perceived as yellow (Ward et al., 2005). This type of synaesthesia is both easily testable and one of the most common subtypes of synaesthesia (estimated between 35-45% of all cases of synaesthesia; Carmichael, Down, Shillcock, Eagleman, & Simner, 2015; Simner et al., 2006). For decades, since synaesthetic experiences were first objectively verified as genuine (Baron-Cohen et al., 1993), researchers have been studying how synaesthetes acquire colour associations for particular letters, and how these consistent letter-colour associations are related to the colours those synaesthetes may experience for whole words (e.g. Simner et al., 2006; Ward et al., 2005; Mills et al., 2002). This study investigates the second question using the colours associated with both letters and words by a group of grapheme-colour synaesthetes. We describe how synaesthetic colour propagates from letters to simple words (e.g. how *R* or *A* may influence the colour of the whole word *rain*), and then how the colours of those words may combine to form the colours of compound words (e.g. how the colours of the constituent words *rain* and *bow* are related to the colours for the compound word *rainbow*). In so doing, we also suggest that we can use these colour patterns to understand how simple and complex words are stored and processed in the brain for everyone. To begin, we will review the current theories on the processing of both simple and complex words in turn and how we can investigate these questions using the synaesthetic colours of letters and words.

How do the letters in a word influence its processing?

If synaesthetic experiences are built on general cognitive and linguistic processing, we would expect to find parallels between processes of visual word recognition (VWR) and the colours a synaesthete experiences. In particular, we might expect that the specific letters important in word recognition (e.g. initial letters) might also be important in synaesthetic word colouring. Theories of VWR in English suggest several different mechanisms of recognising a word from its letters. In serial models, English words are processed from the beginning onwards in a left-to-right scan (e.g. Taft, 1979; Whitney, 2001), which would suggest that the colour of the first letter would be activated first and most strongly for synaesthetes. Conversely, it could be that since the later letters are key to word recognition (i.e. *rain* is recognised only when *N* is recognised, as *rai_* could also be *raid, rail*, etc.), words would take the colour of their later, most recently recognised letter. On the other hand, parallel models suggest that all the letters are processed simultaneously (e.g. Adelman, 2011; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Gomez, Ratcliff, & Perea, 2001), which at first glance would not predict any particular colour to predominate in the colouring of words. Given that first fixations on letters in English words typically land to the left of centre within the word (Rayner, 1979; Shillcock, Ellison, & Monaghan, 2000), any letter in the word might contribute its colour to the whole word, but particularly those near the region that is fixated on first.

However, these predictions for word colour assume that all letters in a word have equal influence on processing, and this is consistently not the case in VWR studies. Rather, across a variety of testing paradigms, the letter in the first position (e.g. *R* in *rain*, *B* in *bow*) has a processing advantage and appears to disproportionately contribute to word recognition in English (Aschenbrenner et al., 2017; Chambers, 1979; Grainger et al., 2016; Johnson & Eisler, 2012). Transposed-letter paradigms, which switch the position of at least two letters in the word (e.g. "anwser" for *answer*) show that these jumbled words can still be read, and even semantically prime related words (e.g. transposed "jugde" primes *court*; Perea & Lupker, 2003). However, they are read more slowly than unscrambled words (Rayner, White, Johnson, & Liversedge, 2006), and this cost is most severe when the initial letter is transposed (White, Johnson, Liversedge, & Rayner, 2008). Critically, the initial-position advantage is consistently reported for strings of letters and numbers (i.e. graphemes) but does not appear for strings of symbols (e.g. §<% £). For example, participants are better able to recall the identity of a symbol in first position when that symbol is a letter or number than when it is a non-linguistic symbol (e.g. §), indicating that this advantage is a particular feature of recognising meaningful strings such as words (Chanceaux & Grainger, 2012; Scaltritti & Balota, 2013; Tydgat & Grainger, 2009). If the first-letter advantage is reflected in synaesthetic colouring, we would expect that whole-word colour is strongly influenced by the colour of the first letter. This is in fact the case – synaesthetes frequently report that the first letter colour dominates the colour of the entire word, so *mother* would tend to take the colour of *M* (Baron-Cohen et al., 1987; Mills et al., 2002). A study by Ward et al. (2005) found that the first letter matched the colour of the whole word for four of the seven synaesthetes they tested. It seems that, if *R* is purple, then a synaesthete may indeed see purple *rain*.

Although the first-position advantage does nicely explain why the first letter colours the whole word for synaesthetes, this is still not the complete picture. A minority of synaesthetes experience whole-word colour based on the first vowel specifically, even in words with initial consonants (e.g. *rain* coloured like *A* rather than *R*). This pattern was reported by one of the synaesthetes in Ward et al.'s (2005) study, with the two remaining synaesthetes showing influences of both consonant/vowel status and position in the word (e.g. the first consonant might be more likely to contribute its colour when it was also the first letter in the word rather than second, as in <u>*m*</u>*other* vs. *u*<u>*n*</u>*cle*). A subsequent study by Simner, Glover et al. (2006), reporting a detailed case study of a synaesthete with whole-word colours based on vowels, emphasised the importance of linguistic information in synaesthetic word colouring. For pairs of words contrasting by stressed syllable (e.g. **'*ca***-*non* versus *ca*-**'*det***), the whole-word colour matched the vowel colour of the stressed vowel, even when it was not in the first syllable (e.g. *ca*-**'*det*** was the colour of *E* rather than *A*). This was the case for both spoken and written stimuli, strongly suggesting that syllabic information was part of both lexical access and synaesthetic colour response. They also found that the synaesthete was faster to name the colour of a word when the vowel was repeated in subsequent syllables downstream than when it differed (e.g. *ether* was given a colour faster than *ethos*). Finally, they reported that in a sample of 20 synaesthetes, three of them reported word colours systematically based on the first vowel. This suggests that at least for some synaesthetes, the category of the letters (i.e. whether they are consonants or vowels) is critically important, and at least as influential, if not more so, than position in the string. Although this pattern of colouring has been documented multiple times in the synaesthesia literature, none of these studies have commented on the fact that this vowel-based pattern of colouring is entirely unpredicted by prominent models of word reading based on transposed-letter effects (e.g. the spatial coding model, Davis, 2010; overlap model, Gomez et al., 2008; open bigram model, Schoonbaert & Grainger, 2004; SERIOL model, Whitney, 2001). None of these explicitly differentiate between consonants and vowels in VWR. We must then ask, is the preference for vowel-based whole-word colours simply a quirk of synaesthetes in particular – and therefore not relevant for a psycholinguistic investigation of word colouring – or might it point to an important aspect of VWR as a linguistic process?

There is ample evidence – besides the clear pattern in synaesthetes – to suggest that the consonant/vowel distinction is important in VWR. First, evidence from aphasics with contrasting difficulties producing consonants versus vowels indicates that they may be represented by different neural structures (Caramazza, Miceli, Chialant, & Capasso, 2000). Lee, Rayner, and Pollatsek (2002) showed that prime words with the same consonant structure as the target (e.g. *lake – like*) shortened gaze times more than primes with the same vowel structure (e.g. *line – like*), suggesting that consonants are processed more rapidly than vowels in word recognition. Studies using event-related potentials (ERP) have shown that the timecourse and location of word recognition differs depending on whether the manipulated letters are consonants or vowels (Carreiras, Vergara, & Perea, 2007, 2009). Furthermore, fMRI evidence suggests that detecting nonwords created by substituting consonants or vowels requires different processes: consonant substitutions activated lexical decision-making areas of the brain, while vowels activated prosody-related regions (Carreiras & Price, 2008). As the identity of a letter as a consonant or vowel may indeed be important to the word recognition process, it is not surprising to find that synaesthetes also are sensitive to this distinction.

In the present study, we will verify the first-letter versus first-vowel source of word colour using colourspace distances, and we will further investigate when and why these particular letters are used as the source of whole-word colours. We will also take into account the interaction between the physical properties (e.g. saturation and lightness) of the colours of the letters and their frequency. Higher-frequency graphemes (e.g. *A, S*) tend to have synaesthetic colours with higher luminance and saturation in German (Beeli et al., 2007) and higher luminance in English (Smilek, Carriere, et al., 2007) compared to lower-frequency graphemes (e.g. *X, Q*). For our current study, this suggests that the saturation and lightness of the letters, along with their frequency, may play a role in whole-word colouring.

How do the constituents in a compound word influence its processing?

As this study is concerned with the synaesthetic colours of both simple and complex words, we turn next to the question of how complex words – in this case, compound words such as *seahorse* or *rainbow* – may be coloured synaesthetically. We further expect that these colours may reflect how compound words are stored and retrieved in normal language processing. Compound words are of particular interest in studying lexical retrieval because they are themselves composed of independent constituent words. This means that the characteristics of the constituent morphemes themselves as independent words (e.g. their frequency, familiarity, or neighbourhood size) may influence their relationship to each other and to the compound word as a whole (Libben, 2006). Consequently, the models of compound processing differ primarily in the degree (and timecourse) of activation of the whole compound and its constituent morpheme representations. That is, when reading *rainbow*, is the word directly recognised as the whole compound *rainbow*, or is it accessed as *rain* and *bow* separately first? This influence of the constituent morphemes in the processing of a compound word is of critical importance to the current study because we expect that a compound's morphemes might contribute their synaesthetic colours to the colour of the whole compound based on the underlying structure of compound words. The following section will review the major theories that may explain the synaesthetic colours for compound words. The reader will kindly note that we will refer to the constituent words of a compound (e.g. *rain* and *bow* in *rainbow*) as "constituent morphemes" or simply "morphemes".

Theories of compound processing can be broadly divided into three categories: prelexical, supralexical, and multiple-route. Prelexical theories propose that compound words are always necessarily decomposed into their constituent morphemes, and only then is the meaning of the whole compound accessed (Stockall & Marantz, 2006; Taft, 2004). Conversely, supralexical theories postulate that the compound is accessed directly as a whole, and the constituent morphemes are accessed subsequently (Butterworth, 1983; see also Giraudo & Grainger, 2001, for affixed words). Multiple-route models propose that both whole-word access and decompositional processes operate simultaneously, weighted by factors such as frequency and familiarity (Kuperman et al., 2009; Schreuder & Baayen, 1995). These theories are often investigated using the well-established frequency effect (Oldfield & Wingfield, 1965), where items that occur with higher frequency in the language are responded to more rapidly across a variety of tasks. If the frequency of a constituent morpheme has an influence on a task using the whole compound (e.g. if high-frequency *rain* makes it easier to read *rainbow* in comparison to other compounds of the same compound frequency), this is taken as evidence that the lexical representation of the constituent morpheme *rain* is activated during the course of processing the whole compound.

We begin with the evidence that the constituent morphemes of a compound word – e.g., the *rain* and *bow* of *rainbow* – are activated as separate words during the course of processing the whole word *rainbow*. Eye-tracking in Finnish (Hyönä & Pollatsek, 1998) showed that when the first morpheme in a compound was higher in frequency, the whole compound had shorter initial fixation and gaze duration. A follow-up study found a similar facilitation effect for higher-frequency second constituent morphemes (Pollatsek et al., 2000). Using a similar paradigm in English, Andrews, Miller, and Rayner (2004) found that the frequencies of both first and second constituent morphemes influenced whole-compound processing. These studies all provided evidence that compounds were morphologically decomposed during processing – that is, that reading *rainbow* necessarily activates the representations of *rain* and *bow* as separate words. Subsequent research in morphologically diverse languages have commonly shown that the morphemes within a compound influence that compound's processing in a variety of experimental paradigms (in English, e.g. Bien et al., 2005; Inhoff et al., 2008; Ji, Gagné, & Spalding, 2011; Juhasz et al., 2003; Libben, Gibson, Yoon, & Sandra, 2003; Libben & Jarema, 2006; MacGregor & Shtyrov, 2013; in Basque and Spanish, e.g. Duñabeitia, Laka, Perea, & Carreiras, 2009; Duñabeitia, Perea, & Carreiras, 2007; in Italian, e.g. Marelli, Aggujaro, Molteni, & Luzzatti, 2012; Marelli & Luzzatti, 2012; in German, e.g. Bronk, Zwitserlood, & Bölte, 2013; Lorenz & Zwitserlood, 2014; Zwitserlood, 1994). This lined up well with evidence from other types of complex words, such as affixed words (e.g. *re-* + *heat* = *reheat*), to indicate that words are decomposed into their morphemic constituents during processing (Stockall & Marantz, 2006; Taft, 1979, 1994, 2004; Taft & Ardasinski, 2006; Taft & Forster, 1975, 1976). MEG studies with compound words (Fiorentino & Poeppel, 2007) also found evidence for early decomposition into constituent morphemes. This "obligatory decomposition" would mean that *rainbow* is always necessarily processed via its constituent morphemes *rain* and *bow* (Marantz, 2013). On the whole, given this overall weight of evidence that the constituent morphemes of a compound are activated during processing, we expect that synaesthetic colouring will reflect this underlying structure. Therefore, when processing a word like *rainbow*, the synaesthetic colours of the words *rain* and *bow* should be individually evoked. Under a single-route, full-decomposition prelexical model like that proposed by Stockall and Marantz (2006), the whole compound *rainbow* should not evoke its own synaesthetic colour independent of *rain* and *bow*. In summary, the well-evidenced influence of constituent morphemes on the processing of the whole compound would predict that compound words such as *rainbow* may have synaesthetic colours directly derived from the colours of their constituent morphemes.

There is also, however, evidence that the frequency of the compound as a whole also modulates its processing, beyond the influence of its individual constituent morphemes. This influence of the whole compound's frequency may indicate *lexicalisation*. That is, high-frequency complex words may be consolidated into a single lexical entry, analogous to the way that high-frequency simplex words become unitised as sight words rather than recognised as a combination of letters (Ehri, 1987; Ehri & Wilce, 1983). Following early work on plural affixes in Dutch (Baayen et al., 1997; Schreuder & Baayen, 1995), both Pollatsek et al. (2000) and Andrews et al. (2004) reported that for compound words, higher frequency of the whole compound had a similar effect that higher frequency of the constituent morphemes did (i.e. reduced gaze duration). Indeed, many of the studies cited above that reported effects of constituent morpheme frequency in compound processing also reported that the frequency of the whole compound had an influence in addition to, or in interaction with, the frequency of the constituent morphemes. However, the influence of compound frequency is not always straightforward: for example, Bien et al. (2005) also reported that compound frequency was a significant influence on response times, but while lower frequencies made response times quicker, higher frequencies tended to rather slow response times. While there is widespread agreement that the constituent morphemes in a compound are accessed during processing, the existence of a lexicalised representation of the compound as a whole that can be accessed directly is still under debate.

If the whole compound does have its own entry in the mental lexicon apart from its constituents, we can also expect to find synaesthetic evidence of this whole-word representation. That is, if there is an entry in the mental lexicon for the word *rainbow* as a whole word, we expect that it would derive its synaesthetic colour from its letters as we described for simplex words, above. For the majority of synaesthetes who derive the whole-word colour from the first letter, this means *rainbow* might be coloured like *R*, rather than like *rain*. The key distinction lies in which compounds are coloured like their first constituent morpheme (i.e. processed via a decomposition route), and which are coloured like their dominant letter (i.e. processed via whole-word lookup). Multiple-route models of compound processing predict that we will find evidence of both these colouring processes, depending on the frequency of the constituent morphemes and of the whole compound (Bertram, Schreuder, & Baayen, 2000; Kuperman et al., 2008, 2009; MacGregor & Shtyrov, 2013). That is, both constituent morphemes *rain* and *bow* would be activated during processing the whole-word representation *rainbow*, and accordingly the synaesthetic colours of *rain*, *bow*, and *R* would also be evoked. However, a higher-frequency whole compound like *rainbow* might be more likely to be lexicalised, or stored as a single unit, and therefore its synaesthetic colour would be more similar to *R* than to *rain*[4]. Lower-frequency compounds like *seahorse* may be less likely to have a strongly lexicalised representation, so their compound colour will be more similar to the colour of the constituent morpheme *sea* rather than *S*. However, we may find that this is modulated by the frequency of the first constituent itself. For compounds with a high-frequency first morpheme (e.g. *hand* in *handcuffs*), the strongly evoked synaesthetic colour of high-frequency *hand* may dominate the colour of the whole compound. The colour dominance of a high-frequency constituent morpheme may be particularly evident if the whole compound itself is low frequency (as is *handcuffs*). Some initial work in synaesthesia suggested that synaesthetic colour does reflect lexical structure; case studies of single synaesthetes in German (Kubitza, 2006) and English (Blazej & Cohen-Goldberg, 2015) have shown that these synaesthetes are more likely to give a compound one colour rather than two when that compound is higher in frequency. However, no investigation has yet addressed the source of these effects by comparing of the synaesthetic colours of the compound word's constituent elements. The current study does this for the synaesthetic colours of compound words, such as *rainbow,* to both the colours of that compound's constituent morphemes (e.g. *rain* and *bow*) as well as its letters (e.g. *R, A,* etc.). We expect that compound colour will often derive from constituent colour, but this may be modulated by the frequency of the two constituent morphemes and the compound as a whole.

Aims

The current investigation will therefore investigate the sources of the colours that synaesthetes experience for both simplex and compound words. This study builds

---

[4] We note here the important consideration that if *rain* is also coloured like *R* due to the letter-to-word colouring process for simple words we described above, it would be difficult to distinguish whether the colour of *rainbow* comes from *rain* or from *R*. We will address this issue directly in our analyses, below.

directly on the results of Mankin et al. (2016 [Chapter 2]) who investigated the colours that synaesthetes reported specifically for compound words. That study obtained colours from synaesthetes for two lists of compound words. One varied on frequency (e.g. high-frequency *rainbow* versus low-frequency *seahorse*) and the other on semantic transparency (e.g. transparent *birdhouse*, whose overall meaning is transparently related to its morphemes, versus opaque *hogwash*). Synaesthetes were asked to give the "strongest, most dominant" synaesthetic colour for each compound (termed the *dominant colour*), and could also report an additional synaesthetic colour if they chose to (termed the *secondary colour*). The results included four major findings. First, higher-frequency compounds were more likely to have one synaesthetic colour rather than two compared to lower-frequency compounds. This points to a connection between lexical frequency and synaesthetic colour experiences. Second, however, semantic transparency did not influence the number of synaesthetic colours. Third, the dominant synaesthetic colour in the compound was found to be more luminant (i.e. lighter) overall than the secondary colour. Finally, the dominant colour's luminance increased with the frequency of the first constituent. This suggested a link between the first constituent and the dominant colour.

The current study will explore the propagation of synaesthetic colouring using a similar experimental setup and a subset of the same participants doing additional tasks. First, we will ask how the colours of individual graphemes are related to the colours of simplex words (here, the constituent words of the compounds used by Mankin et al., 2016). Next, we will investigate how those constituent words pass on their colours to compound words. That is, we will see how the colours reported for *R, A, I, N, B, O*, and *W* relate to the colours for *rain* and *bow* as simple words, and we will then describe the relationship between the colours for the constituents *rain* and *bow* and the colour of the compound *rainbow*. Although these various aspects of letter- and word-colouring have been investigated by previous studies in isolation, this is the first study to document synaesthetic colouring at increasing levels of linguistic complexity with the same participants and words. We expect to find a complex interaction between linguistic variables, such as compound and constituent frequency, and colour elements, such as how similar and different colours combine within a single word. This novel investigation of the relationships between colours across the grapheme → constituent → compound hierarchy will expand our understanding of the mechanics of synaesthesia and effectively utilise synaesthetic colours to investigate psycholinguistic questions.

Study 1: Colour propagation from letter to word

Methods

Participants

We recruited 13 grapheme-colour synaesthetes to participate in our study. All had previously taken part in Mankin et al. (2016 [Chapter 2]). One of these participants was excluded because too many of their letters lacked colours, so we were unable to compare their letter-colours to word-colours reliably. Of the remaining 12 participants (mean age = 24.41, SD = 9.39), 11 were right-handed and 11 were female. All reported experiencing colours for graphemes and had been previously verified as grapheme-colour synaesthetes using the online diagnostic test known as the Synesthesia Battery (described in more detail below; see Eagleman et al., 2007). All participants scored below 1 on this test, which is the widely accepted cutoff to establish a synaesthete as genuine (average score = 0.60, SD = 0.15). This study was given ethical approval by the local university ethics board.

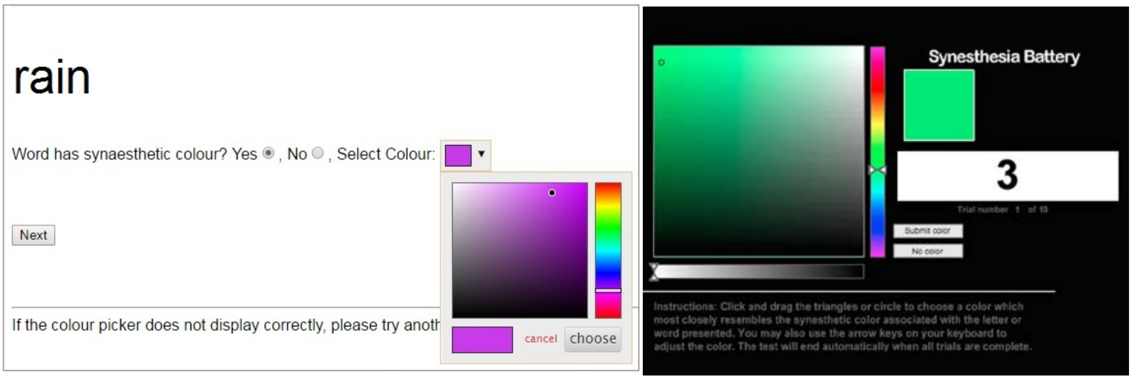
Materials

The words included for this test were the constituent morphemes of the compound words used in the frequency wordlist in Mankin et al. (2016; see Chapter 2). For example, Mankin et al. used the word *rainbow* as a stimulus word, so the present study included *rain* and *bow* separately. After excluding repeats of words that appeared more than once as constituents (e.g. *bow* appeared in both *rainbow* and *bowtie*, but was only included in the testing list once), our list of constituent words comprised 92 words. These words were high in Zipf frequency (M = 4.51, SD = 0.69) and imageability (M = 497.88, SD = 194.84). All words were monosyllabic, with a mean length of 4.10 letters (SD = 0.81; minimum length = 3; maximum = 7). Most words were also monomorphemic, with four exceptions: *clothes* from *clothespin*; *cuffs* from *handcuffs*; *phones* from *headphones*; and *muffs* from *earmuffs*. As we expected that synaesthetes would colour words predominantly by their first few letters, we included these words in the analysis in their original plural forms in order to match the original compound list as closely as possible, and to avoid any changes in connotations (e.g. *clothes* → *cloth*). The full list of the constituent words, along with the psycholinguistic measures for our analyses for both this and the following study, can be found in Appendix A.

Apparatus and procedure

Participants were invited by email to participate in an online test in the same style as the one they had previously completed. Once participants gave informed consent and provided demographic information, words were presented one at a time in the middle of the screen once each in unique random order for each participant. Participants chose their synaesthetic colour for each word from an expandable palette offering 16.8 million colour choices. They were required to select a colour, or indicate that they experienced no synaesthetic colour, before they could proceed to the next word. This test was identical to the one used by Mankin et al. (2016), except that participants were only given the option to pick one colour per word instead of two (see Figure 1, panel A, for a screenshot of the test). The test was not timed, and participants could exit the test and start again later from where they had left off if they wished. After providing a colour response for each of the 92 words, participants were debriefed and paid £12 for their participation. Testing took place approximately eight months after the first test reported in Mankin et al. (2016 [Chapter 2]), and no mention was made to the fact that the current wordlist was composed of the constituents of the wordlist the participants had seen in this earlier test. It is therefore highly unlikely that the participants realized that they were reporting colours for the constituents of words they had seen before.

We also obtained colours for individual letters for each participant through our Sussex Synaesthesia Database. All our participants had given their colours for letters through the online synaesthesia diagnostic test mentioned above (www.synesthete.org) prior to their recruitment for Mankin et al.'s (2016) original study. This test of genuineness is highly similar to the word-colour test described above, except that participants are shown each grapheme (for English speakers, the 26 letters A – Z and the numerals 0 – 9) three times each intermixed in a random order (see Figure 1, panel B). During this test, each synaesthete's letter colours were recorded in RGB colourspace on each of the three trials per grapheme, and the distance between the three colours for each grapheme is compiled into an overall consistency score, where lower values indicate higher consistency (i.e. a smaller distance between colours for the same grapheme; see Eagleman et al., 2007, for calculation). To obtain a single point in colourspace for each letter in each synaesthete's alphabet, we averaged the three values for R, G, and B for each letter colour. That is, since each letter had three different colour responses across the three trials in the test, for the R value we took the average of the three R values for that letter, G was the average of the three G values, and B the average of the three B values. This produced a single centroid RGB colour for each letter, which represented the average colour of that letter across the three trials. If a synaesthete had completed the grapheme-colour test more than once and therefore had more than one set of colours, we used only the first set of colours.



A B

Figure 1. Panel A: Screenshot of the online word-colour test in the present study. For each word, participants had to select "yes" or "no". They could then select a colour for the word by clicking on the drop-down colour palette (shown here expanded) and clicking "choose" when they were satisfied with the colour. Panel B: screenshot of the grapheme-colour consistency test at synesthete.org, used to obtain colour associations for graphemes and verify genuine grapheme-colour synaesthetes (Eagleman et al., 2007).

Results and discussion

Colourspace transformations

For this and the following analyses, the colours for both letters and words provided by each participant were transformed into CIELuv colour values. This colourspace was selected because CIELuv, unlike HSL or RGB, is a perceptually representative colourspace, and can discriminate most accurately between synaesthetes and non-synaesthetes (Rothen, Seth, Witzel, & Ward, 2013). CIELuv has three dimensions: L*, measuring lightness, from 0 (black) to 100 (white); u*, an axis from blue to yellow; and v*, an axis from red to green. These values describe a three-dimensional colourspace that allows calculation of Euclidean distances between any two colours. Due to the pragmatics of testing, each participant completed the test on their own monitor outside of a controlled environment. This means that the reported distances between colours, ΔE, are estimates based on conversions via RGB using a default XYZ setting of 94.81/100.00/107.30 (D65 daylight).

Which letter dominates the synaesthetic colour of the whole word?

To begin, we wanted to know which letter in a word most strongly contributed its colour to the whole word. We found this *source letter* by identifying which of the letter-colours in a word was closest to the whole-word colour. To do this, we measured the Euclidean CIELUV distance between the colour of the whole word (e.g. *rain*) and the colours of its first four letters to maximally capture the first and second consonants and vowels across all words. Four words (*ear, egg, ice*, and *oil*) had a vowel as the first letter. This small number meant we could not make an effective contrast between the consonant- and vowel-initial words, so we removed the vowel-initial words from our analyses, along with a single item with three onset consonants (i.e. *screw*). This left 87 items in our wordlist, all of which had at least one consonant in initial position (e.g. *rain, chair*) and at least one vowel (e.g. *rain, chair*) in the first four letters. We then calculated the distances between the colours of each of the four letters to the colour of the whole word (e.g. the distance between the colour of *R* to *rain*, then *A* to *rain*, and so on). We defined the smallest of these four letter-word distances as the *source* of the whole word's colour, as it was most similar to the colour of the whole word. We also coded whether the source letter was a consonant or vowel, and its order within the word (e.g. *rain* has first consonant *R* and second consonant *N*, first vowel *A* and second vowel *I*).

For each synaesthete, we then counted how many times the first consonant (e.g. *R* in *rain*, *C* in *chair*) was the source letter (i.e. closer to the colour of the whole word than any of the other letters), the number of times the first vowel was the source (e.g. *A* in both *rain* and *chair*), the number of times the second consonant was the source (e.g. *N* in *rain*, *H* in *chair*), and so on. We compared this to the distribution we would expect by chance by dividing the total number of words (N = 87) by the number of possible sources (three consonants and two vowels = 5)[5]. That is, if word colour was not related to any of the letter colours in particular, each letter should be the closest by chance about 87/5 = 17.4% of the time. Using a binomial analysis, we identified which letter-colour sources were significantly associated with word colours above chance. Eight of our synaesthetes had the first consonant as the source letter for

---

[5] Although we only analysed the first four letters of each word, there are five possible sources because some words had three consonants within those four letters. For example, the word *brush* contains only one vowel (*U*) but three consonants (*B*, *R*, and *S*) in its first four letters. Therefore, we had to include third consonant (as in *S* in *brush*) as a possible source.

more than 50% of their words, and for all of them this association was significant [all *p*s < .001]. Since the first consonant was always the first letter, we classified these synaesthetes as *first-letter dominant*. Another two synaesthetes had over 50% of their words with the first vowel as the source, with this association also significant [both *p*s < .001], so they were classified as *first-vowel dominant*. The final two synaesthetes had no single letter used as the source of whole-word colour 50% or more of the time, but appeared to draw on both first consonant and first vowel. The binomial analyses showed that both of these sources were significantly associated with whole-word colour [all *p*s < .002], so we termed these synaesthetes *mixed-dominant*. This distribution is given in Table 1, below.

Table 1. The source of whole-word colour for each synaesthete. Column 2 shows the percentage of words closest in colour to the first consonant (here, "Cons"; total N = 87 words). Columns 3-6 show the same for the second and third consonants and first and second vowels respectively. The highest percentages for each synaesthete, representing the most frequent source of whole-word colour, are shaded dark grey. This most-frequent association was significant for all synaesthetes in a binomial analysis. Additional significant associations are shaded light grey (all corrected *p*s < .01). Based on these results, synaesthetes are grouped as first-letter-dominant (L1); first-vowel-dominant (V1); or mixed (X).

Synaesthete	1st Cons	2nd Cons	3rd Cons	1st Vowel	2nd Vowel	Group
CH	72.41	10.34	4.60	10.34	2.30	L1
FT	79.31	11.49	4.60	2.30	2.30	L1
GC	81.61	10.34	3.45	3.45	1.15	L1
JS	72.41	14.94	8.05	4.60	0.00	L1
LW	83.91	4.60	1.15	8.05	2.30	L1
ME	68.97	14.94	5.75	9.20	1.15	L1
MG	68.97	6.90	6.90	9.20	8.05	L1
SM	75.86	12.64	4.60	5.75	1.15	L1
HO	8.05	10.34	0.00	68.97	12.64	V1
KL	21.84*	4.60	9.20	59.77	4.60	V1
FI	31.03	12.64	10.34	39.08	6.90	X
SS	43.68	13.79	5.75	29.89	6.90	X

*This first-consonant association for KL was marginally significant (p =.059); however, KL chose first-vowel over first-consonant colour nearly three times as often, so we decided on a first-vowel classification.

We also confirmed that for each synaesthete, their most frequently associated source letter was indeed significantly closer in colourspace to the whole-word colour. By simply counting the number of times a particular letter was closest to the whole-word colour, it was possible that the closest letter-word distance (e.g. *R* to *rain*) was not actually significantly closer in colourspace than other letter-word distances. For example, if *R* and *A* were approximately the same distance from *rain*, but *R* was slightly closer, our analysis above identified *R* as the source letter. We wanted to establish that *R* was indeed significantly closer in colourspace to *rain* than *A* was. Since first letter and first vowel were the two most likely sources of whole-word colour among our synaesthetes, we tested whether the first letter and first vowel were indeed significantly different from each other for each synaesthete. To do this, we compared the colourspace distances between the whole-word colour and its first letter colour, and between the same word and its first vowel colour, for each of our 12 participants. For each of the ten first-letter-dominant and first-vowel-dominant participants, the difference in mean distance to the whole-word colour between these two sources (first letter vs first vowel) was significant in paired-samples t-tests [lowest significant $t = 3.37$, highest significant $p = .001$, Bonferroni-corrected α for a total of 12 comparisons = .004]. For the mixed synaesthetes, however, these distances did not differ significantly [higher $t = 0.60$, lower $p = .548$]. This confirmed that most synaesthetes do indeed exhibit a quantitative, systematic correspondence between their colours for certain letters and for whole words, and that this letter-to-word colour relationship differs between individual synaesthetes (for some the first letter and for others the first vowel). Figure 2, below, illustrates the comparison between first-letter and first-vowel colour distances using data from participants GC and HO respectively.

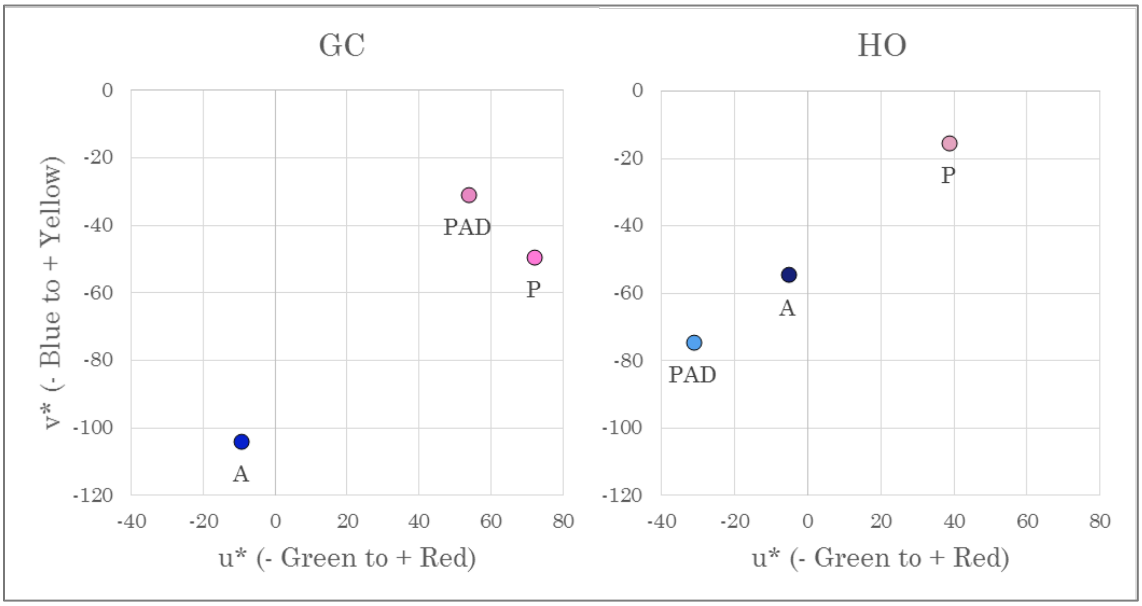


Figure 2. Comparison of two whole-word colour sources in CIELuv colourspace. Each point is a colour labelled with its inducer (i.e. the letter or word that triggered it) and is coloured to match the synaesthetic colour provided by the participants. For clarity, only the hue dimensions u* and v* have been included in the figure, collapsing L (lightness).

Both GC and HO happen to have pink for *P* and blue for *A* in their letter colours; however, they differ in which letter provides the source of the word colour *pad*. For synaesthete GC (left panel), the first letter is consistently closer in colourspace to the word colour than the first vowel, indicating a first-letter source of word colour; here; *P* and *pad* are closer than *A* and *pad*. Synaesthete HO (right panel) displays the first-vowel opposite pattern, with the vowel closer to the word colour than the first letter.

Finally, we return to the question of letter position in a word. Some words had the first vowel in second position (e.g. *rain*) and some in third position (for example *chair*). We were interested in whether the relative position of the first vowel would influence the dominant source of whole-word colouring. To test this, we asked whether the first vowel was more likely to be the source of whole-word colour when it was in second versus third position. Across all participants, we counted how many times the first vowel was in second or third position (e.g. *lamp, sled*), and then how many times within each position that the second vowel was the source of whole-word colour. A chi-square test indicated that position as either the second or third letter in the word did not influence the likelihood of the first vowel being the source letter [$\chi^2(1) = 2.15$, $p = 0.142$]. We tested the same relationship between position and source letter status for the second consonant, which can also vary in position (e.g. second e.g. *chair*, third e.g. *bag*, or fourth e.g. *rain*). We found that likewise, position in the word also did not significantly influence whether the second consonant was likely to be the source letter, [$\chi^2(2) = 4.67$, $p = .096$]. Although nonsignificant, this marginal result indicated that there may be a tendency for second consonants to be more influential further downstream in the word (second consonants that were the source letter when in second position = 8%; in third position = 15%; in fourth position = 12%). Given the tendency for synaesthetic colour to derive from the first letters in a word, this seems unexpected. However, since all the words that we tested here had a consonant in first position, when the second consonant was in second position, it must have been the second of a consonant cluster (e.g. *chair*). In this case, the second consonant may have been overshadowed by the first. However, our wordlist contained no items beginning with a three-consonant cluster (as we had removed *screw*), so in third and fourth position, the second consonant would always have been preceded by a vowel or vowel cluster (e.g. *bag*, *rain*). Therefore, this trend suggests that the second consonant may be more influential in whole-word colour when it is the first in a cluster of consonants, which may increase its salience; but further testing, especially with systematic manipulation of vowel and consonant positions in the words, would be necessary to establish this.

Thus far, we have shown that overall, synaesthetes do systematically colour their words according to particular letters within that word. The first letter in the word (in our dataset, always a consonant) and the first vowel in the word predominated as sources of whole-word colour for all of our synaesthetes. Previous studies have also reported this distinction; among a sample of 20 synaesthetes, Simner, Glover, et al. (2006) found that 15 synaesthetes (75%) used the first letter, three (15%) the initial vowel, and two (10%) used a mixed strategy. Among a sample of seven synaesthetes, Ward et al. (2005) found five (71%) first-letter, two (29%) first-vowel, and one (14%) mixed synaesthete. Our findings point to a similar proportion (8/12, or 67%) following the first-letter pattern, with two (17%) using a first-vowel pattern and the final two using a mix of both. These proportions of letter-to-word colouring strategies are remarkably consistent, despite relatively small samples and varying methodology. Our analysis of source letter position suggested that position downstream in the word (i.e. aside from first position) may marginally influence a letter's likelihood of being the source of whole-word colour. However, while consonant/vowel status appears to be a strong determinant, especially for the first instances of those letters in the word, the influence of downstream letters is less clear and requires further testing.

Why are some word colours different from their source letters?

Having established that synaesthetes tend to have words coloured by certain letters, we next investigate what may cause some words to *mismatch* the colours of their letters. That is, we expect that a word is typically coloured after a particular letter (the first letter for most synaesthetes, the first vowel for others, etc.), but this is not always the case. To show this, we returned to the distance in colourspace between the source letter identified above to the whole-word colour, which we call *source distance*. That is, for the word *ball*, if a synaesthete's letter colour for *B* was closest to the colour they gave for *ball*, we took the distance between their colours for *B* and *ball* in colourspace as the source distance. However, if a different synaesthete had *L* as the closest letter, we used the distance between *ball* and *L* for that synaesthete instead. This (source letter → word) colour distance outcome measure most closely matched a gamma distribution, with most of the observations close to zero and fewer observations with increasingly greater distances. We were interested in what causes these higher distances – that is, under what circumstances is the word not very similar in colour even to its closest letter colour? We hypothesized that these idiosyncratic word colours might stem from changes in word frequency, word imageability, and source letter frequency.

To investigate this, we constructed a generalised linear mixed effects model (GLME) to predict source distance. We conducted this analysis and the one below in Study 2 in R (R Core Team, 2016) using *lme4* (Bates, Maechler, Bolker, & Walker, 2015) and *languageR* (Baayen, 2013) with *p*-values for effects obtained with *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2016), fitting of fixed and random effects performed using *LMERConvenienceFunctions* (Tremblay & Ransijin, 2015) and figures graphed with *visreg* (Breheny & Burchett, 2016). Mixed models allow us to include random effects for participants and items, which is well-suited to modelling the differences between synaesthetes and looking beyond individual random variation to identify overall systematic effects. Indeed, the use of multilevel models like GLME has been shown to be particularly effective for analysing data from synaesthetes specifically (Hamada, Yamamoto, & Saiki, 2017).

We first created a model containing all our predictors of interest and their interactions: word frequency, word imageability, and letter frequency, all grand mean centred. We then forward-fitted random effects for participants and items, testing whether the addition of a random intercept, and further random slopes for each predictor, improved the model. We tested whether the inclusion of a random effect was justified in the model using log-likelihood tests. This procedure compares two models that are identical except for the inclusion of the effect of interest, and only retains the effect if it significantly improves the model fit (Baayen, Davidson, & Bates, 2008). Finally, we back-fitted the predictors in the model, removing any that did not significantly improve the model, beginning with higher-order interactions down. We also removed any random effects that did not significantly improve the model, along with random slopes whose predictors had been removed as fixed effects; this resulted in removing the nonsignificiant random effect of items, leaving only the random effect of participants in the model. The resulting model is summarised in Table 2, below.

Table 2. Summary of the generalised linear mixed effects model (GLME) predicting the distance in colourspace between each word colour and the colour of its source letter, termed "Source Distance." Source letter is defined as the closest of the first four letters in the word to the whole-word colour in CIELuv colourspace, here shortened to "Letter". Estimates and SEs are reported for a gamma distribution with a log link. Significant *p*-values are marked in **bold**.

	Source Distance			
Fixed Effects	Estimate (B)	SE (B)	*t*	*p*(t)
(Intercept)	3.23	0.08	39.40	**<.001**
Word Frequency (Zipf)	-0.08	0.04	-2.05	**.041**
Letter Frequency	-0.02	0.02	-1.14	.253
Letter Type (Consonant or Vowel)	-0.10	0.08	-1.17	.244
Letter Frequency x Letter Type	0.08	0.03	2.90	**.004**
Random Effects	Variance	SD	χ2	*p*(χ 2)
Participant	0.08	0.29	51.14	**<.001**
Residual	0.09	0.31	--	--

The negative main effect of word frequency indicates that more frequent words, such as *man*, tend to be coloured more consistently with their source letter than less frequent words, such as *bulb*. This suggests that frequent exposure to a word solidifies their synaesthetic colour close to the source letter colour, whereas rarer words may rather be subject to more variation, perhaps from the influence of other colours in the word.

We analysed the interaction of letter frequency and letter type by conducting separate GLME analyses for two types of source letters, consonants versus vowels. The models were identical to the main model but excluded the main effect and interaction term for letter type. This analysis showed that when the source letter was a consonant, the frequency of the letter had no significant effect on source distance [b = -0.02, SE(b) = 0.01, t = -1.26, p = .209]. However, when the source letter was a vowel, increasing frequency significantly increased source distance [b = 0.06, SE(b) = 0.02, t = 2.55, p = .011]. This means that when the closest letter-colour is a high-frequency vowel (e.g. *E*), the word will be less consistently coloured like that vowel. In other words, consonants do not influence how closely the word colour matches its source letter colour. However, vowels do, and the lower-frequency vowels (e.g. *U*) are closer to their word colours than higher-frequency vowels.

However, the question still remains *why* synaesthetes would choose a whole-word colour that does not closely match the colour of its most likely source letter. If we assume, as synaesthesia researchers often have, that word colour is a straightforward derivative of the colour of a letter in the word, then why would word-colour vary substantially from the colour of its closest source letter? We suggest two possible explanations, based on an examination of the colours given to us by synaesthetes (depicted below in Table 3). The first is that, at least for some synaesthetes, the word colour is an amalgamation of multiple letter colours (see e.g.

Kubitza, 2006; Mankin et al., 2016; Mills et al., 2002). In our study, synaesthetes could only report a single colour per word, but this may not fully capture their colour experiences for the word. Instead, the synaesthetes may have reported a *combination colour* for the word that combined several letter colours into one (see Table 3, below, for examples). In this case, "source letter" is misleading, as it is the letter that *happens* to most closely match the word colour in colourspace, but may not in fact meaningfully be the source of that colour. For example, in Table 3, first-letter synaesthete FT has a pale pink colour for *B*, which would imply that *bird* and *bag* should also be pale pink. However, while *bird* does indeed closely match the colour of *B*, *bag* is instead a dark pink, closer in colourspace to the dark maroon of *G* than to the pale pink of *B*. Accordingly, *G* has been coded as the source letter for the word *bag*. However, this does not adequately capture the synaesthete's colour for the word *bag*, since it is clear from looking at the colour that it is a dark shade of pink, i.e. the same hue as *B*. It appears that synaesthete FT has captured both the pink of *B* and the maroon of *G* with a dark pink. It is not clear from this data whether FT genuinely experiences this single colour for the word *bag*, or whether this dark pink is a combination colour attempting to capture multiple letter-colour influences.

Another explanation for these idiosyncratic colours that do not closely match their constituent letters is an influence of the real-word colour associations with the word, i.e. its *canonical colour* (for example, green for *pea*). We identified several cases where synaesthetes appeared to be reporting synaesthetic colour associations based on canonical rather than letter-based colour. For example, first-letter synaesthete CH has given green for both *pea* and *plant*, whereas the word colour for *place* shows that these words would typically be yellow like *P*. It may be that when the canonical colour is strong enough (e.g. *pea* strongly associated with green), this canonical colour can influence or even supplant the colour that a synaesthete would typically experience based on the word's letters. However, our wordlist was not balanced to systematically test these contrasts between canonical and letter-based colour, so we can only provide descriptive examples, pending a more controlled experimental investigation (see Mankin & Simner, in prep [Chapter 4]).

Table 3. Examples of possible combination and canonical colours from synaesthetes. "Group" designates the usual source of whole-word colour for each synaesthete, either L1 (first letter) or V1 (first vowel; see Table 1). "Source" identifies the letter in the word that was closest in colourspace to the word colour. The colour examples labelled "Combination" show colours from three synaesthetes: FT, LW, and MG. The top row for each synaesthete shows a "typical" word colour that most closely matched

the first letter in the word. The following row(s) show colours that clearly match the same colour category as the expected source letter (e.g. FT's word colour for *bag* is pink like *B*). However, another letter in the word was identified as the source letter because it was closer in colourspace to the whole-word colour (e.g. *G* for dark pink *bag*, even though *G* is maroon). These "combination" colours therefore appear to be a mix of multiple source letter colours. The colour examples labelled "Canonical" show possible instances of canonical rather than letter-based synaesthetic colours (e.g. green instead of yellow for *pea*).

Type	Synaesthete	Group	Word	Source	Word Colour	Source Colour
Combination	FT	L1	bird	B		
			bag	G		
			bar	R		
			bow	W		
	LW	L1	box	B		
			bow	W		
			bob	O		
	MG	L1	tea	T		
			tooth	O		
Canonical	CH	L1	gull	G		
			glass	S		
			place	P		
			pea	P		
			plant	P		
	MG	L1	hose	O		
			horse	O		
	SS	V1	muffs	M		
			mouse	M		
			post	P		
			pea	P		

Conclusions

This study investigated how the colours that synaesthetes experience for letters propagate to the colours they experience for whole words. We collected a precise measure of these word and letter colours using online colour tests and transformed them into the perceptually accurate CIELuv colourspace in order to compare word colours to letter colours (these latter provided earlier by the same synaesthetes). We then calculated Euclidean distances between the colour of each word and the colours of its first four letters for each participant (e.g. the distance between the colour for *rain* and for *R, A, I*, and *N*). This allowed us to establish which letters were closest in colourspace to the whole-word colour, which we defined as that word's *source letter*. Eight synaesthetes primarily used the colour of the first letter to colour the entire word, while two others used the first vowel, and two others showed no strong preference for either strategy but rather employed both. We then investigated how similar the source letter colour and whole-word colour were using the distance in colourspace between those two colours, which we called *source distance*. Our generalised linear mixed effects model indicated that overall, higher-frequency words were coloured more consistently like their source letters. However, when the source letter was a higher-frequency vowel, there was more distance between the word and source letter colour. We further suggested that there may be other influences at work, including multiple colours per word and the canonical colour of the word. Having investigated the influence of individual letters on whole-word colour for simple words, we will now explore the sources of synaesthetic colour in morphologically complex words – in this case, compound words.

Study 2: Colour propagation from word to compound

In this second study, we next turn to the colours that synaesthetes experience for compound words. In Mankin et al.'s (2016 [Chapter 2]) study, synaesthetes reported their synaesthetic colours for compound words (e.g. *rainbow*). They could report up to two colours for each compound word: the "strongest, most dominant" colour in the word, which was accordingly termed the *dominant colour* of the compound; and any additional colour, which was termed the *secondary colour*. Mankin et al.'s study looked at the number of synaesthetic colours each compound was reported to have, and the physical characteristics of those colours (e.g. their luminance and saturation). However, this study did not ask synaesthetes to report the "first" colour in a compound, but rather simply the "dominant" colour. This distinction was intentional, as this wording deliberately avoided the implication that the first-reported "dominant" colour must map onto the first constituent morpheme of the compound (i.e. that the dominant colour of *rainbow* should necessarily come from *rain*). As the current study has now obtained the colours for the constituent morphemes for these words from the same synaesthetes, we can test whether the dominant and secondary colours do indeed map onto the first and second constituent morphemes of the compound.

This study reports two sets of analyses: first for the dominant colour, then for the secondary colour. For each, we compare the compound colour (i.e. dominant or secondary) to the colour of that compound's morphemes and to its letters. This allows us to determine whether the dominant and secondary colours of a compound correspond to the colours of its first and second constituent morphemes; for example, does the dominant colour of *rainbow* come from the colour of *rain*, and the secondary colour from *bow*, or from the colours of the letters within *rain* and *bow*? In summary, the following analysis will investigate the sources of the two colours of a compound word, and what we can learn about compound processing using those colours.

Dataset

As in previous analyses reported here, all data came from the same 12 participants as in Study 1. For this study, we utilized three sets of colours for each synaesthete: colours for letters, for constituent morphemes, and for compound words (e.g. colours for *R, A, I, N, B, O, W, rain, bow*, and *rainbow*). Letter colours were obtained from the Sussex Synaesthesia Database, as described in Study 1, using the data from the grapheme-colour consistency test at www.synesthete.org. The data for constituent morpheme colours were the same as Study 1, above. The dominant and secondary compound colours came from the study reported by Mankin et al. (2016 [Chapter 2]). Both the compound colours from Mankin et al. (e.g. for *rainbow*) and the constituent morpheme colours obtained by this study (e.g. for *rain* and *bow*) were obtained from the same synaesthete participants using an identical online test apparatus, including an identical colour palette. The only methodological difference between the compound and constituent morpheme tests was that the latter (e.g. what colour is *rain*?) required the selection of a single colour, while the former (e.g. what colour is *rainbow*?) allowed up to two colour choices, including reporting no colour at all. We transformed the compound word colours collected by Mankin et al. into CIELuv colourspace, as we had done for letters and for constituent words in Study 1, to allow us to compare them directly. As in Study 1, this transformation from HSL to CIELuv via RGB used a default XYZ setting of 94.81/100.00/107.30 (D65 daylight) to estimate the reported distances between colours, ΔE.

We had previously excluded the four vowel-initial constituent morphemes (i.e. *ear, egg, ice*, and *oil*), as well as *screw*, from our analyses in Study 1. We therefore removed the seven compounds that included these morphemes from the subsequent analyses reported here (i.e. *earmuffs, earring, eggplant, icecream, iceskate*, *oillamp,* and *corkscrew*). As in Study 1, we included three items that were, strictly speaking, composed of two morphemes themselves, for the reason described previously: *clothes* from *clothespin*; *cuffs* from *handcuffs*; and *phones* from *headphones*. (The fourth, *muffs* in *earmuffs*, had been removed due to the vowel-initial morpheme *ear*.) We will refer to these bimorphemic items, along with the rest of our monomorphemic wordlist, as the “constituent morphemes” of our compound words. This left us with a final list of 87 constituent morphemes and 52 compound words. A full list of all of these items and relevant psycholinguistic measures can be found in Appendix A.

Results and discussion

What is the source of the dominant colour of a compound word?

We begin by investigating the relationship between the dominant, first-reported colour in a compound word and several possible sources of word colours. We calculated the Euclidean distance in CIELUV colourspace from the dominant compound colour (e.g. the first colour that synaesthetes reported for the whole compound *rainbow*) to twelve other colours: first morpheme colour (e.g. the whole-word colour of *rain*); second morpheme colour (e.g. the whole-word colour of *bow*); and the letter-colours of the five possible sources in each morpheme (i.e. first, second, or third consonant, and first or second vowel) [6]. For example, we measured the distance between the dominant colour of *rainbow* and to the colours of *rain, bow, R, A, I, N, B, O,* and *W,* and identified which of these colours was the closest. As we had done in Study 1, we then counted how many times each possible source (e.g. first morpheme, second vowel, etc.) most closely matched the dominant compound colour for each synaesthete (see Table 1). This distribution indicated that the dominant colour of the compound word is most similar to the colour of the first morpheme for the majority of synaesthetes. Using our terminology from Study 1, the colour of the first constituent, for example *rain*, is usually the *source* of the dominant colour of the compound (here, *rainbow*).

As we had done in Study 1 for the letter sources of whole-word colour, we then evaluated the likelihood of obtaining this distribution by chance. However, we first calculated the baseline probability of each possible source matching the dominant colour by chance for each synaesthete. This had not been necessary in Study 1, since we had required synaesthetes to report one and only one colour for each of the morphemes. In the current study, we had to calculate this baseline probability for each synaesthete, as the compound colour test run by Mankin et al. (2016) had allowed synaesthetes not to report any colour for a particular compound (i.e. a “no colour” response), so for some synaesthetes the total number of coloured compounds

---

[6] Our preferred example, *rainbow*, illustrates only the first two consonants and vowels (e.g. two each in *rain*). However, the wordlist also included morphemes such as *brush*, which contains a third consonant within its first four letters.

was less than 52. Therefore, this baseline probability of a match by chance was calculated by dividing the total number of coloured compounds that each synaesthete had reported by the twelve possible sources of colour. These twelve sources comprised the whole-word colours of each morpheme (e.g. the colours of *rain* and *bow* for *rainbow*) and maximally the first three consonants and two vowels in each of these morphemes. For the eight synaesthetes who had reported a dominant colour for all 52 compounds, the baseline probability of a match between dominant colour and one of the twelve sources by chance was 52/12 = 4.33%. The binomial analysis for each synaesthete confirmed that the first-morpheme word colour was the source of the dominant compound colour (e.g. the whole-word colour of *rain* was closest to the dominant colour of *rainbow*) significantly more frequently than chance for all twelve synaesthetes [all *p*s $< .001$]. The first consonant (e.g. *R* in *rainbow*) was also a significant source of dominant colour [$p = .001$ or below] for all synaesthetes except HO and SS, for whom this association was not significant. Other significant sources of dominant colour were: the first vowel of the first morpheme (e.g. *A* in *rainbow*) for JS, FI, and SS [$p = .002$ or below]; the second consonant of the second morpheme (e.g. *W* in *rainbow*) for SM, ME, and GC [$p = .018$ or below]; the first vowel of the second morpheme (e.g. *O* in *rainbow*) for HO [$p = .030$]; and the whole-word colour of the second morpheme (e.g. *bow*) for KL [$p = .049$]. Table 4 summarises these results.

Table 4. Sources of dominant compound colour for each synaesthete. For each morpheme, possible sources are whole-word colour ("Word"), the first or second consonant within the morpheme ("C1" and "C2"), or first or second vowel ("V1" or "V2") [7]. Columns give the frequency with which each possible source most closely matched the dominant compound colour. This frequency is given as the percentage out of the total number of coloured compounds for each synaesthete (given in "Total Items"). For example, for synaesthete CH, the dominant colour of the whole compound came from the word-colour of the first morpheme for 23.08% of compound words. Dark grey shading identifies the most frequent source of dominant colour for each synaesthete, all of which were significantly associated at $p < .001$; other significant associations are shaded light grey. For each morpheme, the "Total" column sums up the frequencies for all sources within that morpheme (e.g. for CH, 69.23% of compounds had a source in the first morpheme). In the "Syn" column, first-letter synaesthetes are identified by regular type, first-vowel synaesthetes by *italics*, and mixed synaesthetes by **bold**.

[7] Third consonant colour was never the most frequent source nor significantly associated with dominant colour (highest percentage = 6.25) and has been omitted for legibility, but is included in each constituent morpheme's totals.

	First Morpheme						Second Morpheme						Total Items
Syn	Word	C1	C2	V1	V2	Total	Word	C1	C2	V1	V2	Total	
CH	23.08	38.46	0.00	7.69	0.00	69.23	7.69	5.77	7.69	3.85	1.92	30.77	52
FT	45.10	29.41	5.88	3.92	0.00	84.31	1.96	1.96	5.88	1.96	0.00	15.69	51
GC	46.15	25.00	5.77	0.00	0.00	76.92	5.77	1.92	11.54	1.92	0.00	23.08	52
JS	35.42	16.67	4.17	14.58	6.25	83.33	2.08	8.33	2.08	2.08	0.00	16.67	48
LW	48.08	32.69	1.92	3.85	1.92	88.46	3.85	1.92	3.85	1.92	0.00	11.54	52
ME	23.08	23.08	5.77	9.62	7.69	69.23	1.92	5.77	11.54	7.69	0.00	30.77	52
MG	48.08	23.08	1.92	5.77	1.92	80.77	5.77	0.00	5.77	1.92	1.92	19.23	52
SM	38.46	25.00	1.92	0.00	3.85	69.23	1.92	5.77	13.46	3.85	0.00	30.77	52
HO	66.67	6.25	2.08	4.17	2.08	81.25	4.17	0.00	0.00	10.42	4.17	18.75	48
KL	18.18	48.48	6.06	0.00	6.06	78.79	9.09	6.06	6.06	0.00	0.00	21.21	33
FI	26.92	15.38	5.77	26.92	1.92	80.77	3.85	1.92	5.77	3.85	1.92	19.23	52
SS	25.00	7.69	7.69	23.08	5.77	71.15	3.85	1.92	7.69	3.85	7.69	28.85	52

To summarise, these results indicate that the dominant synaesthetic colour of a compound typically derives from a source in its first constituent morpheme, and most frequently from that morpheme's whole-word colour. That is, the strongest, most dominant colour of *rainbow* tends to be the colour of *rain*. However, the use of distances in colourspace to identify these colour sources requires an important additional comparison. We will again illustrate this using our favoured examples: *rainbow, rain,* and *R*.

Our results in Study 1 showed that *rain* is typically coloured by *R*, and this relationship is particularly strong for high-frequency words like *rain*. Consider the possibility that *R* gives its colour to *rain*, and therefore both are very similar shades of purple, for instance. We then want to measure the distance between both *R* and *rain* to the dominant colour of *rainbow*. Because both *R* and *rain* are already similar in colour to each other, which of the two is closer to the dominant colour of *rainbow* may not be meaningful, but rather down to minute variations in colourspace. That is, because *rain* takes its colour from *R*, comparing both to the dominant colour of *rainbow* may be irrelevant. For example, the distance between *R* and *rainbow* may be 10, and between *rain* and *rainbow* 9. (We note here that in CIELuv colourspace, 1 is the minimum distance between two colours that can still be perceptually distinguished by a typical observer; Mokrzycki & Tatol, 2012). In this situation, *rain* would be coded as the source for *rainbow,* even though the difference between *rain* and *R* is negligible. And in fact, *R* might be the true source of the dominant colour of *rainbow* via the colour of *rain*. Therefore, in order for us to confirm that our distinction between first constituent morpheme and first letter sources in Table 4 is in fact meaningful, we must first establish that there is a genuine difference between *rain* and *R* as sources of dominant colour.

To clarify this, we looked at the subset of compound words that had a first-morpheme source of dominant colour (e.g. for which *rain* was the closest of the twelve possible sources to *rainbow*; N = 223). We then identified which of the letters in the compound was closest in colourspace to the dominant colour of the compound if we ignored the first morpheme as a possible source. For example, we calculated which of the letters in *rainbow* would be the source letter without considering the colour of *rain*. In our scenario above, this would identify *R* as the closest letter. If there was no real difference between *rain* or *R* as the source of *rainbow,* then the distances between *rain* → *rainbow* and *R* → *rainbow* should not differ significantly from each other. We tested this by conducting a paired-samples t-test between the two sources of dominant colour: the first morpheme (*rain*) versus the closest letter (*R*). The results showed that first-morpheme colour was significantly closer to the compound's dominant colour than the next closest letter colour [mean distance to closest letter colour = 32.18, SD = 21.67; mean distance to first-morpheme colour = 17.54, SD = 14.30; paired-samples $t(222) = 14.57$, $p < .001$, paired-samples $d = 1.02$]. This shows that the identification of *rain* as the source of the dominant colour of *rainbow* was not a quirk of colourspace distances. Rather, *rain* was consistently significantly closer to the dominant colour of *rainbow* than the closest letter was. We can therefore be confident that the dominant colour of a whole compound tends to come specifically from the colour of its first constituent morpheme, more so than the colour of the closest letter in the compound. Importantly, this means that there is a meaningful link between the strongest, most dominant colour that synaesthetes report for a compound word, and the morphological structure of that compound.

Our analysis thus far has established that the dominant colour of the whole compound tends to be meaningfully derived from the colour of its first constituent morpheme. As Table 4 above shows, however, for most synaesthetes, this was not *always* the case; other letters in the compound, and particularly its first letter or vowel, were also significantly associated with the dominant compound colour. Therefore, we asked why some of these dominant colours derived from their first constituent (e.g. *rainbow* coloured like *rain*) while other dominant colours derived from other letters in the word (e.g. *rainbow* coloured like *R* or *A*). As we suggested in the introduction, this distinction could be indicative of different processing routes depending on the frequency of the constituents or the compound. Therefore, we tested whether the frequency of the whole compound and the frequency of the first constituent, along with their interaction, influenced the probability of the compound having the first constituent word colour as the source of its dominant colour. We again used a generalised linear mixed effects model (GLME), using the same statistical methodology as described in our previous GLME analysis in Study 1. However, for the current analysis, our model had a binomial outcome: either a first-morpheme colour source or a letter colour source. That is, we asked whether the frequency of the compound *rainbow* and of the morpheme *rain* influenced the probability that *rainbow* would be coloured most like *rain* or like one of its first letters (i.e. *R, A, I,* or *N*)[8]. We restricted this analysis only to compounds that were coloured either by their first morpheme or by another letter source within the first morpheme, excluding compounds that had closest colour sources in the second morpheme. This was because we were specifically interested in differentiating between these two possible sources (i.e. first morpheme vs the first letters of a compound). As before, we tested the inclusion of random effects for participants and items using chi-squared model comparisons (Baayen et al., 2008), and found no significant random effect of items, so this was removed. The model is summarised in Table 5, below.

Table 5. Summary of the binomial GLME model predicting the probability of a compound having a first-constituent source for its dominant colour, compared to having its source in one of its letter colours. Both frequency measures were grand mean centred. Significant p-values are highlighted in **bold.**

	Word vs Letter Colour Source			
Fixed Effects	Estimate (B)	SE (B)	z	$p(z)$
(Intercept)	-0.10	0.21	-0.47	.678
Compound Frequency	0.23	0.16	1.40	.161
First Morpheme Frequency	-0.09	0.21	-0.44	.658
Compound x Morpheme Frequency	-1.07	0.43	-2.49	**.013**
Random Effects	Variance	SD	χ^2	$p(\chi^2)$
Participant	0.46	0.68	22.19	**<.001**

This analysis suggested that the interaction of the frequencies of both the first constituent morpheme and the whole compound had an influence on how likely it was that the dominant compound colour came from its first morpheme (e.g. *rainbow*

[8] Four of our compound words (i.e. *clothespin, firehose, lifevest,* and *nailfile*) had no frequency rating in the SUBTLEX-UK corpus and were therefore excluded from this analysis.

coloured like *rain*) rather than one of its letters (e.g. *rainbow* coloured like *R*). To break down this interaction, we conducted separate binomial GLME analyses on the same data split between higher- and lower-frequency first morphemes (mean Zipf frequency = 4.75). When the first constituent morpheme was higher frequency (e.g. *hand* in *handcuffs*), whole-compound frequency had no effect on the source of its dominant colour [b = -0.08, SE(b) = 0.17, z = -0.45, p = .650]. However, when the first morpheme was lower frequency (e.g. *dough* in *doughnut*), increasing compound frequency marginally increased the probability that the dominant compound colour would come from the morpheme rather than any of the letters [b = 0.58, SE(b) = 0.32, z = 1.81, p = .071]. This was unexpected given our prediction that higher compound frequency would increase the likelihood of a letter-based dominant colour source. That is, if higher frequency compounds are lexicalised, we expected that they would be processed more like whole words, and therefore derived their dominant colour from a letter rather than the first morpheme. Here we instead found that higher-frequency compounds tend to be coloured more like their first morpheme, but only when that morpheme is lower-frequency.

It may be that this relationship does indeed capture the letter effect we expected, but it is obscured by the fact that we collected only one colour per word for the constituent morphemes. In Study 1, we showed that lower-frequency morphemes were not as closely coloured by their letters as higher-frequency morphemes (e.g. *rain* coloured more closely to *R* than *dough* coloured like *D*). We suggested that the synaesthetic colours of lower-frequency words may represent the influence of multiple letter colours, which are combined into a single colour that does not closely match any particular letter. That is, this idiosyncratic word colour for lower-frequency morphemes might be synthesised from multiple letter colours (e.g. *bag* coloured dark pink to represent the pink of *B* and the maroon of *G*; see Table 3, above). However, these are the same letter influences we expect will influence the colour of a whole compound. That is, if *suitcase* is lexicalised as a whole, we expect that it will derive its dominant colour from the initial letter(s) – e.g. *S, U,* etc. However, as its first morpheme *suit* is relatively low in frequency, the synaesthetic colour of *suit* would also derive from its first few letters, which are also *S, U*, etc. That is, the idiosyncratic combination colour of the low-frequency first morpheme may in fact reflect the underlying letter colours, which propagate to the dominant colour of the compound. Why is it then that the whole compound *suitcase* does not have a colour based on one particularly dominant letter? We had expected that it would, if it was lexicalised as a whole, since Study 1 showed that higher-frequency words tended to derive their colour more closely from a particular letter. This may be due to the fact that compound words, both in our list and in English in general, tend to be low in frequency. This would mean that even "high frequency" compound words might still not be frequent enough to have a single colour derived from one particular letter. Overall, this analysis points to the frequency of the compound and the morpheme having an effect on the whole compound's colour, but the underlying processes require further clarification.

In summary, we have shown that the dominant, first-reported colour of a compound word (e.g. *rainbow*) typically derives its colour from the colour of its first constituent (e.g. *rain*). However, whether *rainbow* is coloured more like *rain* or one of its letters depends on the frequency of both words. In particular, compounds with a lower-frequency first morpheme but higher whole-compound frequency (e.g. *doughnut*) tended to be coloured more like their first morpheme (e.g. *dough*) than like the closest letter. On the whole, our data suggest that the representation of *rain* is activated when reading the word *rainbow*, and contributes its colour to the whole word. We next ask whether there is also an influence of *bow* in *rainbow* – that is, when a compound is given two colours, does the secondary colour analogously derive from the colour of the second constituent?

What is the source of the secondary colour of a compound word?

The previous analysis has clarified how the colours of simple words, such as *rain*, and letters propagate to dominantly colour a compound word in synaesthesia. However, an unusual feature of Mankin et al.'s (2016) study was that synaesthetes could choose up to two colours for each compound, and indeed the majority of compound words in that study were given two colours. We next ask what the source is for the *secondary colours* that synaesthetes chose. We expect that the secondary colour may be likely to come from the second constituent, but it could equally come from another letter in the first constituent. Therefore, as above for dominant colours (see Table 4), we compared the distance between the secondary colour and each of the first four letters in both of its constituents. The distribution of most frequent source letters is summarised in Table 6, below.

Table 6. Sources of secondary compound colour for each synaesthete. For each morpheme, possible sources are whole-word colour ("Word"), the first or second consonant within the morpheme ("C1" and "C2"), or first or second vowel ("V1" or

"V2") [9]. Columns give the frequency with which each possible source most closely matched the secondary compound colour. This frequency is given as the percentage out of the total number of coloured compounds for each synaesthete (given in "Total Items"). For example, for synaesthete CH, the secondary colour of the whole compound came from the word-colour of the second morpheme for 12.5% of compound words. Dark grey shading identifies the most frequent source of secondary colour for each synaesthete, all of which were significantly associated at $p < .001$; other significant associations are shaded light grey [all *p*s < .05]. For each morpheme, the "Total" column sums up the frequencies for all sources within that morpheme (e.g. for CH, 66.67% of compounds had a source in the second morpheme). In the "Syn" column, first-letter synaesthetes are identified by regular type, first-vowel synaesthetes by *italics*, and mixed synaesthetes by **bold**.

	First Morpheme						Second Morpheme						Total
Syn	Word	C1	C2	V1	V2	Total	Word	C1	C2	V1	V2	Total	Items
CH	8.33	4.17	4.17	8.33	8.33	33.33	12.50	33.33	4.17	16.67	0.00	66.67	24
FT	7.14	0.00	0.00	0.00	0.00	7.14	57.14	35.71	0.00	0.00	0.00	92.86	28
GC	1.92	3.85	3.85	0.00	0.00	9.62	48.08	32.69	0.00	3.85	1.92	90.38	52
JS	0.00	0.00	0.00	11.11	0.00	11.11	22.22	44.44	0.00	0.00	11.11	88.89	9
LW	1.96	0.00	19.61	21.57	21.57	72.55	5.88	5.88	3.92	3.92	7.84	27.45	51
ME	2.00	4.00	4.00	4.00	2.00	16.00	26.00	40.00	6.00	4.00	2.00	84.00	50
MG	6.98	0.00	2.33	30.23	11.63	53.49	6.98	2.33	2.33	18.60	16.28	46.51	43
SM	2.22	4.44	13.33	2.22	0.00	24.44	35.56	35.56	0.00	2.22	0.00	75.56	45
HO	2.17	2.17	2.17	4.35	4.35	15.22	63.04	2.17	0.00	13.04	6.52	84.78	46
KL	17.86	3.57	3.57	3.57	14.29	42.86	21.43	7.14	3.57	17.86	3.57	57.14	28
FI	10.81	0.00	2.70	2.70	0.00	16.22	24.32	13.51	8.11	29.73	2.70	83.78	37
SS	7.84	3.92	1.96	7.84	9.80	39.22	9.80	17.65	7.84	17.65	5.88	60.78	51

As this table shows, secondary colour most often comes from a source in the second morpheme for ten of our synaesthetes. However, we note two exceptions: first-letter synaesthetes, LW and MG, tended to use subsequent letters from the first constituent as the source of secondary colour. This was especially characteristic of LW, who had a colour source within the first morpheme for 73% of their secondary colours, whereas secondary colour was split for MG (53% from the first constituent and 47% from the second). This suggests that these two synaesthetes may have reported the colours that they experienced in sequence (e.g. dominant colour from the first letter, secondary colour from the second letter), ignoring any underlying morphological structure. Overall, however, synaesthetes appear to be sensitive to the morphology of compound words, such that their two colours tended to come from

[9] Third consonant colour was never the most frequent source nor significantly associated with dominant colour (highest percentage = 7.14) and has been omitted for legibility, but is included in each constituent's totals.

the two morphemes. As we had done for dominant colours, above, we again established that the second morpheme in particular was uniquely contributing its colour to the secondary colour of the compound. That is, as with *R* and *rain* in the previous analysis, we checked that there was a meaningful difference between *B* → *rainbow* and *bow* → *rainbow* as sources of secondary colour. To do this, we again isolated the compounds with a second-morpheme source of secondary colour (N = 130) and identified the closest letter source of secondary colour. We then compared these two sources (i.e. second morpheme versus closest letter) using a paired-samples t-test. As in Study 1, the results confirmed that the secondary colour of the compound came specifically from the second morpheme, rather than the colour of the closest letter [mean distance to second morpheme = 17.12, SD = 14.63; mean distance to closest letter = 31.05, SD = 20.94; paired-samples $t(129) = 11.92$, $p < .001$, paired-samples $d = 1.15$]. Together with our previous analysis, this indicates that the dominant and secondary synaesthetic colours of a compound do systematically map onto the first and second morpheme of that compound respectively[10].

Conclusions

This analysis has investigated the sources of the multiple colours that synaesthetes experience for morphologically complex words – here, specifically noun-noun compound words in English, such as *rainbow*. Mankin et al. (2016) collected two colours (the dominant and secondary colour) from synaesthetes for a list of compound words, but were unable to establish the sources of these two colours. Here we expanded this investigation by comparing the colours for each compound to the colours that synaesthetes reported for the compound's constituent morphemes (e.g. the colours of *rain* and *bow* separately for *rainbow*; see Study 1) and for the letters making up those morphemes. We found that the dominant, first-selected compound colour typically derives from a colour source in the first morpheme, and in particular the first morpheme's whole-word colour. We further confirmed that it was the first morpheme in particular, distinct from the colours of its letters, which contributed to the dominant colour. This suggests that the colour associated with *rain* is activated when reading *rainbow*, implying that the lexical representation of the word *rain* is similarly activated. Our linear mixed effects model also showed that the dominant colour was more likely to come from the first morpheme as opposed to the closest

[10] We do not report a second GLME for secondary colour here because we did not have clear hypotheses as we did for dominant colour.

letter when the compound was high frequency but the first morpheme was lower frequency (e.g. *suitcase*). This account fits with decompositional or multiple-route accounts of compound processing.

We further showed that the secondary compound colour tended to come from a source in the second morpheme, either the second morpheme colour itself (e.g. *bow* for *rainbow*) or the dominant letter in that morpheme (e.g. *B* for first-letter synaesthetes, *O* for first-vowel synaesthetes). However, two synaesthetes preferred a subsequent letter in the first morpheme (e.g. *A* in rainbow) over a second-morpheme source. Together, these findings imply that the colours that synaesthetes experience for compound words correspond to the underlying morphological structure of compounds.

General discussion

In this study, we have explored how grapheme-colour synaesthetes experience the colours for simplex and complex words, based on the colours of individual letters. Our first study collected the colours that a group of synaesthetes experienced for morphologically simple words (e.g. what colour is *rain*?) using an online colour palette. We then compared these whole-word colours to the colours that the same synaesthetes had given for the individual letters in those words. We measured the distance in colourspace between each word and its letters (e.g. the distance between the colour of *rain* and the colours of *R, A, I,* and *N*) to identify which of these letters was the closest, i.e. the most similar in colour, which we called the *source letter*. We showed that synaesthetes exhibit systematic preferences for which letter in the word is the source letter, and further confirmed this using a binomial analysis to identify significant source letter preferences. We found that the majority of our synaesthetes (8 out of 12, or 67%) had a significant tendency for the first letter in the word to be the source letter. In our dataset, all our words had a consonant in the first position (e.g. *rain, bow*), so we were unable to distinguish between first position and first consonants. Another two synaesthetes (16%) used the first vowel in the word (e.g. *rain, chair*), and the last two synaesthetes had significant preferences for both the first letter and the first vowel, indicating a mixed strategy. Previously, Ward et al. (2005) found that four of their seven participants had a strong tendency to colour the whole word by the first letter, regardless of vowel/consonant status, with another of their participants colouring words primarily by first consonant, indicating a proportion of 71% for this strategy overall. Similarly, Simner, Glover, and Mowat (2006) found in a sample of 20 synaesthetes with word colours derived from letter colours that 75% used the first letter (always a consonant in their data as well), 15% used the first vowel, and 10% reported a mix of both strategies. Therefore, our results using a more sensitive colour distance measure support these findings showing the existence of first-letter, first-vowel, and mixed synaesthetes.

As we suggested in the introduction, the first-letter pattern of synaesthetic colouring lines up well with studies on reading and visual word recognition (VWR). As we suggested in the introduction, the importance of the first letter colour for synaesthetic whole-word colouring reflects the importance of that letter for word recognition in typical processing (e.g. Aschenbrenner et al., 2017; Chanceaux & Grainger, 2012; Scaltritti & Balota, 2013). This supports our argument that grapheme-colour synaesthesia is mapped systematically onto underlying linguistic processes. That is, the same features that are salient in word recognition for everyone are accordingly salient in synaesthetic colouring. This means that synaesthesia can be used to investigate normal language processing, and may even pre-empt findings in non-synaesthetes. For instance, whole-word synaesthetic colouring based on the first vowel in the word has been noted by synaesthesia researchers for more than a decade (e.g. Simner, Glover, et al., 2006; Ward et al., 2005). However, as we mentioned in the introduction, many theories of VWR do not yet account for the different roles of consonants and vowels, although there is also evidence outside of synaesthesia to suggest that such a differentiation is warranted (cf Carreiras, Duñabeitia, & Molinaro, 2009; Duñabeitia & Carreiras, 2011). Ours and other studies showing that some synaesthetes rely primarily on the vowel for word colour indicate that this distinction may be central to word processing and recognition in everyone.

This vowel-based pattern of synaesthetic word colouring further suggests that the syllabic structure of words may influence their processing, as vowels typically form the syllabic nucleus of English words. Recently, Taft, Xu, and Li (2017) investigated syllabic structure influences on VWR by embedding real words inside nonwords (e.g. *furb*, *teap*). When the embedded real word ended in a consonant (e.g. *fur* in *furb*), participants were slower to reject it as a nonword than when the embedded word ended with a vowel (e.g. *tea* in *teap*). This indicates that syllabic structure is a meaningful component of VWR. This is exactly what Simner, Glover, and Mowat (2006) found in their case study of a grapheme-colour synaesthete, whose whole-word colour changed depended on the stressed syllable (e.g. **'*ca*-***non* coloured like *A* versus *ca-'**det*** coloured like *E*). Altogether, this evidence suggests that complex information about syllable structure, stress, and consonant/vowel letter status is available to readers during VWR, and the colours that a synaesthete experiences upon seeing a word are modulated by this information. We may then ask how the phonological rules of English in particular may influence this relationship between syllable structure and synaesthetic colouring. That is, the vowels in unstressed English syllables are frequently reduced to schwa /ə /, so the change in word colour based on stress is confounded with changes in the pronunciation of the typically dominant vowel, since schwa can be represented by any of the vowel letters in English (Roach, 2009). This could be investigated using vowel-based synaesthetes in languages that do not have this reduction to schwa in unstressed syllables (e.g. Spanish), or in English using words with more than two syllables.

After identifying which letter in a simplex word was the source of its whole-word synaesthetic colour (e.g. whether *R* or *A* was the source of the colour of *rain*), we also used a linear mixed effects model to investigate why some simplex words were closer in colourspace to their source letters than others. We asked what might predict when a word colour was *not* derived from a particular letter (i.e. higher distances between words and source letters). We found that higher-frequency simplex words (e.g. *rain*) tended to be coloured more similarly to their source letters, and lower-frequency words (e.g. *brief*) more idiosyncratically coloured. These idiosyncratic word colours may derive from competition between different source graphemes in the word. Synaesthetes frequently report experiencing the colours of multiple letters in word; that is, although the colour of a particular letter predominates, the other colours are still perceived (e.g. Baron-Cohen et al., 1993; Mankin et al., 2016; Mills et al., 2002). Our results suggest that the salience of these additional letter colours is modulated by the frequency of the word. On the one hand, as a word is seen more often (i.e. higher frequency), its colour derives more strongly from a particular letter in the word. Conversely, the source letter in lower-frequency words is not as predominant, and the colours of other letters in the word have an influence on the colour of the word overall. This pattern of colouring is strongly analogous to the process of *unitisation*, in which very common words are not processed as combinations of individual letters but as single units or "sight words" (e.g. Ehri, 2005; Ehri & Wilce, 1983, 1985). Since such words are not recognised by their letters but rather as a whole, we accordingly found a highly consistent synaesthetic colour "fossilized" from one particular letter, indicating that the word has that colour as a whole unit. In summary, our finding that synaesthetic colour responses depend on the frequency of the word indicates that these colours systematically and meaningfully reflect the underlying processes in visual word recognition, and can therefore be used to investigate those very processes.

Finally, we examined the actual colours provided by synaesthetes for simplex words and their source letters, and described possible reasons for idiosyncratic word colours that frequency could not account for. One of these was that word colours were not always straightforwardly derived from a single letter (e.g. *bird* is pink because *B* is pink), but rather appeared to be influenced by other letters (e.g. *bag* is dark pink because of a maroon *G*; see Table 3). Since we collected only one colour per word, we were not able to distinguish whether this colour is really what the synaesthete perceives (e.g. the word colour for *bag* is actually dark pink), or whether this is an attempt to represent multiple colour impressions at once (e.g. *bag* is both light pink and dark maroon → dark pink). This can be easily remedied by allowing synaesthetes to report multiple colours for words, a clear improvement in synaesthesia research methodology that is only beginning to gain traction. To date, only two published studies have even offered more than one colour as possibility (Blazej & Cohen-Goldberg, 2015; Mankin et al., 2016 [Chapter 2]). Until this option is offered to synaesthetes to better capture their experiences, distinguishing between a genuine single colour and a multi-colour combination must remain speculative.

In Mankin et al.'s (2016 [Chapter 2]) study on compound words, synaesthetes were asked to provide the "strongest, most dominant" colour in the compound, but no reference was made to order (i.e. they were never instructed to provide the *first* colour they experienced, only the most dominant). Study 2 followed up on Mankin et al.'s findings by comparing the compound's dominant colour to the colours of its letters and to the colours for its constituent morphemes. That is, we measured the distance in CIELuv colourspace between the dominant colour of *rainbow* and the colours of each of its constituent morphemes (i.e. *rain* and *bow*) and its letters (i.e. *R, A, B, O*, etc.) and again identified the closest distance as the source of the dominant colour. For all of our synaesthetes, the colour of the first morpheme was significantly associated with the dominant compound colour (e.g. the dominant colour of *rainbow* came from *rain*). We further checked whether the dominant colour was indeed derived from the morpheme colour as a word, and was not really a letter-colour influence. That is, if *rain* is coloured like *R*, and *rainbow* is coloured like *rain*, then the distance between *rainbow* and *rain* should be indistinguishable from the distance between *rainbow* and *R*. This would mean the preference for *rain* to contribute the dominant colour of *rainbow* would be down to minute variations in colourspace. However, this was not the case; we found that for the compounds with a first-morpheme source, the morpheme colour was significantly closer to the dominant colour of *rainbow* than the closest letter colour. This indicates that when reading *rainbow*, the whole-word representations of the constituent morphemes are individually activated, which is marked by the "strongest, most dominant" colour in the compound deriving specifically from the word colour of the first morpheme.

We also asked what influenced whether the dominant colour of a compound was closer to its first morpheme or its closest letter source. To investigate this, we tested whether the frequency of the first constituent morpheme or of the whole compound could influence this difference in source. We found that the dominant colour was more likely to come from the first constituent morpheme when that morpheme was low frequency but the compound was high frequency (e.g. *suitcase* coloured like *suit* rather than *S*). Why might this be the case? We suggested above that lower-frequency constituents, e.g. *suit*, may have "combination" colours composed of multiple letter-colour influences. However, higher-frequency morphemes, e.g. *hand*, tend to have a whole-word colour very similar to their dominant letter. Overall in English, compound words are low in frequency, and therefore may be more likely to have a "combination" colour than a single colour fossilised from a particular letter. If this is the case, then the combination colour of both *suit* and *suitcase* would be based on the same letters, e.g. *S, U,* etc., and would therefore match each other closely. This underscores the importance of obtaining more accurate and nuanced colour associations, including more than one colour per word, to allow these comparisons. We also note that while the frequency of the compound words was intentionally contrasted, the frequency of the constituent morphemes was not explicitly balanced; specifically, all of the constituent morphemes were high in frequency, so confirmation of this effect should be pursued with a specifically designed wordlist.

Having established overall that the dominant colour of a compound does indeed map onto its first morpheme, we next identified the source of the compound's secondary colour. We showed that when compounds were given a second colour by synaesthetes, that secondary colour typically matched the word colour of the second constituent morpheme (e.g. the second colour of *rainbow* usually was closest to the colour of *bow*), and this significantly closer than the next closest letter colour. This corroborates accounts of compound processing which claim that the representation of each constituent morpheme as an individual word is accessed during the processing of a compound word (Andrews et al., 2004; Kuperman et al., 2008; Taft, 1988). That is, if a compound is processed as a lexicalised whole without regard to its morphological structure, we would expect that its colours would derive from its initial letters rather than its morpheme colours (e.g. the colours of *rainbow* deriving from *R* and *A*, rather than from *rain* and *bow*). The preference for constituent colour suggests that the lexical entry for the constituent morphemes must be activated during compound processing, inducing the synaesthetic colour associated with those morphemes. Our results provide further evidence from synaesthesia that compounds are broken down into their constituent morphemes during processing, which fits with both multiple-route or decomposition models of compound processing (Kuperman et al., 2008, 2009; MacGregor & Shtyrov, 2013; Stockall & Marantz, 2006; Taft, 2004). We did not find clear evidence in our analyses for the lexicalisation of higher-frequency compounds, which would have pointed to a multiple-route process. However, we suggested that future studies investigating the relationship between synaesthetic colour and language should collect multiple colours per word to allow a more detailed and nuanced analysis.

In summary, our studies have investigated the synaesthetic colours for words at increasing stages of morphological complexity. It is clear from these results that all words are not coloured equally; both their linguistic and colour characteristics interact to produce the overall word colour. The patterns of letter-to-word colour propagation supported recent evidence that consonants and vowels are processed differently, a distinction that may be due to individual differences or systematic features of visual word recognition. We also described colour evidence that canonical colour may further influence synaesthetic colour experiences, and that future studies of word colouring must allow for multiple colours in a word. Our results underline the critical role that psycholinguistic factors play in forming the synaesthetic colours for words, and explored the potential of synaesthesia as a psycholinguistic phenomenon and a tool for investigating language processing.

Chapter 4

Semantics in grapheme-colour synaesthesia: Exploring the alien colour effect

Abstract

Grapheme-colour synaesthesia, a condition wherein people experience colours associated with letters or words, is a promising but largely unexplored tool to understand language and embodied cognition. Here, we investigate the conflict that synaesthetes experience between two evoked colours: the synaesthetic colour derived from the letters in a word (e.g. purple for "red" if *R* is synaesthetically purple) and the semantic colour of those words (e.g. the red of "red", or the orange of "fire"). In the first experiment, we examine the colours that synaesthetes experience for colour terms and matched control words (e.g. "red"/"reed"). We conducted a linear mixed effects analysis to investigate whether the canonical, semantic colour of a word could exert a pull on its synaesthetic colour (e.g. the synaesthetic colour of "red" pulled closer to the canonical colour compared to "reed"). We found that higher-frequency colour terms like "red" were pulled closer to their canonical colours than lower-frequency colour terms (e.g. "azure"), while higher saturation of the canonical colour pushed the colour term away. In the second experiment, we investigated these effects in non-colour-term target words that were easy to mentally image (e.g. "fire") in comparison to control low-imageability words (e.g. "fine"). We found that, again, higher-frequency target words were pulled more strongly toward their canonical colour. We also found that the lightness and saturation of the control word were systematically related to the colour of the target word. Our study provides the first systematic investigation of the conflict between semantic/canonical and synaesthetic colours (i.e. the *alien colour effect*) and also suggests that canonical colour is automatically evoked in the normal course of language processing (Barsalou, 1999, 2008; Mannaert, Dijkstra, & Zwaan, 2017).

Introduction

Grapheme-colour synaesthesia is a condition wherein letters and words evoke automatic and consistent experiences of colour for a small portion of the population (Ward & Mattingley, 2006). To date, studies of grapheme-colour synaesthesia have primarily focussed on the colours associated with individual letters (e.g. Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007; Ward, Simner, & Auyeung, 2005), but interest is now turning to the synaesthetic colours associated with whole words. The synaesthetic colour of a word is primarily based on the colours of the letters that compose that word, and in particular the initial letter(s) (e.g. the words "fire" and "fine" would both be coloured by *F*; Asano & Yokosawa, 2012; Barnett, Feeney, Gormley, & Newell, 2009; Mankin & Simner, in prep [Chapter 3]; Simner, Glover, et al., 2006; Ward et al., 2005). However, some studies have found that the meaning of the word as a whole influences its synaesthetic colour (Asano & Yokosawa, 2012; Goodhew & Kidd, 2017; Gray et al., 2002, 2006). This lines up with anecdotal reports that the synaesthetic colour of the whole word might match the canonical colour of the denoted concept (e.g. "banana" might be yellow despite a blue *B*; Rich, Bradshaw, & Mattingley, 2005). In order to disentangle these findings, here we investigated what synaesthetic and linguistic characteristics influence the synaesthetic colours for words that already have a strong semantic, or *canonical*, colour (e.g. "red", "fire"). We will first review the influence of imageability and canonical colour on language processing more broadly, and then explore how this may interact with synaesthetic colour associations.

Imageability is a semantic characteristic generally defined as the ease with which a word can be mentally pictured or imagined (Paivio, Yuille, & Madigan, 1968). For example, it is easier to generate a mental image of the word "fire" (high imageability) than the word "freedom" (low imageability). High-imageability words are recalled better (Paivio, Walsh, & Bons, 1994; Rothen et al., 2012), have faster lexical decision times (Morrison & Ellis, 2000) and naming times (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Schwanenflugel & Akin, 1994), and are produced faster (Prado & Ullman, 2009) than low-imageability words. A classic explanation of these effects is dual coding theory, which suggests that highly imageable, concrete words are processed using both symbolic, verbal representations and perceptually-based sensory representations. However, low-imageability, abstract words are only encoded in the verbal semantic system (Paivio, 1969, 1971, 1991)[11]. This has since been supported by studies reporting that highly imageable concrete nouns are processed in different parts of the brain than abstract words (Bedny & Thompson-Schill, 2006; Sabsevitz, Medler, Seidenberg, & Binder, 2005) and with a different pattern of ERP activation (Xiao, Zhao, Zhang, & Guo, 2012). It is therefore no surprise that highly imageable words are named faster and more accurately by children (Masterson, Druks, & Gallienne, 2008) and are acquired earlier (Masterson & Druks, 1998), above and beyond the well-known advantage for nouns to be acquired before verbs (McDonough, Song, Hirsh-Pasek, Golinkoff, & Lannon, 2011). In sum, this evidence indicates that high-imageability words are easier to process, produce, and learn than low-imageability words, and this advantage may stem from additional perceptual or sensory information available for high-imageability words.

The clear mental image associated with words like "fire" can also have an additional interesting quality, namely a canonical colour. By *canonical colour*, we mean a strongly preferred, prototypical colour of the word's referent, such yellow for "banana", green for "leaf", blue for "sky", and so on. For clarity in this paper, we will denote the specific shade of a canonical colour in capitals (e.g. the canonical colour of the word "red" is the colour RED). Despite the variability of real-world colours (e.g. leaves are canonically GREEN but turn red or yellow in the autumn, and in the authors' part of the world, the sky is grey far more often than it is BLUE), a theory of cognition grounded in sensorimotor experience would suggest that the colour associated with a concept may be evoked as part of its representation (Barsalou, 1999, 2008). Several studies utilising the Stroop paradigm (MacLeod, 1991; Stroop, 1935) have tested whether words with canonical colours produce interference in naming the ink colour of that word, as colour terms do. That is, Stroop interference occurs when a word like "blue" is printed in a mismatching colour ink (e.g. orange), and the participant must respond with the ink colour ("orange"). The question then becomes whether the canonical colour of a non-colour-term word (e.g. the BLUE of "lake") has a similar disruptive effect when "lake" is also printed in orange ink. Studies have demonstrated that words with canonical colours like "lake" do produce greater

---

[11] As Connell and Lynott (2012) point out, the terms *concrete* and *imageable* are often used interchangeably in the literature, although different rating scales exist for each (e.g. Altarriba, Bauer, & Benvenuto, 1999; Coltheart, 1981). Here we will focus on imageability, as we are specifically interested in the colour impressions evoked in the visual modality, but we include evidence from studies on concreteness as well.

interference in naming the ink colour than neutral words such as "seat", but less than colour terms themselves (e.g. "blue"). This suggests that the canonical colour is indeed evoked by the word, although not as strongly as by the colour term directly (Klein, 1964; Proctor, 1978; Risko et al., 2006; Schmidt & Cheesman, 2005). This use of perceptual systems in conceptual processing was further supported by an fMRI study by Simmons et al. (2007), which found overlapping patterns of activation in the left fusiform gyrus for both colour perception and colour property judgments (e.g. whether taxis are YELLOW). Furthermore, Richter and Zwaan (2009) asked participants to judge whether two patches of colour were the same, separated temporally by the presentation of a colour word or a non-colour control word. They found that participants slowed in their colour judgments when there was a mismatch between the category of the colour patches (e.g. two shades of blue) and the intervening colour term (e.g. "red"); however, participants were no faster when the colour term matched the colour patches than with an intervening control word. Overall, it seems, colour terms themselves do automatically activate perceptual colour systems, but does the canonical colour evoked by non-colour-terms like "fire" do the same when no colour is explicitly mentioned? If this is the case – that is, if the fiery orange of "fire" is evoked even when there is no mention of "orange" – we expect that this automatically simulated canonical colour will have an influence on the synaesthetic colour experience as well.

Studies in this area have indicated that the canonical colour is indeed evoked in normal reading, even with no explicit mention of colour. Two studies (Connell, 2007; Zwaan & Pecher, 2012) tested this by presenting sentences with a canonical object colour (e.g. "Sarah stopped in the woods to pick a leaf off a tree"). Critically, there was no explicit mention of the colour, only one implied by context (e.g. a fresh leaf is typically GREEN). After reading each sentence, participants were asked to indicate whether a pictured object (here, a leaf) was mentioned in the preceding sentence in match (a green leaf) vs mismatch (an orange leaf) conditions. The results indicated that participants were faster and more accurate when the implied colour of the object matched the picture (Zwaan & Pecher, 2012; but see Connell, 2007). Connell and Lynott (2009) also showed this match advantage by presenting sentences that implied a particular canonical colour (e.g. "Joe was excited to see a bear at the North Pole") and then a target word (e.g. "bear"). When the priming sentence implied the typical colour of the target (e.g. BROWN for "…a bear in the woods"), naming times were faster for the target printed in that typical colour (e.g. "bear" in brown ink) than in an atypical or unrelated colour (e.g. "bear" in white or yellow ink). However, when the priming sentence implied an atypical colour (e.g. WHITE for "…a bear at the North Pole"), both typical and atypical colours were named faster than the unrelated colour. This experiment showed not only that the specific implied colour is simulated (e.g. a white bear at the North Pole), but the typical colour of the target is also evoked alongside the context-implied colour (e.g. "bear" evokes BROWN even when the context implies WHITE). This indicates that the canonical colour associated with a word is automatically evoked during processing, so we can reasonably expect that these canonical colours may influence or interfere with the synaesthetic colours for the same words.

Most recently, Mannaert, Dijkstra, and Zwaan (2017) addressed some irregularities in Connell's (2007) findings and subsequent colour-match advantages by manipulating two aspects of the experiment. First, they used the same experimental paradigm as Connell (2007) and Zwaan and Pecher (2012) with more carefully chosen, real-world colour pictures rather than line drawings (e.g. a picture of a green traffic light in a match context ["The driving instructor told Bob to go at the traffic light"] or a mismatch context ["…to stop at the traffic light"]). Here, they again found that participants were faster to respond when the implied colour and the colour in the picture matched than when they mismatched. Next, they reduced the saturation of the pictures to the lowest level at which the hue could still be recognised and again tested how quickly participants responded in the same paradigm. With desaturated pictures, participants responded as quickly in the mismatch condition as they did in the match condition. That is, the delay in response caused by the conflict between implied and picture colours (e.g. "…stop at the traffic lights" followed by a green traffic light) disappeared when the colour was less salient. This evidence suggests not only that colour will be automatically evoked by the presentation of a high-imageability word with a canonical colour, but that the influence of this canonical colour may depend on its saturation or perceptual salience.

Grapheme-colour synaesthetes provide a particularly interesting opportunity to study the influence of canonical colour for two reasons. First, canonical colour may be a particularly salient influence for synaesthetes because they have been frequently reported to have elevated mental imagery abilities compared to non-synaesthetes (Barnett & Newell, 2008; Brang, Miller, McQuire, Ramachandran, & Coulson, 2013; Chun & Hupé, 2016; Janik-McErlean & Banissy, 2016; Mealor et al., 2015; Price, 2009; Spiller, Jonas, Simner, & Jansari, 2015; but see Havlik, Carmichael, & Simner, 2015). Therefore, synaesthetes may experience the mental images evoked by high-imageability words even more vividly. This in turn may make an imageability influence in synaesthetic colours even stronger. However, in one of the implied colour experiments described above (e.g. responding to "…pick a leaf off a tree" with either a green or orange leaf), Zwaan and Pecher (2012) found that there was no correlation between mental imagery ability and the speed with which participants responded. It may therefore be the case that although mental simulation (i.e. the perceptual information stored as part of a concept's mental representation, such as the GREEN of a newly picked leaf) is an integral and routine part of cognition, mental imagery may be an unrelated conscious process (Barsalou, 1999). If this is the case, the higher mental imagery ability of synaesthetes may not cause a difference in their synaesthetic colour experiences.

The second and more intriguing aspect of studying canonical colours in word processing using synaesthetic experiences derives from the possibility for inherent conflict between the canonical colour of the mental simulation and the synaesthetic colour of that word based on its spelling. Several studies report a canonically-based synaesthetic colour association for certain words that does not match the word's letter-based synaesthetic colour. For example, Rich, Bradshaw, and Mattingley (2005) mention that one synaesthete in their large sample experienced yellow for the word "banana" although its grapheme colours were blue and brown. Baron-Cohen, Wyke, and Binnie (1987) reported that for word-colour synaesthete EP, "elephant" was synaesthetically grey, matching the canonical colour of elephants. Asano and Yokosawa (2012) found that ideographic characters in Japanese tended to be synaesthetically coloured according to their canonical colour (e.g. the characters for "red", 赤, and "blood", 血, were given a red colour by all of their participants). However, these ideographic characters represent a concept without reference to spelling or pronunciation, so the conflict with synaesthetic colours for the associated Japanese syllabary may be minimal. Goodhew and Kidd (2017) found that synaesthetes tended to report similar colours to non-synaesthetes for a list of words that often matched their canonical colour (e.g. yellow for "sun", brown for "mud", etc.). However, to date no study on synaesthesia has directly compared a word with a canonical colour to an orthographically matched word without one (e.g. canonically orange "fire" vs non-canonically-coloured "fine"). Such a comparison would allow us to understand how synaesthetes resolve this colour conflict. The question then becomes: why are certain words coloured canonically for synaesthetes, and what qualities of those words make the canonical colour strong enough to overcome the letter-based synaesthetic colour? The resolution of this colour conflict is not only a matter of synaesthetic colour associations: the factors that pull a synaesthete's associated colour away from the letter-based colour and toward the semantic canonical colour may be the very influences that are most important in canonical colour and mental simulations in general.

We will examine this influence of imageability and canonical colour in two experiments. To begin, we will investigate the clearest possible influence of canonical colour: the synaesthetic colours evoked by colour terms such as "red", "purple", "beige", etc. When synaesthetes see colour terms, they can experience either the denoted colour of the colour term (e.g. RED for "red") or the synaesthetic colour based on the colours of the letters the synaesthete experiences (e.g. purple for "red" if the synaesthete has a purple *R*). This latter is known as the *alien colour effect,* or ACE (Gray et al., 2002, 2006). Gray et al. (2002) first investigated this effect with participants reporting varying degrees of ACE (e.g. "red" → synaesthetically purple). These participants were given a conventional Stroop task (e.g. "blue" coloured yellow, with the correct colour response being "yellow") with an additional "negative priming" condition. In these negative-priming trials, the correct answer was the name of the colour that had to be inhibited on the previous trial (e.g. after "blue" coloured yellow, the next trial might be "red" coloured blue, so "blue" is the correct response). This study found that colour naming times increased with the degree of ACE, such that synaesthetes who reported the highest degree of ACE also displayed the greatest interference in the Stroop task in both conditions. This could be because, rather than inhibiting only one response in the Stroop task (e.g. inhibiting the response "blue" and instead responding "yellow"), high-ACE synaesthetes must inhibit an additional response, namely the alien colour they experience for that word (e.g. pink for "blue" if *B* is blue). This creates additional interference in colour naming, which is reflected in their increased reaction times. In a subsequent experiment, Gray et al. (2006) confirmed the reality of the ACE using fMRI. In the critical trials, synaesthetes again completed a conventional Stroop task, naming the ink colours of either colour words or rows of Xs. In comparison to both non-synaesthete controls and non-ACE synaesthetes, synaesthetes who experienced the ACE evinced greater activation in the supplementary motor area and the right hippocampus. Gray et al. (2006) suggest that the increased activity in the supplementary motor area reflects the inhibition of the additional, synaesthetic colour name, while the hippocampal activation represents an increased demand for goal conflict resolution in the Stroop task for ACE synaesthetes. Altogether, these two studies provide both behavioural and neurological evidence for the reality of the alien colour effect. However, neither study obtained their synaesthete participants' associated colours for colour terms, and the prevalence of ACE was based on self-report alone.

We will address this by gathering synaesthetes' colours for a list of colour terms (e.g. "red") and a matched list of non-colour-term control words (e.g. "reed"). Control words are used as an example of the "typical" synaesthetic colouring of similarly spelled words, so we can then compare how the colours reported for colour terms differ from their matched counterparts. This allows us to ask, for instance, whether high-frequency colour terms such as "red" or "blue" might be more likely to be canonically RED and BLUE respectively than less common colour terms, such as "maroon" or "beige", are to be MAROON and BEIGE. We would expect this if, as suggested above, canonical colours are automatically and perceptually evoked by a colour term. In this case, more frequent colour terms may have a "stronger" canonical colour relative to the letter-based colour for synaesthetes if they are evoked more often. In the second experiment, we make the same comparison between non-colour-term words contrasted by imageability, such as "fire" versus "fine". We expect that similar influences as those for colour terms will operate on the synaesthetic colours for these imageable words.

To summarise, the current study will investigate the influence of semantics and imageability on synaesthetic colouring in two experiments. Overall, these experiments will test whether and why the canonical colour associated with a colour term (e.g. "red"; Experiment 1) or highly imageable word (e.g. "fire"; Experiment 2), can influence the colour that a synaesthete automatically and consistently experiences.

Experiment 1: The alien colour effect

We begin with an investigation into the so-called alien colour effect, or ACE, by comparing the synaesthetic colours for colour terms (e.g. "red", "tan") with a matched list of control words (e.g. "reed", "tack"). This experiment tested the conflict between the canonical and letter-based colour of a word and the psycholinguistic characteristics of these words that may influence synaesthetic colour choice.

Methods

Participants

Twenty-three synaesthete participants were recruited from the Sussex Synaesthete Database, all of whom reported experiencing colours for both letters and words and were native speakers of English. These participants were between 18 and 74 years of age (M = 37.13 years, SD = 16.66), 20 female and 19 right-handed. In order to be included in the database, all participants had already been verified as genuine synaesthetes by completing the grapheme-colour synaesthesia test at www.synesthete.org, a standardised test for quantitatively verifying synaesthetic experiences (see Eagleman, Kagan, Nelson, Sagaram, & Sarma, 2007, for details of this diagnostic test)[12]. We used the score cut-off of 1.43 recommended by Rothen, Seth, Witzel, & Ward (2013) to determine genuineness; all participants scored below this cut-off on the test (M = 0.75, SD = 0.25, range = 0.37 – 1.18). Participants were each paid £10 for their participation in Experiments 1 and 2, below.

Materials

We selected the 20 most frequent colour terms for which there were both age of acquisition (AoA) and imageability ratings available in the literature, as we wanted to match colour terms and control words as closely as possible on these measures. We excluded colour terms referring to real-world objects like "gold" to avoid inflated values (e.g. "gold" will have a higher frequency due to references to the metal, not the colour). As synaesthetic colours for words are commonly based on the first consonant or vowel (e.g. "mother" usually matches the colour of *M*, and less commonly *O*; Mankin & Simner, in prep [Chapter 3]; Ward et al., 2005), we then selected control words for each colour term with similar spelling. These control words were all monomorphemic nouns beginning with the same onset grapheme(s) (e.g. "red" was paired with "reed"). All words were orthographically matched on their first letter and vowel, and some words were matched on up to 5 graphemes (e.g., "violet"/"violence"). The mean number of matches in initial graphemes was 2.8 (SD = 0.95). We also matched words pairwise as closely as possible on imageability ratings, AOA, and length, and group-wise on word frequency (using the Zipf

[12] In brief, this test presents each English letter A – Z and 0 – 9 three times in random order. For each grapheme, the participant selects a colour on a 16.8-million-colour palette. The distance in colourspace between the three presentations of each grapheme is calculated, with a closer distance indicating a more similar colour. This results in a consistency score across all graphemes, with a smaller score indicating higher colour consistency.

frequency measure; van Heuven, Mandera, Keuleers, & Brysbaert, 2014) and number of syllables. Table 1, below, summarises these measures. A full list of target and control words, with relevant psycholinguistic measures, is given in Appendix B.

Table 1. Descriptive statistics for the psycholinguistic and orthographic measures for target colour terms ("Colour") and control words ("Control"). The *t*-statistic reported is from independent-samples *t*-tests for each measure between colour and control groups.

Measure	Word Type	M	SD	*t*	*p*(t)
Imageability[13]	Colour	550.8	53.6	0.70	.486
	Control	538	61.3		
AoA	Colour	339.4	121	-0.54	.596
	Control	356.6	72.9		
Length	Colour	5.2	1.1	0.55	.583
	Control	5	1.1		
Frequency (Zipf)	Colour	4.6	1	1.60	.122
	Control	3.7	0.6		
Number of syllables	Colour	1.5	0.6	1.10	.290
	Control	1.3	0.6		

Apparatus and procedure

Participants were invited via email to complete an online test. After completing basic demographic information, they began an online colour selection task (see Figure 1, below). Participants were asked to indicate the "strongest, most dominant" synaesthetic colour that they experienced for each word. The colour terms and control words were interleaved with the 40-item wordlist for Experiment 2, below, for a total of 80 items presented in unique random order for each participant. Each item appeared in large, bold letters, above the question, "Word has synaesthetic colour?" If they selected "Yes", participants could open an expandable colour palette with 16.8 million colours. They first choose the desired colour using a sliding hue scale, then clicked on the palette to change saturation and lightness and confirm their choice by clicking "choose." If they selected "No", they could proceed to the next word without choosing a colour, and the response was coded as no colour. Participants could choose a maximum of one colour per word. After the participants had seen all 80 words, they were also asked to briefly describe where their

[13] One of our control words, "aztec" (for "azure"), did not have an imageability rating in the literature. To obtain one, we asked 14 native English speakers, naïve to the hypotheses of this study, to give imageability ratings for "aztec" and 13 other distractor words. We used the same instructions, procedure, and method of calculation as the original imageability rating studies (Bird et al., 2001; Stadthagen-Gonzale & Davis, 2006) to produce this rating.

synaesthetic colours for words came from. Finally, they were offered an open response box to give any additional information and feedback.

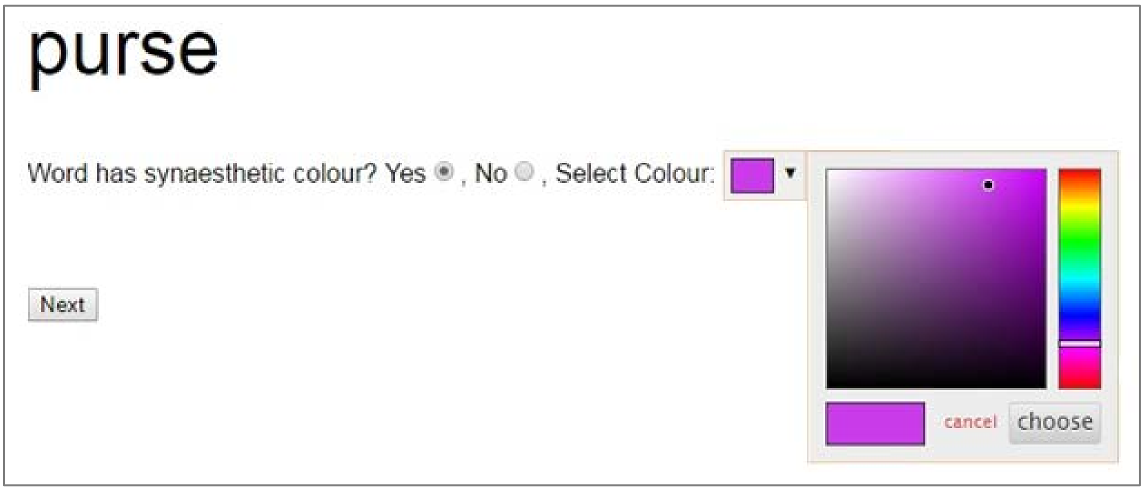


Figure 1. The online colour selection task. Each item appears in large letters in the middle of the screen (in this case, "purse"). For each item, participants indicated whether they experienced a synaesthetic colour (by selecting "Yes" or "No") and used the drop-down colour palette, shown expanded, to choose their synaesthetic colour.

Canonical colours from non-synaesthetes

We also obtained canonical colours for each of our colour terms by having ten colleagues, all non-synaesthetes and naïve to the hypotheses of the study, complete the same word-colour test described above. These were all native speakers of English, six female and eight right-handed, with a mean age of 32.1 years (SD = 8.70 years). They were instructed to "select the first colour you think of - whatever pops into your head first" and this should be the "first, most automatic" colour response to each word. Participants again used the same colour palette as described above to make their colour selections.

Results and discussion

Data preparation

Visual inspection of the data revealed that due to a technological error, three participants had recorded black for every item, so they were removed from the analysis, leaving 20 synaesthetes. To begin, we transformed all colours to the CIELuv colourspace. In this colourspace, the L dimension, lightness, runs from 0 (black) to 100 (white) for all colours; the other two dimensions, u* and v*, are red-green and blue-yellow axes respectively. This means that using CIELuv, we can calculate the Euclidean distance between two colours in a perceptually accurate, three-dimensional colourspace. Due to the pragmatics of testing, each participant completed the test on their own monitor outside of a controlled environment. Therefore the reported distances between colours, ΔE, are estimates based on conversions via RGB using a default XYZ setting of 94.81/100.00/107.30 (D65 daylight). For the purposes of later analysis, we also calculated the saturation of each colour using the following equation (Schanda, 2007):

$$\frac{\sqrt{(u^*)^2+(v^*)^2}}{L}$$

We also obtained values for the canonical colour of each colour term. Three of the non-synaesthetes did not complete the test as directed and chose black for most of the colour terms and were therefore removed, leaving legitimate colour choices from seven non-synaesthetes. We converted the colours given by our seven non-synaesthetes into CIELuv colourspace, as above, and then took the average value each of L, u*, and v* to find the centroid for each colour. This gave us an empirically calculated canonical colour for each colour term (e.g. a RED for "red"). These twenty colours are given in Table 2, below; further reference to canonical colours in this paper (e.g. RED, INDIGO) refers to these particular colours. We chose to obtain canonical colours in this way, rather than using standard locations in colourspace, because we wanted to use colours obtained with the same test apparatus that our synaesthetes had seen, to match their experience as closely as possible.

Table 2. For each of the 20 colour terms, the centroid colour obtained from seven non-synaesthetes to be used as a canonical colour. RBG values are provided for reference.

Colour Term	Canonical Colour	R	G	B	Colour Term	Canonical Colour	R	G	B
BLACK		0	0	0	PINK		244	124	204
GREY		153	149	149	SCARLET		232	54	44
WHITE		252	249	249	VIOLET		169	55	211
RED		206	10	14	CRIMSON		202	28	38
ORANGE		244	140	11	TAN		176	144	92
YELLOW		239	240	21	MAUVE		195	85	177
GREEN		40	198	44	BEIGE		229	206	162
BLUE		47	57	240	MAROON		130	16	44
PURPLE		152	36	195	INDIGO		134	45	224
BROWN		133	88	34	AZURE		38	189	224

Finally, we calculated the Euclidean distances in CIELuv colourspace for the synaesthetic colours provided by each synaesthete participant. This gave us three distances in colourspace for each item triple (e.g. "red"/"reed"/RED). These were (a) the distance between the synaesthetic colour for each colour term (e.g. "red") and the canonical colour for the same colour term (e.g. RED); (b) the distance from the synaesthetic colour for each control word (e.g. "reed") to the same canonical colour; and (c) the distance between the two synaesthetic colours for the colour term and control word (e.g. between "red" and "reed"). These distances will be used in our analyses below. A smaller colourspace distance means that the colours are more similar to each other; a distance of zero between two of the colours would mean that the exact same colour was chosen for those two words. Four of the synaesthetes – AG, BM, CT, and MT – had 10 or fewer complete pairs of colour distances (i.e. no colour was given for one or both of the colour term/control word pair, so comparison was impossible), and were excluded from further analysis in this experiment. This left 16 synaesthetes whose data are analysed below.

Data analysis: What synaesthetic colour do colour terms have?

We first wanted to know, for colour terms, whether synaesthetes more often experienced the canonical colour (e.g. "red" is also synaesthetically red) or a colour derived from the letters in the word. Because "red" and "reed" contain the same letters, we expected that they would elicit similar letter-based synaesthetic colours if there were no semantic influences at all. That is, we assume that the colour of "reed" captures the synaesthetic colour that each synaesthete would normally experience for "red" if "red" were not a colour term. By making this assumption, we treat control words like "reed" as the baseline colours for comparison. Therefore, systematic differences between the colours of "reed" and "red", for example, are attributed to the possible pull, or influence, of the canonical colour RED on the synaesthetic colour of "red".

To explore this, we returned to two of the colourspace distances we measured above: (a) target word to canonical colour (e.g. "red" → RED) and (b) control word to the same canonical colour (e.g. "reed" → RED). We first compared these two distances in a paired-samples t-test, and found that target word to canonical colour distance (e.g. "red" → RED) was overall smaller than control word to canonical colour distance (e.g. "reed" → RED) [Mdiff = -14.41; paired-samples $t(294) = -5.21$, $p < .001$, paired-samples $d = 0.35$]. This means that across synaesthetes, target words were more similar to their canonical colours than control words were (e.g. "red" was coloured more like canonical RED than "reed" was). We then looked at individual differences between synaesthetes. For example, for each synaesthete, we compared the distance between "red" and "reed" to the distance between "red" and the canonical colour RED. If "red" was closer to "reed" in colourspace than it was to canonical RED, then we coded this as *letter-coloured*; this "alien colour" (i.e. based on the letter colours rather than the canonical colour of the colour term) characterises the alien colour effect (Gray et al., 2002, 2006). If the reverse were true, and "red" was closer to RED, this was accordingly coded as a *canonical-coloured*. We made this comparison for every colour term for each synaesthete; the totals and percentages are summarised in Table 3, below. Importantly, all synaesthetes had at least some letter-based (i.e. alien) colours, and indeed some of them had alien colours for most or all of the colour terms. Therefore, this analysis showed that *all* synaesthetes experienced at least some degree of alien colour effect (ACE) for these colour terms.

Table 3. Counts and percentages of colour-term colours closer to the canonical colour ("red" is like RED → *canonical-coloured*) or the control word colour ("red" is like "reed" → *letter-coloured*), with totals for each synaesthete. Note that most synaesthetes had at least one missing colour within the colour term/control word pair, so many totals are less than the original number of matched pairs (N = 20). As the ACE is defined as experiencing a colour for a colour term based on the synaesthetic rather than canonical colour, the ACE percentage is the number of letter-coloured colour terms out of the total number of complete pairs for each synaesthete.

Synaesthete	Canonical-coloured	Letter-coloured	Total	Percentage ACE
CSP	0	11	11	100.00
AM	1	19	20	95.00
TR	1	14	15	93.33
CS	2	18	20	90.00
JM	2	18	20	90.00
GC	3	17	20	85.00
HO	3	11	14	78.57
JO	5	12	17	70.59
CD	8	12	20	60.00
SM	8	11	19	57.89
SS	8	11	19	57.89
MG	9	11	20	55.00
AK	9	10	19	52.63
L	10	10	20	50.00
SB	7	7	14	50.00
BH	11	6	17	35.29
MEAN	**5.44**	**12.38**	**17.81**	**70.07**

Although this binary categorisation into canonical-coloured and letter-coloured is useful for obtaining an overview of general preference, it does not capture important nuances. For instance, while the synaesthetic colour associated with a colour term such as "red" may be closer in colourspace to the control word "reed" than to canonical

RED, the canonical colour may yet be having an influence on more subtle aspects, such as its lightness and saturation. Figure 2 illustrates this below.

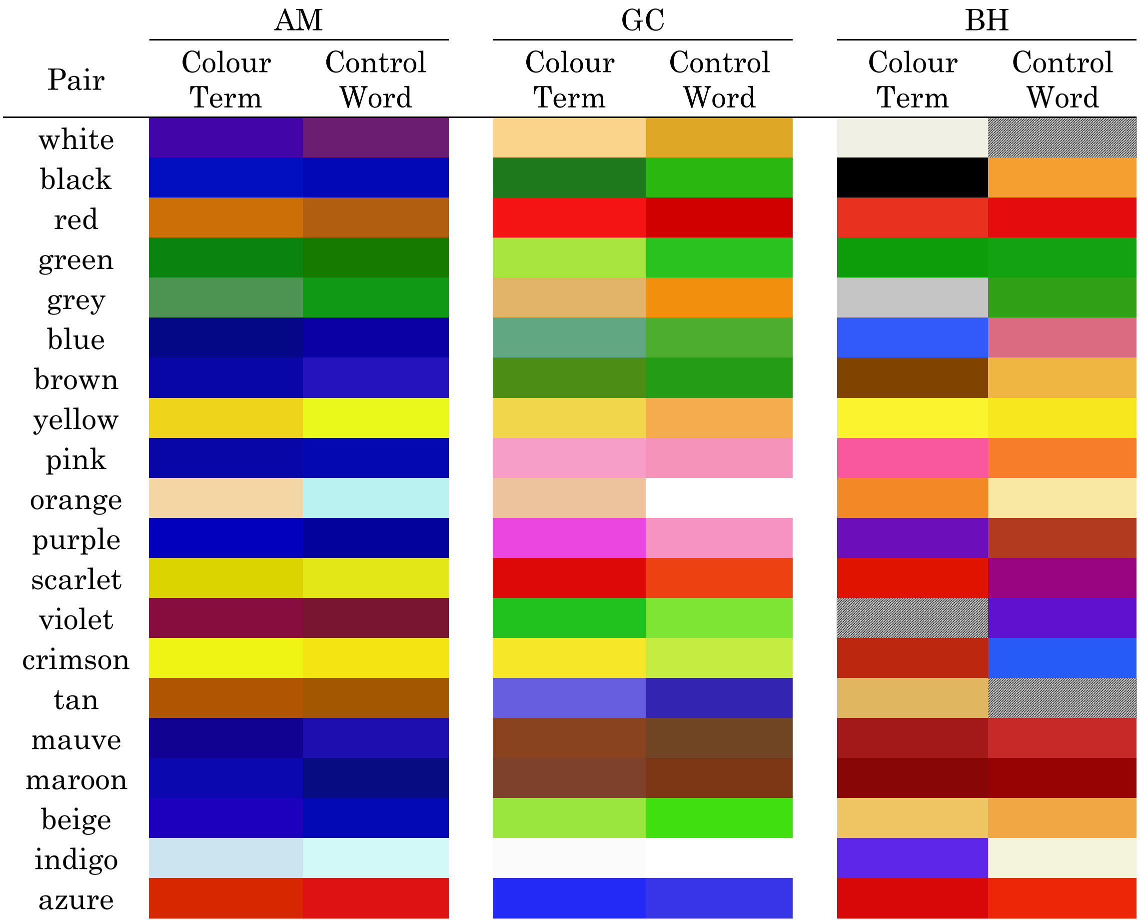


Figure 2. Colours given for colour terms (left side of each colour column) and control words (right side of each column), in order from highest to lowest frequency colour term. Colour term/control word pairs (e.g., "red"/"reed") are labelled in the first column by their colour term only; a full list of word stimuli can be found in Appendix B. Diagonally barred cells indicate no synaesthetic colour given. The three synaesthetes demonstrate strong letter-colouring (AM), influence on the colour term by the canonical colour (GC), and strong canonical-colouring (BH).

Synaesthete AM (first colour column) shows an almost perfect correspondence between control words and colour terms. Here, the similarity between "blue" and BLUE, or "yellow" and YELLOW, would be due to the common pattern among synaesthetes of associating the first letter of a colour term with its denoted colour (e.g. *B* with blue, *R* with red, etc.; c.f. Mankin & Simner, 2017 [Chapter 5]; Rich et al., 2005; Simner et al., 2005). This shows that AM colours colour terms by their letters and not by semantics. Synaesthete GC (centre colour column) also has similar colours for control words and colour terms; however, the colour for "white" is lighter (i.e. closer to WHITE) than the colour for "whip", "black" is darker than "blade", "red" is brighter red, "blue" is bluer, "grey" is greyer, etc. This illustrates the possibility that the canonical colours of colour terms "pull" the synaesthetic colours closer to the canonical colour. Finally, synaesthete BH shows a strong preference to use the canonical rather than the letter colour, but with a possible influence of colour term frequency. That is, high-frequency colour terms are coloured canonically, while unusual colour terms are similar to their control words. For example, very low-frequency "azure" is coloured like control word "aztec" rather than canonical AZURE (i.e., sky-blue).

What pulls the synaesthetic colour towards the canonical colour?

To explore these influences in more detail, we next constructed a linear mixed effects (LME) model. LME models account for both random variation (i.e. from participants and items) and fixed effects (e.g. frequency and imageability measures) and are therefore well-suited to the analysis of psycholinguistic effects (Baayen et al., 2008). They are also particularly useful for studying synaesthesia, as the random effect of participants can help account for individual differences between synaesthetes (Hamada et al., 2017). The current LME analysis and that in Experiment 2 were conducted in R (R Core Team, 2016) using *lme4* (Bates et al., 2015) and *languageR* (Baayen, 2013) with stepwise model testing and *p*-values for effects obtained with *lmerTest* (Kuznetsova et al., 2016) and interaction figures graphed with *visreg* (Breheny & Burchett, 2016).

For this model, we wanted to know when the canonical colour might have a measureable impact on the synaesthetic colour of the colour term. To measure this influence, we first calculated an outcome measure to quantify the pull of the canonical colour. We can illustrate this with the colour term "indigo" and its control word "inch". We created this *canonical pull* outcome measure by first calculating the Euclidean distance between the colour term and its canonical colour (e.g. the distance between the colours of "indigo" and INDIGO) and between the control word and the same canonical colour (e.g. "inch" and INDIGO). We then subtracted the first distance from the second. This captured the change in distance due to the "pull" of the canonical colour on the colour term (see Figure 3, below). So, assuming that "inch" represents a typically synaesthetically coloured word, this allows us to model what might pull the synaesthetic colour of "indigo" towards INDIGO, and whether this affects some synaesthetes more than others. For example, for synaesthete AM (left panel of Figure 3, below), the colour term "indigo" is coloured like its control word, so the canonical colour makes little difference. Accordingly, the canonical pull of

INDIGO is small (difference between the distances = 10.68). For synaesthete BH (right panel), the canonical colour exerts a strong pull on the colour term, so “indigo” is much closer to INDIGO than its control word is (difference = 106.03).

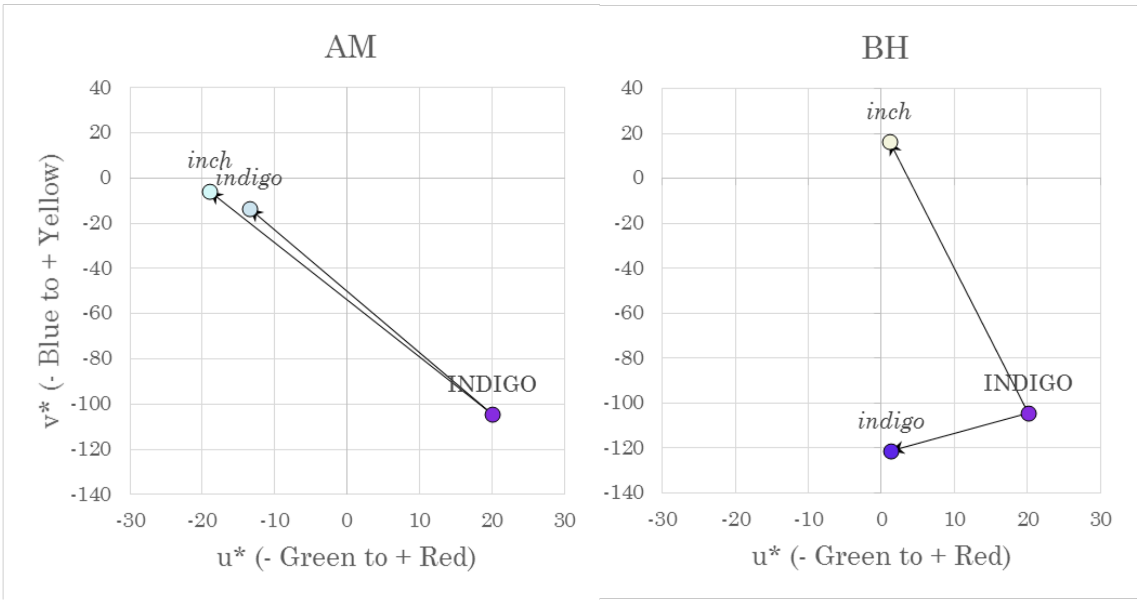


Figure 3. Illustration of *canonical pull* for two synaesthetes, AM (left panel) and BH (right panel), and the “indigo”/“inch” inducer pair. Points labelled “indigo” and “inch” represent a location in CIELuv colourspace and are coloured to match the synaesthetic colour provided by the participants. Points labelled INDIGO are the canonical colour of indigo obtained from non-synaesthetes (same in both panels). The length of the arrows connecting INDIGO to both “indigo” and “inch” represent the Euclidean distance in colourspace between them. The canonical pull outcome measure is calculated by subtracting the length of the INDIGO → “indigo” line from the INDIGO → “inch” line for each synaesthete. For clarity, only the hue dimensions u* and v* have been included in the figure, collapsing L (lightness), and u* and v* axes are not to scale.

So what might pull a colour term closer to its canonical colour? First, we predicted that increasing frequency of the colour term would predict a greater canonical pull; that is, the more common a colour term is, the stronger the pull of the canonical colour would be. This would mean that high-frequency colours like “red” would tend to be synaesthetically red, while low-frequency terms like “indigo” may not be indigo-coloured. However, we also tested for an interaction with the distance between the control word and canonical colour (e.g. the distance between “inch” and INDIGO in Figure 3, above), which we called *control-canonical distance*. This was to control for the instances where the colour term (e.g. “red”) was already close to canonical RED simply because the control word started with a similarly coloured letter (e.g. *R* is red, so “reed” is also). This would obscure the canonical pull that RED might otherwise have for some synaesthetes.

We also explored the influences of lightness and saturation. We wanted to know whether the strength of the canonical pull was modulated by the colours themselves. We therefore entered terms and interactions for the lightness of the canonical colour (RED) and the lightness of the control word colour (“reed”), and another for the saturation of both. Finally, we predicted that the frequency of the colour term might be more influential when the canonical colour was also a very light or saturated colour (i.e. more salient). We accordingly tested whether the frequency of the colour term interacted with the lightness and saturation of the canonical colour. We used a stepwise comparison to remove nonsignificant predictors one at a time, starting with the highest-order interactions; nonsignificant main effects were retained if they were part of a significant interaction. The final model is summarised in Table 4, below.

We also specified random intercepts for both participants and items. This element of the mixed effects model gives each participant or item its own intercept in the linear model. We tested the inclusion of the random intercepts using log-likelihood ratio tests between models with and without the random effects (Baayen et al., 2008). The random intercepts for participants significantly improved the model, indicating that synaesthetes experienced the canonical pull to different degrees; however, random intercepts for items did not improve the model and were removed.

Table 4. Summary table of the LME model describing the predictors that influence *canonical pull*, expressing how much closer the target word (e.g. a colour term like “red”) is to its canonical colour (e.g. RED) than its control word is (e.g. “reed”). Significant *p*-values are marked in **bold**.

	Canonical Pull			
Fixed Effects	Estimate (B)	SE (B)	*t*	*p*(t)
Intercept	-35.52	6.03	-5.89	**<.001**
Canonical Colour Saturation	-4.39	2.09	-2.10	**.036**
Target Word Frequency	-1.31	4.55	-0.29	.774
Control-Canonical Distance	0.54	0.04	12.53	**<.001**
Colour Term Frequency x Control-Canonical Distance	0.09	0.04	2.16	**.031**
Random Effects	Variance	SD	χ^2	$p(\chi^2)$
Participant	273.9	16.55	32.31	**<.001**
Residual	1186.3	34.44	--	--

The model has two main elements, which we will discuss in order. First, there was a negative influence of canonical colour saturation. This means that when the canonical colour is more vividly coloured (i.e. higher in saturation), canonical pull decreases. In other words, high-saturation canonical colours like RED do not pull the synaesthetic colour of "red" toward themselves, but rather push them away toward "reed". Why might this be the case, when we had expected the opposite (i.e. that high-saturation RED would exert a stronger pull on "red" than BLACK on "black")? We suggest that this captures a preference for letter-colouring when forced to make a choice between two very different colours for a particular word. Consider, for example, the colour term "black", which has a low-saturation canonical colour, and a synaesthete for whom words beginning with *B* are usually blue, including the control word "blade". When given only the option to report a single colour for "black", the synaesthete must choose between two simultaneous colour sensations (canonical BLACK and letter-based blue). Here, the synaesthete can compromise by reporting a darker shade of blue, which effectively captures both colours. However, in the case of a colour term like "crimson" with both a highly saturated canonical colour and a very clear mismatching control word colour (e.g. "yellow" for cricket), there is no way to compromise. Halfway between the two (here, orange) would require a change of colour category, the way a darker blue for "black" does not. Therefore, the synaesthete is forced to report a colour that matches *either* the canonical colour (CRIMSON) *or* the letter-based colour (yellow), a choice made starker as the canonical colour becomes more saturated. Our results show that synaesthetes in this situation choose a colour for the colour term that matches their synaesthetic letter-colour more closely. The saturation of the control word did not affect this, as the interaction between control word and canonical colour saturation was not significant.

The second significant element in our model was the interaction of colour term frequency and control-canonical distance. As the model summary in Table 4 shows, the effect of colour term frequency was not significant on its own. In other words, the frequency of "red" alone did not influence whether it was pulled towards canonical RED or not. However, as we had expected, this effect was significant in combination with the distance between the control-word and canonical colours (e.g. between "reed" and RED, which we have called *control-canonical distance*). Figure 4 illustrates this interaction with simple slopes. The interaction indicates that when the control word is already close in synaesthetic colour to the canonical colour (e.g. "reed" is similar in colour to RED, in our terms a low control-canonical distance), then the frequency of the colour term "red" does not affect canonical pull. That is, the frequency of "red" has no effect because "red" and "reed" are both already similarly coloured to RED. However, when the control word is very different from the canonical colour (i.e. high control-canonical distance), higher-frequency colour terms experience a stronger canonical pull as we predicted. In this case, if “reed” is synaesthetically purple because of a purple *R* (i.e., control-canonical distance is high), the higher frequency of the colour term “red” will mean that its synaesthetic colour is pulled more strongly toward canonical RED. However, low-frequency colour terms like “azure” are not pulled towards their canonical colour so strongly. In essence, the synaesthetic colours of higher frequency colour terms are in fact pulled more strongly towards their canonical colours, but only when we account for how similar those colours are to begin with.

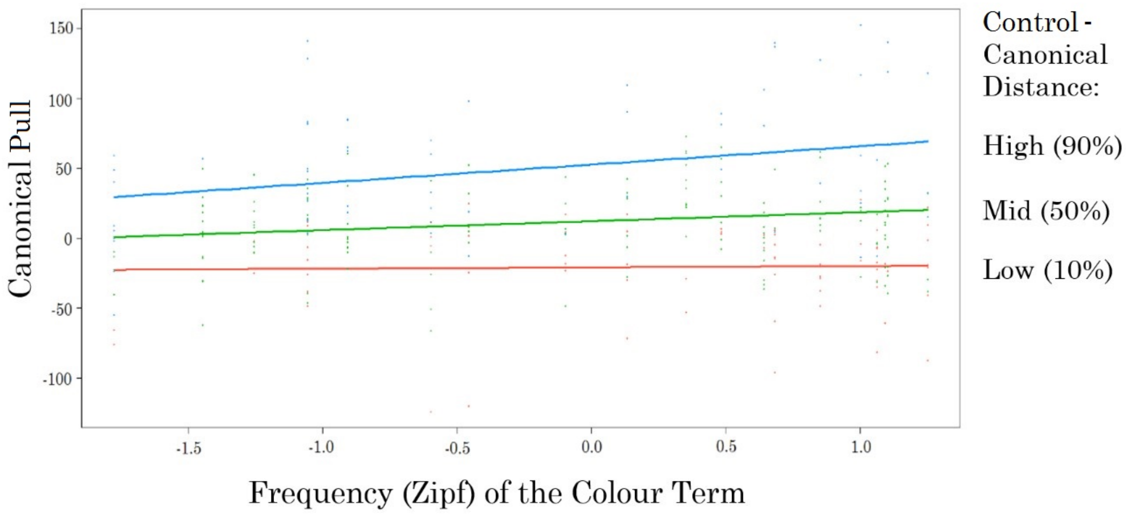


Figure 4. Interaction of colour term frequency (grand mean centred) and control-canonical distance from the LME model, depicted with simple slopes. This illustrates that when control-canonical distance is low, colour term frequency has no effect. However, when control-canonical distance is high (i.e., the control word and canonical colour term are far apart from each other in colourspace), the influence of colour term frequency is clearly evident.

Conclusions

Our study shows that both the colour and linguistic frequency of colour terms like “red” influence the colours that synaesthetes experience for them. First, we found that colour terms like “red” are pulled towards their semantic colour (e.g. canonical RED), and this varies in degree from synaesthete to synaesthete. We found evidence of the alien colour effect in all our synaesthetes, which suggests that the conflict between canonical and letter-based colours for colour terms is a common experience among synaesthetes. We next used a linear mixed effects model to show that when the canonical colour is highly saturated, creating a strong categorical distinction between the two possible colours (e.g. canonical RED vs letter-based purple for “red”), synaesthetes tend to report their letter-based colour over the canonical colour. Our results also indicated that higher-frequency colour terms, such as “red”, tended to be more like their canonical colour than uncommon colour terms like “azure” were. This could only be seen, however, when we controlled for the similarity between the letter-based colour and the canonical colour (e.g. how close "reed" was to RED in colourspace). When the letter-based, control-word colour was very different to the canonical colour, the influence of colour term frequency became apparent. These complex influences suggest that synaesthetes experience at least two strong colour associations for colour terms – the canonical colour and the letter-based colour – either of which may exert a stronger influence on a particular colour term depending on the saturation of the colours and the frequency of the colour term. Overall, our results show clear support for the idea that the canonical colour of words – here, the colour denoted by a colour term – has a measurable impact on synaesthetic colour experiences. We will next ask whether the mental images associated with highly imageable words might have a similar influence on synaesthetic colour.

Experiment 2: Imageability and canonical colour

Our second experiment continues our investigation of imageability in synaesthetic colour associations. As we demonstrated above, synaesthetes experience a conflict in the associated colour for words with a strong canonical colour – that is, colour terms and their canonical colours, such as "red" and the canonical colour RED. The strongly evoked canonical colour can cause differences in the synaesthetic colour between words that would otherwise be coloured very similarly to each other based on their spelling (e.g. "red" and "reed"). We described this as the *pull* of the canonical colour. Here we will continue our exploration of this effect by comparing the synaesthetic colours for matched pairs of words, but this time contrasted primarily by their imageability. Imageability is the ease of forming a mental image of the word's concept (e.g. high-imageability "fire" vs low-imageability "fine"). We expect that, as with colour terms and control words, these two groups of words may differ in their synaesthetic colours due to the colours of the mental images they may evoke.

Methods

Participants

The same twenty-three participants completed this experiment as in Experiment 1, above. As in Experiment 1, three were removed from subsequent analysis because the test erroneously recorded black for every item. Participants were each paid £10 total for their participation in both this and the previous experiment.

Materials

The wordlist comprised 20 pairs of matched monomorphemic words (total N = 40). One member of the pair was high in imageability and the other low (e.g., high-imageability "fire" vs low-imageability "fine"). Imageability ratings were obtained from the N-Watch psycholinguistic research tool (Davis, 2005) collated from several norming studies (Bird et al., 2001; Gilhooly & Logie, 1980; Stadthagen-Gonzale & Davis, 2006). These ratings range between 100 (very difficult to mentally picture) and 700 (very easy to mentally picture). Our high-imageability target words all had ratings greater than or equal to 475 (lowest rating: "dirt", 475) while the low-imageability control words were all rated below 425 (highest rating: "whine", 424). We matched words on their initial graphemes such that both words in a pair shared at least their first consonant and first vowel (e.g., "prong"/"prow", "chain"/"change"). The only minor exception was "wine"/"whine"; while both the first consonant and vowel did match, "whine" had a second consonant, *H*, between them. Besides this, all words had the same first two letters at least (mean number of matched initial letters = 2.74, SD = 0.65). We also matched overall spelling as closely as possible (mean number of letters in common = 3.1, SD = 0.64; mean word length = 4.23 letters, SD = 0.67). These words were groupwise balanced on Zipf frequency from the SUBTLEX-UK corpus and on length, but were strongly differentiated by age of acquisition (AoA; see Table 5). AoA is known to be highly correlated with imageability (Bird, Franklin, & Howard, 2001), so this was expected. A full list of the target and control words with balancing measures can be found in Appendix C.

Table 5. Descriptive statistics for the psycholinguistic and orthographic measures for high-imageability target words and low-imageability control words, designated "Target" and "Control" respectively. We report independent t-tests for each measure between target and control. Significant *p*-values are highlighted in **bold**.

Measure	Word Type	M	SD	*t*	*p*(t)
Imageability	Target	576.2	56.09	14.71	**< .001**
	Control	315.3	56.05		
AoA	Target	292.29	68.51	-5.76	**< .001**
	Control	463.13	110.56		
Length	Target	4.10	0.64	-1.21	.236
	Control	4.35	0.67		
Frequency (Zipf)	Target	4.07	0.90	0.27	.787
	Control	3.99	0.97		

Apparatus and procedure

Data collection for both this and Experiment 1 took place at the same time, using the same apparatus and procedure described previously. The wordlists for both experiments were interleaved together and presented in unique random order for each participant.

Canonical colours

The same ten non-synaesthetes completed the word-colour test as described in Experiment 1, above. These were all native speakers of English, six female and eight right-handed, with a mean age of 32.1 years (SD = 8.70 years). As described in Experiment 1, they were instructed to “select the first colour you think of - whatever pops into your head first” and this should be the “first, most automatic” colour response to each word. Otherwise, the test and procedure were identical.

Results and Discussion

Data validation

First, we transformed all colours to the CIELuv colourspace, using the same specifications as in Experiment 1. As before, we also calculated the saturation of each colour (Schanda, 2007). As in Experiment 1, our primary outcome measure was the *canonical pull* for each high-imageability word. This measure reflects the change in the high-imageability target word’s colour due to the influence of its semantic-based canonical colour. We measured canonical pull by calculating the distance between the synaesthetic colour of the target word, for example “fire”, and its canonical colour (FIRE), and similarly between the control word “fine” and FIRE. We then subtracted the first distance from the second. For example, the synaesthetic colour of the word “fire” may be pulled toward the canonical fiery orange of FIRE with reference to a similarly spelled control word like “fine”, which we assume is typically coloured by its letters (e.g. blue if *F* is blue). Since colours for both the high and low-imageability words in a pair were necessary to calculate this measure, we examined how many of each synaesthete’s responses out of the 20 target/control word pairs were missing one or both colours. On average, synaesthetes had 16.05 complete pairs (80.25%), but four synaesthetes (AG, BM, CT, and MT) were missing 10 or more pairs. These four synaesthetes were removed, leaving 16 synaesthetes in the following analysis.

Calculating canonical colours

We were next interested in calculating the canonical colours for our list of high-imageability target words (e.g. “fire”), as we had done for the canonical colours of colour terms in Experiment 1 (e.g. “red”). However, many of the non-synaesthete participants had selected black for many items, unlike in Experiment 1. This difference in colour responses between experiments may mean that some participants found it easier to provide colours for colour terms (e.g. "red") than for high imageability words, the colour associations for which may be less clear. Therefore, we could not exclude participants based on the number of black responses, as we had done before. In order to retain enough data for this canonical colour calculation, we instead removed all black colour responses across all of the non-synaesthete participants. In this way, we retained the maximum amount of colour information[14]. This meant that some canonical colours were calculated using fewer colour responses than others (minimum = 5, maximum = 9). We then calculated each centroid CIELuv canonical colour by taking the average of each of the L*, u*, and v* dimensions, which gave us a single canonical colour for all 20 of the high-imageability words. These canonical colours are presented in Table 6, below.

Table 6. Canonical colours for the twenty high-imageability words, calculated as the centroid colour for each word. Canonical colours are ordered alphabetically. RGB values are provided for reference.

Word	Canonical Colour	R	G	B	Word	Canonical Colour	R	G	B
BOAR		111	88	92	HIVE		180	197	106
BONE		246	243	234	IRON		163	142	135
BULL		162	97	97	MINT		111	194	124
CHAIN		179	174	157	PIG		231	178	210
DIRT		109	74	42	PRONG		147	155	162
FACE		215	196	203	RAT		111	100	91
FEAST		153	114	109	SEAT		181	78	59
FIRE		235	117	57	STUMP		138	84	65
FLAME		221	129	93	WAX		232	226	190
HARE		179	153	121	WINE		167	38	50

Data analysis: What synaesthetic colour do high-imageability words have?

We first wanted to see if the colours that synaesthetes experience for whole words can be influenced by the colour of the mental image associated with that word. As we had done in Experiment 1, we began with a comparison of colourspace distances. We measured the Euclidean distance between each high-imageability target word and its canonical colour, which we had obtained from the controls (e.g. the

---

[14] This calculation may have also penalised potentially darker canonical colours (e.g. "bull", "rat"), for which some of the black responses may have been genuine. However, we could not reliably distinguish these from problematic ones (e.g. black for "fire", or "mint"), so we chose a method that favoured consistency across all target words, as the goal was to calculate canonical colours objectively.

synaesthetic colour for "fire" to the canonical colour FIRE). We then compared this to the colourspace distance between the control word and the same canonical colour (e.g. "fine" → FIRE). Using a paired-samples t-test as in Experiment 1, we found that high-imageability target words (e.g. "fire") were closer to their respective canonical colours (e.g. FIRE) than their control words were ["fire" → FIRE vs. "fine" → FIRE Mdiff = -5.78, paired-samples $t(283) = -2.98$, $p = .003$]. This is the same result as in Experiment 1 for colour terms (e.g. "red"), which suggests the same process may be at work, although the effect size for this difference was small in general as well as smaller than the same comparison for colour terms [paired-samples $d = 0.20$]. We then asked whether there was a difference in canonical pull for colour terms (from Experiment 1) versus high-imageability words. That is, does RED exert a stronger influence on the synaesthetic colour of "red" than FIRE does on "fire"? To test this, we compared the canonical pull outcome measures from Experiment 1 ("red") versus the current experiment ("fire"). This analysis confirmed that on the whole, colour terms like "red" were pulled more strongly towards their canonical colour than high-imageability words like "fire" were [Mdiff = -8.63, Welch's $t(523.26) = 2.56$, $p = .011$], although this effect was also small [$d = 0.22$]. Altogether, these analyses showed that canonical colours have an effect on the synaesthetic colours of both colour terms and high-imageability words, but this influence is slightly stronger for colour terms.

To further explore this, we next looked at individual differences between synaesthetes by comparing the distances between each target word and its canonical colour (e.g. "fire"→ FIRE) or its control word (e.g. "fire" → "fine"). If the target word "fire" was closer to the colour of canonical FIRE than it was to its control word "fine", we coded this as *canonically coloured*. If the reverse were true, and "fire" was coloured more like its control word "fine", we coded this as *letter-coloured* (as it more closely matches the typical synaesthetic colour based on the colours of the letters). This letter-colour source is analogous to the *alien colour* we investigated for colour terms in Experiment 1. We then calculated the percentage of the total number of target words that were closer to the low-imageability control word rather than the canonical colour. The results of this comparison are summarised in Table 7, below, which show that synaesthetes again varied in the extent to which their synaesthetic colours are pulled towards the meaning-based canonical colour of words. As we had made the same comparison between canonical- vs letter-colour sources for colour terms (e.g. "red") in Experiment 1, we compared the percentage letter-coloured across the two experiments (called "Percentage ACE" in Experiment 1). The correlation between the percentage of letter-coloured words between experiments showed that the same synaesthetes who strongly experienced the alien colour effect in Experiment 1 also tended to have more letter-coloured words in the current experiment [Spearman's ρ= 0.65, p = .007].

Table 7. Counts and percentages of high-imageability word colours closer to the canonical colour ("fire" closer to FIRE coded as *canonically coloured*) or the low-imageability word colour ("fire" closer to "fine" coded as *letter-coloured*), with the total number of complete pairs. Note that many synaesthetes had at least one missing colour within the colour term/matched word pair, so some totals are less than the original number of matched pairs (N = 20). Here "Percentage Letter-Coloured" is the number of letter-coloured target words out of the total number of complete pairs for each synaesthete.

Synaesthete	Canonically Coloured	Letter-Coloured	Total	Percentage Letter-Coloured
GC	0	20	20	100.00
TR	0	19	19	100.00
AM	1	19	20	95.00
MG	1	19	20	95.00
CSP	1	13	14	92.86
CS	2	18	20	90.00
JM	2	18	20	90.00
HO	2	13	15	86.67
SM	3	17	20	85.00
AK	5	14	19	73.68
BH	5	9	14	64.29
CD	8	12	20	60.00
SS	8	12	20	60.00
L	8	11	19	57.89
SB	10	4	14	28.57
JO	12	3	15	20.00
MEAN	**4.25**	**13.81**	**18.06**	**74.93**

What pulls the synaesthetic colour towards the canonical colour?

As we had done in Experiment 1, we returned to our *canonical pull* outcome measure for each target/control word pair for each synaesthete. We calculated the Euclidean distance in CIELuv colourspace between the high-imageability word ("fire") and its canonical colour (FIRE), and we subtracted this from the distance between the low-imageability word ("fine") and the same canonical colour. The resulting difference expressed the change in the synaesthetic colour of "fire" due to the influence, or *pull*, of the canonical colour. Our analysis asked what would exaggerate this pull – that is, what might cause "fire" to be coloured more like FIRE when it would typically be coloured like "fine". In essence, we treated the canonical colours of our high-imageability words (FIRE) like the canonical colours of colour terms from Experiment 1 (RED). Therefore, as in Experiment 1, we constructed a linear mixed effects model with similar predictors, with canonical pull as the outcome measure.

In our model we expected that the same factors would be pertinent as in Experiment 1. That is, the strong correlation reported above between percentage ACE from Experiment 1 and percentage letter-coloured in Experiment 2 indicated that the same influences were at work for both colour terms and high-imageability words. However, we expected that the effects would be stronger and clearer for colour terms than they are for high-imageability words. This is because the canonical colour associations for high-imageability words are not as salient in word processing as those of colour terms, as we discussed in the introduction (see especially Risko et al., 2006; Schmidt & Cheesman, 2005). To test this, we looked for the same effects in this model as were significant in the model reported in Experiment 1. Specifically, we hypothesized that target words with higher frequency (e.g. "fire") would be pulled more strongly towards their canonical colours than lower-frequency target words (e.g. "prong"). Unlike with colour terms, however, we did not need to control for the distance between the control word and the canonical colour (i.e. control-canonical distance). This is because the control words in Experiment 1, such as "reed" for "red", were already likely to match the canonical colour RED because *R* is likely to be red for synaesthetes (and *B* blue, *G* green, etc.; Rich et al., 2005; Rouw et al., 2014; Simner et al., 2005; Simner & Ward, 2008). Here, the control words have no such inherent association – that is, there was no a priori reason to assume that the synaesthetic colour of "fine" would be likely to be coloured like FIRE based on its spelling. Our second finding in Experiment 1 was the main effect of canonical colour saturation, so we also included canonical colour saturation in our model.

As before, we included random intercepts for participants and items in the model, but the random effect of items did not significantly improve the model and was removed. The random effect of participants did significantly improve the model and was retained. Similarly, the fixed effect of canonical colour saturation was also non-significant and was removed [B = -3.76, SE(B) = 3.16, t = -1.19, p = .236], yielding the final model summarised in Table 8, below. However, model diagnostics indicated significant deviations from normality of the model residuals [W = 0.93, p < .001], and further investigation revealed that that the model did not capture the higher and lower ends of the distribution well (i.e. a leptokurtic distribution of residuals). As this had not been the case with our previous model for colour terms (see Experiment 1), it suggests that additional factors may be influencing high-imageability word colour that we have not been able to account for, and therefore the model should be interpreted with caution. We will address this more fully in our general discussion.

Table 8. Summary table of the LME model describing the predictors that influence *canonical pull*, expressing how much closer the high-imageability target word (e.g. "fire") is to its canonical colour (FIRE) than its control word is (e.g. "fine"). Significant *p*-values are marked in **bold**.

	Canonical Pull			
Fixed Effects	Estimate (B)	SE (B)	*t*	*p*(*t*)
Intercept	6.10	2.70	2.26	**.038**
Target Word Frequency	8.45	3.04	2.79	**.006**
Random Effects	Variance	SD	χ^2	$p(\chi^2)$
Participant	60.3	7.77	5.70	**.017**
Residual	980.4	31.31	--	--

The model showed that the frequency of the target word was again a strong predictor of canonical pull. This means that more frequent target words, such as "fire", tended to be pulled closer to their canonical colour than low-frequency words such as "prong". This is what we had expected, as we found the same effect of frequency for colour terms in Experiment 1 once we accounted for control-canonical distance[15]. Together, these frequency effects indicate that the colour of these words (e.g. the RED of "red" or the fiery colour of "fire") is evoked every time they are processed, and the more often they are evoked, the stronger the influence of that canonical colour becomes.

Unlike Experiment 1, the saturation of the canonical colour had no significant influence in this model. This may reflect the fact that overall, the canonical colours gathered from non-synaesthetes (e.g. FIRE, BONE, WAX, FACE) were less saturated than the synaesthetic colours of low-imageability words ("fine", "bond", "wad", "facet") [canonical vs low-imageability saturation: MDiff = -0.78, Welch's $t(481.04) = 11.18$, $p < .001$, $d = 1.02$]. They were also overall less saturated than the canonical colours for colour terms from Experiment 1 [MDiff = 0.78, Welch's $t(32.13) = 2.84$, *p*

[15] That is, in Experiment 1 we had to control for the fact that the control word was likely to be canonically coloured (e.g. "reed" is often red due to a red *R*), so the effect of target word frequency came out only in the interaction with control-canonical distance (e.g. the distance between "reed" and RED). Here, we did not need to control for this because "fine" is not often fiery orange, so this was a simple main effect that captured the same influence of frequency.

= .008, d = 1.00]. We suggest that the lower saturation of these canonical colours meant that they were not as salient the canonical colours in Experiment 1. Therefore, it is not surprising that canonical colour saturation was not a significant influence here.

Conclusions

This experiment extended our investigation of semantic influences on synaesthetic colours to words contrasted primarily by their imageability (e.g. high-imageability "fire" vs low-imageability "fine"). Here we compared the colours that synaesthetes experienced for these words to discover whether the canonical colour associated with the high-imageability target word (e.g. fiery orange for "fire") might influence the synaesthetic colour experienced for that word. Our LME analysis modelled the pull of the canonical colour – that is, how does the canonical colour of FIRE influence the synaesthetic colour for the word "fire"? We found that words like "fire" are indeed pulled towards their canonical colour, but that this varies in degree from synaesthete to synaesthete. We also found that increasing frequency of the target word ("fire") pulled the high-imageability word closer to its canonical colour in colourspace relative to its control word. This lines up well with our results from Experiment 1, indicating that language processing is an intrinsic element of synaesthetic word colouring.

General discussion

This study set out to investigate how the canonical colours associated with a word (e.g. the colour of "red", or the fiery orange colour for the word "fire") may influence its synaesthetic colours. In Experiment 1, we looked at the clearest example of canonical colour associations, namely colour terms (e.g. "red"). Our second experiment did the same, but using high-imageability words (e.g. "fire"). For each set of words, we analysed the colours reported by synaesthetes using CIELuv colourspace distances. We empirically established the canonical colour of our colour terms (e.g. what is canonical RED, or FIRE?) by asking non-synaesthetes to give colours for these words. We then investigated whether our synaesthete participants tended to colour the words "red" and "fire" more like their canonical colour (e.g. "red" coloured RED) or like the control word (e.g. "red" coloured like "reed"). In this comparison, we assumed that the colour of the control word represents the typical synaesthetic colour that the synaesthete would experience for any word with similar spelling. We found that overall, target words like "red" and "fire" were closer in colourspace to their respective canonical colours (i.e. RED and FIRE) than their control words were. We then quantified this influence of the canonical colour as *canonical pull*. We defined this measure as the difference between two distances: the distance between the control word and canonical colour (e.g. between "reed" and RED), and the distance between the target word and the same canonical colour (e.g. between "red" and RED). This difference expressed the change in the synaesthetic colour of "red" due to the influence, or *pull*, of the canonical colour. In this way, we investigated how the semantic associations of a word might influence its synaesthetic colour.

When canonical pull was weak in Experiment 1, this represented the *alien colour effect* (ACE; Gray et al., 2002, 2006), in which synaesthetes experience a colour based on the letters in the word (e.g. "red" coloured like *R*) rather than the canonical colour (e.g. "red" coloured RED). Here, we give a detailed explanation of this effect in grapheme-colour synaesthetes for the first time. Our results showed that all of our synaesthetes experienced the ACE to some degree – that is, for all of our participants, at least 30% of the synaesthetic colours for colour terms were closer to their letter colour (i.e. control word) than their canonical colours. These results indicate that the typical synaesthetic colour of a word based on its letters is a very strong influence on the word's colour, frequently overriding the canonical colour of the colour term. This, along with feedback from participants volunteered in the online test (discussed further below), suggests that there is a very real conflict between the canonical and letter-based synaesthetic colour in determining the colour for the whole word. That is, it appears that synaesthetes are aware of both the canonical and letter-based colours, and must struggle to determine which is the synaesthetic colour of the whole word depending on both the colours and the words involved.

We then built a linear mixed effects model to investigate what influences may pull the colour of the target word (e.g. "red") closer to the canonical colour (RED). Our results show that there is a complex interaction between word frequency and canonical colours that together influence the synaesthetic colour overall. We found that a colour term is coloured more strongly like its canonical colour (e.g. "red" is coloured like RED) when it is a high-frequency colour term, as we had expected. However, this effect only became clear when we extracted out situations where the synaesthetic colour of a word and its canonical colour were already very similar to each other. That is, if "reed" was already red because of a red *R*, then the frequency of "red" was irrelevant. In this case, there was no conflict to resolve because both "red" and "reed" were already coloured RED. On the other hand, if "reed" was a very different colour (for example, purple), then "red", being a high-frequency colour term, tended to be pulled more strongly towards canonical RED. However, a low-frequency colour term like "azure" was more likely to match the colour of its control word ("aztec") even when "aztec" was a very different colour from canonical AZURE. This supports our hypothesis that colour terms like "red" are more likely to synaesthetically match their canonical colours due to the combination of frequency and strongly evoked canonical colour. This can be seen in the schematic below, which shows a clear canonical pull only when the control word ("reed") is very different from the canonical colour (RED) to begin with, and the paired colour term ("red") is high in frequency.

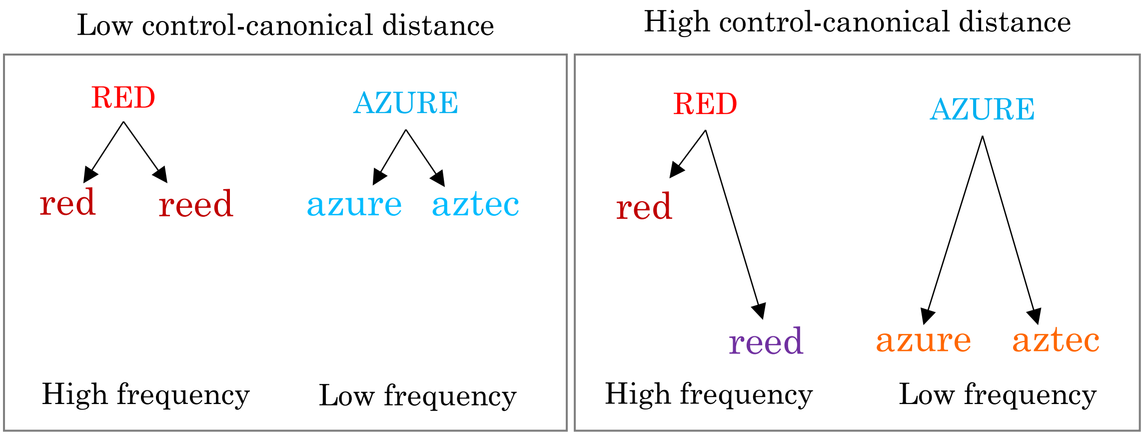


Figure 5. Schematic of the frequency x control-canonical distance interaction.

In our analysis of colour terms, we had also expected that more saturated (i.e. more vividly coloured) canonical colours would pull the synaesthetic colour of a colour term towards them. For example, "red" would be pulled towards RED because RED is a strong, vivid colour, while BEIGE would have a weaker pull on "beige". We instead found the opposite effect. This meant that more saturated canonical colours actually *pushed* colour terms like "red" away towards their letter-based colour. We suggest that this may be due to the experimental design only allowing the synaesthete to choose a single colour for each word. Our results show that when synaesthetes were forced to choose between the usual synaesthetic colour and a vibrant canonical colour, the letter-based synaesthetic colour predominated. In particular, this could partly be due to the instructions, which asked synaesthetes to report the *synaesthetic* colour they experienced, and may have been interpreted as asking them to prefer their letter-based synaesthetic word colour in case of this sort of conflict. This effect highlights the difficulty with restricting synaesthetes to a single colour choice, when this does not reflect their actual synaesthetic experience. Indeed, at the end of our experiment, where synaesthetes were given space for feedback, a common point of contention was that a colour palette with only one colour per word was inadequate to represent their synaesthesia accurately. For example, one synaesthete explained that "[t]he color … appears as more of a combination of all the colors in my head." Another characterised their word colours as "dynamic and shifting colours." It is clear that the single-colour study does not capture some aspects of synaesthetic colour that the synaesthetes themselves find meaningful, and the development of a better apparatus warrants further attention.

Overall, our first experiment showed that the alien colour effect is quantifiable using colourspace distances and dependent on both linguistic and perceptual influences. In our second experiment, we then explored whether non-colour-terms were also subject to such effects. Using the same experimental paradigm and analysis, we found that the canonical colours of high-imageability words (e.g. "fire") do have an influence on the synaesthetic colours for those words, analogous to the canonical colours of colour terms. We showed that the percentages of letter-based word colours for the same synaesthetes across our two experiments were strongly correlated, suggesting that the same synaesthetic colouring processes are at work for both colour terms and high-imageability words. However, by comparing the size of the canonical pull between the two experiments, we showed that canonical colour had a stronger influence on colour terms than on high-imageability words – that is, "red" was pulled more strongly towards RED than "fire" was pulled towards FIRE. In other words, both colour terms and high-imageability words evoke a canonical colour in a similar way, but the canonical colour is more influential for colour terms like "red" than it is for high-imageability words like "fire". This lines up well with results from Stroop studies with non-synaesthetes, which have shown that "colour associates" like "fire" cause more interference in colour naming than unrelated words (e.g. "fine"), but less than colour terms like "red" (Klein, 1964; Proctor, 1978; Risko et al., 2006; Schmidt & Cheesman, 2005). Why might it be that canonical colours have a greater impact on the processing of colour terms than on high-imageability words? We suggest that this may be because the primary – indeed, the only – defining characteristic of the meaning of the word "red" is its colour. However, for words like "fire", colour is only one part of its evoked meaning, alongside other visual information such as brightness, tactile information such as heat and pain, etc. In other words, the meaning of “red” is completely composed of the colour RED, whereas “fire” is only partially composed of the colour FIRE, and therefore this canonical colour is a less salient influence. This produces a testable prediction: words for which colour is a more salient element of their meaning will produce stronger effects of canonical colour. For example, a word like “sky”, for which colour may be the most salient element, might produce stronger canonical colour effects than a word like “lake”, which has other salient connotations (e.g. wetness, depth, cold, etc.)[16].

Finally in Experiment 2, we constructed another linear mixed effects model to describe the canonical pull for high-imageability words. This analysis showed that higher-frequency words tended to be more like their canonical colour (e.g. high-frequency “fire” coloured more like FIRE) than were low-frequency words (e.g. “prong”). This stronger canonical pull for higher-frequency words reinforces the same finding for frequency that we identified in Experiment 1, once we controlled for the baseline distance between control words and canonical colours (i.e. as in Figure 5, above, for “red” in the right-side panel). Together, the strong frequency effects we found in Experiments 1 and 2 support the idea that the influence of the canonical colour increases with repeated exposure. Overall, it seems that the repeated exposure to the automatically evoked canonical colour of high-frequency words (e.g. “red”, “fire”) solidifies the association with the canonical colour, to the point that it can override the influence of grapheme colour. This may be particularly effective for synaesthetes, who as a group have above-average mental imagery ability (e.g., Barnett & Newell, 2008; Chun & Hupé, 2016)[17] and are both familiar with and attentive to the experience of words evoking colours. We infer that lower-frequency words do not have their canonical colour brought to mind as often and are therefore coloured like regular words, based on their grapheme colours. This clearly demonstrates that both the canonical colour association and the frequency of a word

[16] The use of colour associates vs colour terms in understanding the Stroop effect has been a subject of longstanding debate, which is outside the scope of this paper (see e.g. Roelofs, 2003; Schmidt & Cheesman, 2005). However, we do note that theories explaining the “semantic gradient” of Stroop interference (i.e. increasing interference based on the proximity of colour to the meaning of the words, e.g. unrelated “fine”→ associated “fire” → pure colour term “orange”) are primarily based on the timecourse of the interference and delays in response times, which were irrelevant here. Furthermore, the characteristics of these colour associates (such as their colour proximity, or frequency) are seldom systematically or incrementally varied in previous studies.

[17] At least, this appears to be the case among synaesthetes who volunteer for this sort of study; see Simner (2012a, 2012b).

directly influence the synaesthetic colour of that word. Indeed, some Stroop studies that contrasted the frequency of the printed colour terms (e.g. high-frequency "red" vs low-frequency "lavender") reported less interference for the lower-frequency colour terms (Proctor, 1978). If these same influences are pertinent for mental simulation and language processing in general, then we would predict that future studies would find a stronger effect of canonical colour in picture-verification experiments such as Connell (2007), Connell and Lynott (2009) and Mannaert, Dijkstra, and Zwaan (2017) when the target words are higher frequency. This effect may be especially apparent in individuals with high mental imagery. Although mental simulation and mental imagery of colours evoked by words are thought to be different processes (Barsalou, 1999, 2008; Zwaan & Pecher, 2012), it may still be the case that high mental imagery reinforces the colour evoked in the course of automatic colour simulation.

An important consideration for future research on canonical colour influences, both in synaesthetes and in the general population, comes from a difference in findings between our two experiments. In contrast to Experiment 1, we did not find a significant influence of the saturation of the canonical colour in Experiment 2. We suggested that this was because the canonical colours of our high-imageability target words in Experiment 2 were much lower in saturation, as well as higher in lightness, than their control words. Looking at the target words, it is clear that this is due to the specific choice of words (see Appendix C for a full list). That is, the words we chose often refer to objects that are desaturated, especially black or grey (e.g. "rat", "chain", "hare", "iron"), brown ("dirt", "stump", "feast") or very light (e.g. "wax", "bone", "face"). As we are the first to investigate these canonical colour effects by contrasting with control words, it was imperative to match our stimuli as closely as possible on as many measures as we could (e.g. length, spelling, frequency, etc.), and therefore we were limited in our choice of possible word pairs. However, our results here show that future research in this area must take the actual canonical colour of the word into account, as well as (minimally) spelling and word frequency. We particularly note that the diagnostics on our LME model in Experiment 2 indicated problems with model fit – that is, the model did not capture some important influences. This was not the case with the LME model in Experiment 1, signifying that this issue was specific to the data on high-imageability words (e.g. "fire"). We suggest that this missing element may be rooted in the canonical colours of these words.

In our study, we chose highly imageable words, but we suggest now that it is not enough that the word is easy to mentally picture; it must also have a clear, strong colour associated with it. This can be difficult to capture even for words like "fire" that, on the surface, appear to have an obvious colour association. Throughout this paper, we have used the word "fire" as an example of an easily imageable word with a clear colour association. However, the question is: *which* colour is fire? We used seven colours from our non-synaesthete controls to generate the canonical colour FIRE that we then used in our Experiment 2 calculations. These seven colours ranged from yellow to dark red, which, when averaged, gave us the orange colour FIRE we have reported. This was less of an issue with "bone" (all white or pale yellows or pinks), but very striking with "feast" (a mix including white, turquoise, purple, red, and green). This was likely not an issue in Experiment 1 for colour terms because although our control participants may have differed somewhat in the exact shade of red or beige, the colours were clear and highly consistent. For words like "fire" or "feast", however, differences in mental imagery or personal experience may lead to very different colour responses. For further study into canonical colour effects in synaesthesia, it is of critical importance to choose target words by their canonical colour, and ideally with some values held constant (e.g. all highly saturated canonical colours). To the best of our knowledge, there is no standardised measure or database of canonical colours, and with the increasing interest in mentally simulated canonical colour on both synaesthetic associations and embodied cognitive processes in general, such a resource would be invaluable.

Another critical consideration for future work is evaluating how well our task captured the normal synaesthetic experiences of word colours. In other words, it is possible that the very nature of our task, and the terms of "pull" by which we have characterised the interaction of canonical and synaesthetic colour, may suggest conscious decision on the part of our synaesthetes. One could then argue that the systematicity we observed in our results is actually conscious strategy on the part of our synaesthetes. We argue that any test collecting synaesthetic colours necessitates that synaesthetes consciously observe their internal colour experiences and then decide the most apt analogue of that experience given the limitations of the testing apparatus. For example, a synaesthete might experience pink for words beginning with *B*, and therefore a conflict that forces them to consciously decide between pink and canonical BLUE for the word "blue". However, the conscious choice to indicate blue rather than pink was implicitly influenced by the frequency of the word "blue".

That is, choosing between two conflicting colour associations does not mean that our results are artificially strategized by our synaesthetes; it rather highlights how the resolution of that conflict may be influenced by subtle and implicit properties of the words and colours under consideration.

In both experiments, our results showed that there is a complex interaction of canonical colour and linguistic features that together form the synaesthetic colour of a word. That is, the idea that a high-imageability word like “fire” is either the colour of its graphemes (i.e. coloured like “fine”) or like its canonical colour (i.e. like the typical colour of “fire”) is too simple. The frequency of the word, whether it is already similar to its canonical colour, how clear and consistent that canonical colour is, and how saturated the synaesthetic and canonical colours are, all exert a pull on the colour a synaesthete experiences for that word. The primary message of this complexity is that while synaesthetic colouring is by no means mechanical or straightforward, it is based on measurable, systematic perceptual and conceptual factors. That is, each word is uniquely but systematically influenced by its connotations and its imagery as well as its spelling. This supports the idea that synaesthetic colours capture and reflect conceptual and linguistic processes, even implicit characteristics such as frequency, and can therefore be a useful tool for understanding and researching how these processes work.

In conclusion, this study has shown that the canonical colours of both colour terms and high-imageability words have systematic influences on the synaesthetic colours associated with those words. The impact of these canonical colours, stemming from the mental image evoked by the word, was modulated by the frequency of those words. These linguistic, and particularly semantic, influences shed light on the systematic intertwining of language and synaesthesia, and emphasise the need for greater complexity and nuance in the methods used to capture synaesthetic experiences.

Chapter 5

A is for *apple*: The role of letter-word associations in the development of grapheme-colour synaesthesia

Abstract

This study investigates the origins of specific letter-colour associations experienced by people with grapheme-colour synaesthesia. We present novel evidence that frequently observed trends in synaesthesia (e.g. *A* is typically red) can be tied to orthographic associations between letters and words (e.g., "*A* is for apple"), which are typically formed during literacy acquisition. In our experiments, we first tested members of the general population to show that certain words are consistently associated with letters of the alphabet (e.g. *A* is for *apple*), which we named *index words*. Sampling from the same population, we then elicited the typical colour associations of these index words (e.g. apples are red) and used the letter → index word → colour connections to predict which colours and letters would be paired together based on these orthographic-semantic influences. We then looked at direct letter-colour associations (e.g., *A* → red, *B* → blue…) from both synaesthetes and non-synaesthetes. In both populations, we show statistically that the colour predicted by index words matches significantly with the letter-colour mappings: that is, *A* → red because *A* is for *apple* and apples are prototypically red. We therefore conclude that letter-colour associations in both synaesthetes and non-synaesthetes are tied to early-learned letter-word associations.

Introduction

People with synaesthesia experience consistent and automatic quasi-perceptual experiences, such as experiencing taste or colour sensations when they hear words (Simner, 2012b; Ward & Mattingley, 2006). The condition has enjoyed a recent surge of interest since its scientific "rediscovery" in the 1970s and 1980s (Cytowic, 1989; Cytowic & Wood, 1982; Marks, 1975). One idea that has gained traction is that experiences in synaesthesia often reflect intuitive, cross-modal associations common to synaesthetes and non-synaesthetes (Sagiv & Ward, 2006; Spector & Maurer, 2009; Ward et al., 2006). Hence, for both synaesthetes experiencing synaesthesia, and non-synaesthetes making intuitive associations, brighter colours are associated with higher musical pitch (Ward et al., 2006), darker colours with rougher and harder surfaces (Simner & Ludwig, 2012; Ward, Banissy, & Jonas, 2008), and numbers with particular spatial locations (Jonas, Spiller, Jansari, & Ward, 2014). Studying synaesthesia can therefore elucidate universal cross-modal structures and cognitive processes. In the current study we look at similarities between synaesthetes and non-synaesthetes in the way they associate colours with graphemes (letters and numbers). We shall see that such associations are not random for either population, and can be predicted in part by linguistic influences (see also Mankin, 2017; Simner, 2007) and in particular, by early-learned letter-to-word associations (e.g., *A* is for apple).

The current study focuses on grapheme-colour synaesthesia, a common variety of synaesthesia wherein graphemes (here, particularly letters) give rise to automatic associations with colours (e.g. *E* might be leaf green or *D* brown; Baron-Cohen, Burt, Smith-Laittan, Harrison, & Bolton, 1996; Simner, Glover, & Mowat, 2006; Ward, Simner, & Auyeung, 2005). This synaesthesia recruits the cognitive processes involved in reading, which themselves involve a learned association between abstract symbols and sound or meaning. A common thread in large-scale investigations of synaesthesia is that synaesthetes and non-synaesthetes tend to agree on certain colour associations at above chance levels. For example, *A* tends to be red for both populations, *L* tends to be yellow, and so on. The largest studies showing these trends have been conducted in English (Jonas, 2010; Rich et al., 2005; Simner et al., 2005; Witthoft et al., 2015), although similar trends have been found in other languages such as Dutch and Hindi (Rouw et al., 2014), Japanese (Nagai, Yokosawa, & Asano, 2015), German (Emrich, Schneider & Zedler, 2002; Simner et al., 2005) and Ukrainian (Lavrynenko, 2014). To explore how these colour-letter pairing trends are formed, we will first briefly review the previously identified sources of these trends. We will then investigate an as-yet-untested possibility: that the colours for letters may originate from early-acquired letter-to-word associations. We name this proposal, which has been raised previously but never tested, the 'A is for apple' hypothesis: simply put, *A* is red because *A* is for *apple* and apples are red.

We turn first to the colour-letter trends identified in synaesthetes. Three studies (Jonas, 2010; Rich et al., 2005; Simner et al., 2005) asked English-speaking synaesthetes for their letter-colour experiences and identified the colour that occurred at a higher-than-chance level for each letter. A fourth study (Witthoft et al., 2015) reported the most frequent (i.e. *modal*) colour choice for each letter from a large population of synaesthetes. Two of the four studies listed above (namely, Rich et al., 2005; Simner et al., 2005) also gave the same letter-colour association test to non-synaesthetes. Although a question such as "What colour is the letter A?" may seem nonsensical to non-synaesthetes, these participants nonetheless showed agreement not only among themselves, but also with synaesthetes. The sources of some of these widespread associations were more obvious than others. There was a significant tendency for both synaesthetes and non-synaesthetes to associate a letter with the colour that begins with that letter: *R* with red, *Y* with yellow, *G* with green, *B* with blue, *V* with violet, and *P* with pink. Non-synaesthetes also strongly associated *W* with white and *O* with orange, while synaesthetes' associations here were not explicable by colour-name association: *O* with white, but *J* with orange. Furthermore, both groups showed strong shared associations across other letters as well: *A* with red, *D* with brown, *F* with green, *L* with yellow, *U* with grey, *X* and *Z* with black, and *I* with white and/or black. Disregarding for the moment the letter-colour pairs that are easily explicable by the initial letter of the colour name (e.g. red for *R*), how can we explain these trends across synaesthetic and non-synaesthetic populations?

First, there is some evidence that associations can be explicitly acquired from childhood toys or books featuring coloured letters. After a few synaesthetes reported letter-colour associations highly similar to coloured alphabet magnets (Witthoft & Winawer, 2006, 2013), Witthoft et al. (2015) found that in a large sample of 6,588 synaesthetes, 400 (about 6%) had 10 or more letter-colour associations that matched a well-known alphabet magnet set. Furthermore, just one in 150 synaesthetes showed similarities to childhood alphabet books in a study by Rich et al. (2005). A second possibility is that these common letter-colour pairings are indicative of more general associations, not specific to graphemes but to shapes and concepts. Pre-literate children consistently pair *X* with black and *O* with white, but show no inclination towards associating *A* with red and *G* with green, while literate children and adults do both (Spector & Maurer, 2008). Hence, some grapheme-colour pairings may be based in literacy (e.g. *G* → green), while others may be naturally biased shape-colour pairings (e.g. *X* → black). In a follow-up study, pre-literate children also further associated *I* and amoeboid shapes with white, and Z and jagged shapes with black, implying a more general natural bias for spiky or sharp shapes with black, and round or smooth shapes with white (Spector & Maurer, 2011). Finally, Brang, Rouw, Ramachandran, and Coulson (2011) showed that graphemes with similar visual features tended to have more similar colours, so the visual characteristics of graphemes do appear to have some influence on their associated synaesthetic colours.

Another explanation is that these shared associations might come from implicit linguistic, rather than explicit perceptual, characteristics of these colours and graphemes. Two studies found that the saturation and luminance of the colours associated with graphemes by synaesthetes are modulated by how frequently those graphemes appear in the synaesthetes' native language (in German, Beeli, Esslen, & Jäncke, 2007; in English, Smilek, Carriere, Dixon, & Merikle, 2007). Specifically, graphemes that are high in frequency (e.g., *A, S*) have colour associations that are more saturated (i.e., richer in colour; Beeli et al., 2007) and more luminant (i.e. brighter; Smilek et al., 2007). Simner et al. (2005) suggested that these effects may be better explained by also considering the linguistic frequency of the colour term (see also Simner & Ward, 2008): for example, that *A* is red for synaesthetes because *A* is a high frequency letter and *red* is a high frequency colour term, while low-frequency letters like *Q* tend to be paired with lower-frequency colour terms like *purple*. Although Simner et al. (2005) showed that high-frequency letters tend to be paired with high-frequency colour terms, they could not explain why those particular combinations arose: why is high-frequency *A* consistently red but not another high-frequency colour like blue? In other words, what is special about the connection between *A* and red in particular? The current study will attempt to answer this question by proposing that at least some of the letter-colour pairings of synaesthetes and non-synaesthetes are based on word associations acquired during early alphabet acquisition.

Our study will show that for the average person, each letter of the alphabet becomes associated with a particular word or words during alphabet acquisition, particularly through alphabet books (see Nodelman, 2001). These books commonly present a letter of the alphabet with a word beginning with that letter using the phrase "A is for…; B is for…" as a way to encourage children to make the connection between sound, spelling, and words. We will refer to these associated words, which are explicitly linked to the identity of the letter through repeated reading, as *index words*. Here we propose that the prototypical colour of the index word for each letter becomes associated with the letter itself. In short, this can be exemplified as, "*A* is red because *A* is for *apple*, and apples are red." This suggestion has been mentioned by studies investigating common grapheme-colour associations (e.g. Hancock, 2013; Spector & Maurer, 2011) but, to the best of our knowledge, it has never been empirically tested. Here we ask whether this sort of orthographic-semantic medation (*A* → *apple*: orthographic mediation; apples → red: semantic mediation) has any basis in psychological reality when it comes to how letters are internally represented, both for synaesthetes and non-synaesthetes.

In order for this approach to be viable to explain letter-colour commonalities, three connections must be established. First, it must be the case that a letter (e.g. *A*) is consistently associated with a specific word beginning with that letter (e.g., *apple*) across a large proportion of the population. Second, it must also be the case that this index word (*apple*) has a consistently associated prototypical colour for most of the population (e.g., apples are predominantly conceptualised as red). Third, the prototypical colour of the index word must also be the preferred colour for the letter when people are asked to give direct letter-colour associations (e.g., *A* is red). If we find that all three are true, this will support an index-word explanation for grapheme colours. Therefore, the current study will elicit the index words for each letter, the colour of the index word's referent, and the colours directly associated with each letter for both synaesthetes and non-synaesthetes.

Experiment 1: *A* is for apple

Here we will ask whether certain words are consistently associated with each letter of the alphabet (e.g. "*A* is for <u>*apple*</u>"). We will refer to these highly-associated words as *index words*. Experiment 1 prepares for our subsequent investigation into whether index words have prototypical colours (Experiment 2) which influence letter-colour judgements (Experiment 3). As well as identifying index words, this first study will also explore linguistic characteristics of index words, and what determines their selection above other words in the language.

Method

Participants

Our participants comprised 315 non-synaesthete native English speakers from the USA. Participants were 43% female with a mean age of 33.2 years old (SD = 9.9); all were older than 20. We recruited our participants from Amazon's Mechanical Turk (see below). Since this platform is open to workers around the world, we necessarily selected our target sample (N = 315 non-synaesthete native speakers of American English) from a larger population by additionally testing the following participants who were subsequently removed: 71 American English-speakers who self-declared synaesthesia (see *Procedure*), and participants who took the test but who were either non-native speakers (N = 55) or who were non-American English speakers, these being from India (22 total, 9 reporting synaesthesia), unspecified national origin, e.g. "white" (22 total, 5 reporting synaesthesia); and a further 18 from various national backgrounds. This left our final sample of 315 participants.

Materials and procedure

Participants were recruiting using Amazon's Mechanical Turk (hereafter MTurk; www.mturk.com), a self-termed online "marketplace for work" for tasks requiring human intelligence. Workers can preview and complete experiments on the website, and are compensated with a small financial reward once their submissions are approved by the requester. MTurk has been validated as an effective research tool (Bankieris & Simner, 2015; Goodman, Cryder, & Cheema, 2013), and the reward we offered ($.20 per completed test) falls within the typical rate (Buhrmester, Kwang, & Gosling, 2011).

After giving demographic data (gender, age, nationality, and native and additional languages), participants began our test, which consisted of a series of phrases in the format "[Letter] is for…" (e.g. "A is for…"). Each letter was followed by a text box, and letters appeared in alphabetical order. This ordering was intentionally selected to evoke alphabet books and early literacy learning. Participants were given the following instruction: "In the box below each phrase, write the first English word beginning with that letter that you think of. Please answer as quickly and instinctively as possible." Participants then completed each sentence in alphabetical order and this continued to the end of the test ("Z is for…"), where a final question asked participants if they experienced synaesthesia (defined as "lifelong colours for letters or digits").

Results

Data validation

Our dataset comprised a series of words associated to letters, and we first minimally cleaned our data using the following criteria. First, responses that clearly referred to the same concept were combined (e.g. "apple" and "apples" both fell under "apple"). This was not done when the plural morpheme created two different words (e.g. "new" and "news" were not combined) nor when any other affixation gave rise to different concepts (e.g. "killer", "kill", and "killing" were not combined). We corrected spelling mistakes where the intended word was clear (e.g. "giraffee" was combined with "giraffe"). However, ambiguous responses were left as they were, and therefore counted as unique responses (e.g. "ca", which could have been intended as "cat", "car", "can", etc.).

Identifying index words

We next asked whether each letter had a particularly dominant *index word* from among the *response words* given by our participants. To begin, we calculated the agreement for each response word across our participants to measure whether different people gave the same word for each letter. Here, *agreement* indicates the percentage of participants who agreed on a response for any given letter (e.g., over 80% of participants agreed that *A* is for apple). Figure 1 shows the most commonly chosen word for each letter according to this metric.

From Figure 1, it is immediately clear that while some letters have a clear index word out of all response words (i.e. *apple, dog, xylophone,* and *zebra* all have over 50% agreement), other letters do not (e.g. *pie* at 3.2%). We identified the three response words for each letter that had the highest agreement as potential *index words.* A list of these three index words for each letter and their percentage agreement can be found in Appendix D. Our main focus is how these index words may help shape grapheme-colour associations (Experiments 2, 3), but in the following section we briefly explore what psycholinguistic factors underlie English speakers' choice of index words.

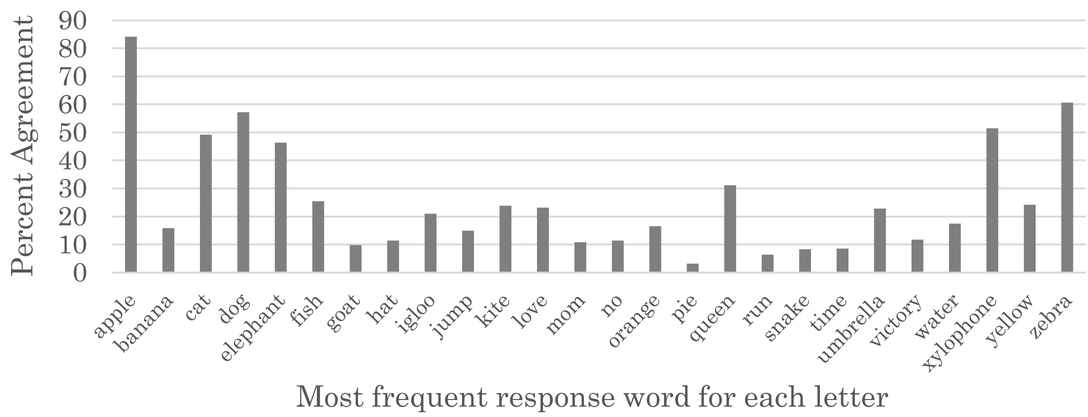


Figure 1. The highest-agreement index word for each letter, listed in alphabetical order with its agreement (as a percentage of all responses for that letter).

Characteristics of response and index words

We examined several possible predictors for how participants chose response words for any given letter. In this section, we first analysed the entire set of response words given by our participants (e.g., *A* is for *apple*, *animal, aardvark,* etc.) to increase our dataset. We considered several factors that might make these response words not only different from other words in the language (e.g., *animal* was a response word but *annex* was not) but which might also distinguish those that were chosen very often from those chosen less often (*apple* was chosen more often than *aardvark*). This analysis may help us understand why particular words might be more likely to contribute their prototypical colour to grapheme-colour trends. To this end, we examined several possible predictors for response word agreement, beginning with word frequency.

Higher-frequency words are reliably elicit quicker responses in behavioural tasks (e.g., Oldfield & Wingfield, 1965) so we first tested whether there was a tendency for response words to be high frequency. For this we entered lexical frequency as a predictor of agreement in response words across our participants in a multiple regression using frequency measures from several different corpora (CELEX: Baayen, Piepenbrock, & van Rijn, 1993; Kučera-Francis: Kučera & Francis, 1967; HAL: Lund & Burgess, 1996; and SUBTLEX-US: Brysbaert & New, 2009). However, none of these measures predicted response word agreement [for all predictors, $t < 1.42$, $p > .155$], meaning that higher-frequency words were no more likely to have higher agreement among our participants than lower-frequency words. However, these frequency measures do not capture the task demands of our experiment. That is, we asked our participants to give a response *within each letter*, whereas the above frequency measures are all calculated from entire corpora across all letters. This means that while *xylophone*, for instance, may be one of the highest-frequency words that begins with *X*, it has a very low frequency in the language as a whole, because the frequency measures used above do not group frequency by spelling.

We therefore developed a new frequency measure: the frequency of each word per million within all occurrences of words *beginning with the same letter*, which we call *by-letter frequency*. To calculate by-letter frequency, we made use of the SUBTLEX-US database of American English film subtitles because this contains the same variety of (American) English used in our experiment, and it has been shown to predict response times to lexical decision and naming tasks better than older, more widely-used corpora (Brysbaert & New, 2009). For each of the *response words* generated in our study (e.g., *apple*, *animal, aardvark,* etc.) we divided its total count in the corpus by the sum of all counts for every word sharing that initial letter. For example, *apple* appears in the corpus 1,207 times, so we divided this raw frequency count by the total count of all words in the corpus beginning with *A*. We then multiplied by a million to produce a by-letter frequency per million (e.g., for *apple*, this was 303.9). As the resulting distribution was highly skewed, we also $\log_{10}$-transformed the values to reach a final log by-letter frequency for all of the words in the corpus.

Having established the *$\log_{10}$-transformed by-letter frequency* for all of the words in the corpus, we then compared the response words generated by our participants to the rest of the words in the corpus to see whether their log by-letter frequency differed. We conducted a 2 (Word-type: Response/Non-response word) x 26 (Letter: *A* – *Z*) ANOVA predicting the log by-letter frequency. Our most important finding was a main effect of response word [$F(1, 74234) = 7842.73$, $p < .001$], indicating that response words have higher frequency [M = 2.62, SD = 1.03] than non-response words [M = 0.76, SD = 0.003; $d = 2.56$]. That is, when people are asked to name the first word that they think of beginning with a particular letter, they tend to choose one of the most common words in English within that letter category (e.g., *A* is for *apple*, not *annex*). There was also a main effect of letter, [$F(25, 74234) = 407.82$, $p < .001$], reflecting the fact that the mean frequencies of the words in each letter group differed between letters. Finally, there was a significant interaction between response word and letter [$F(25, 74234) = 5.73$, $p < .001$]. Bonferroni-corrected post-hoc tests showed that mean log by-letter frequency was significantly higher for response vs non-response words for all letters [*p*s < .001, *α* corrected for 26 comparisons = .002] except the letter *X* [$t(29) = -1.39$, $p = .175$]. This failed to reach significance because of the small number of words in English beginning with *X* combined with a very high frequency of *x* counted on its own as a word [raw by-letter frequency per million = 750714.3, log by-letter frequency = 5.88]. The difference between response vs non-response word frequencies by letter is clearly illustrated in Figure 2.

We also calculated the percentile rank of each word within all words beginning with the same letter to evaluate this further. Out of the 2024 response words, only 56 (2.77%) fell below the median log by-letter frequency, and 257 (12.70%) fell outside the 75th percentile. This underscores that for the vast majority of response words, high frequency is a defining characteristic.

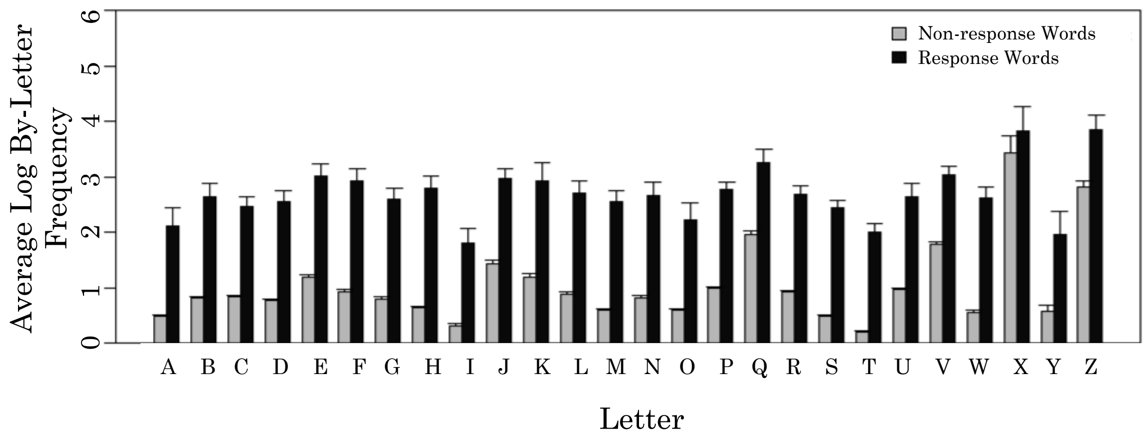


Figure 2. Comparison by letter between average log by-letter frequency of response words (dark bars) and non-response words (light bars). This difference was significant for every letter except *X*.

Having shown that response words tend to be high frequency, we now ask what determines the degree of agreement among respondents – that is, why some response words came up more often than others. The relationship between log by-letter frequency and agreement is illustrated in Figure 3, showing that while the majority of response words were single, unique instances, the higher-agreement words tend to also have higher log by-letter frequency.

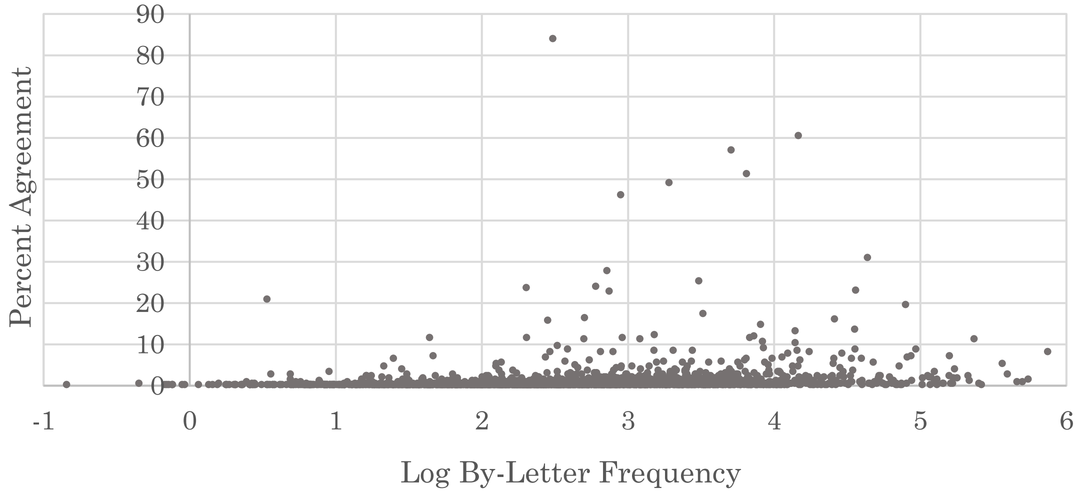


Figure 3. Scatterplot of log by-letter frequency versus percent agreement for all response words.

We repeated our multiple regression for frequency to predict the percentage agreement among participants, but used our new log by-letter frequency as a predictor along with 6 other predictors which we hypothesised may contribute to response word agreement: age of acquisition, imageability, familiarity (Bird et al., 2001; Gilhooly & Logie, 1980; Stadthagen-Gonzale & Davis, 2006), neighbourhood size (i.e., the number of words that differ from the target word by just one letter change, such as *car* and *mat* for *cat*), and behavioural reaction times in lexical decision and naming tasks (English Lexicon Project; Balota et al., 2007). Our final model (see Table 1) shows that only log by-letter frequency and imageability were significant predictors. In summary then, the most widely agreed-upon index words across participants tend to be the most frequent words for that letter, and also the more highly imageable.

Table 1. Summary of the regression model predicting percent response word agreement. $R^2 = .059$, $F(2, 1217) = 38.44$, $p < .001$.

Predictors	Estimate (*B*)	SE (*B*)	*t*	*p*
Intercept	-5.245	0.794	-6.609	< .001
Imageability	0.008	0.001	7.148	< .001
Log by-letter Frequency	0.968	0.146	6.617	< .001

Discussion

This experiment investigated whether letters of the alphabet have associations with words that are shared among language users. In our study, we asked participants to complete phrases of the type “A is for ____”. We classified words that were generated by our participants as response words, and calculated how much agreement there was among participants for each response word. We found that some letters of the alphabet are indeed consistently associated with particular words and have high agreement, and we have termed these index words (e.g. *A* is for *apple*). We further demonstrated that the total set of response words from our participants was higher in frequency than the remaining words in English, but only if frequency is considered within each initial letter. To do this, we created a new frequency measure: a log-transformed frequency per million within each letter based on the SUBTLEX-US corpus (Brysbaert & New, 2009) that more accurately related to the task that we had set our participants of choosing a word beginning with a particular letter (i.e., *xylophone* has a high log by-letter frequency because it is one of the more frequent words beginning with *X*, even though it is low frequency within the language overall). Using this log by-letter frequency, we found that the most widely agreed upon index words are those that are the most frequent by this measure, and are also highly imageable. These types of words are the central feature of alphabet books commonly used in literacy pedagogy with demonstrable success (e.g. Nowak, 2015), and our findings suggest that at least some of these letter-word associations endure into adulthood. We will next ask whether the connection between index words and letters is strong enough to account for direct letter-colour associations in synaesthetes and non-synaesthetes. We do this by now establishing the colours of index words.

Experiment 2: Apples are red

In this experiment, we will focus on the top three highest-agreement response words that we identified in Experiment 1, which we have termed *index words.* We will seek to establish whether the index words for each letter refer to entities that have consistent, prototypical colours (e.g. what is the prototypical colour of an apple?). If these words do indeed have consistent colour associations, we will then be able to compare these colours with the colours associated with letters directly (see Experiment 3, below).

Method

Participants

Our participants comprised 146 English-speaking American non-synaesthetes, 49.3% female ($N = 66$) with a mean age of 37.7 years ($SD = 12.8$ years, range = 19 to 75 years). As in Experiment 1, all participants were recruited using MTurk. These participants had not taken part in Experiment 1. In order to match cultural and linguistic background with the index words gathered in Experiment 1, we excluded participants who were non-native speakers of English and/or were not Americans (*N* = 12). MTurk allows the requester to specify geographic location, so we required that our test would only be available to workers in the United States. (We had not specified this in our first experiment because we were unsure which language or cultural group would dominate our initial sample.) Therefore, non-specific responses to nationality, e.g. "white", "black", were this time included in the analysis, so long as the geographic location was our target location. As before, we screened participants for self-reported grapheme-colour synaesthesia at the end of the test, using the question described in Experiment 1. We also tested but subsequently removed a further 90 participants because they had already taken part in Experiment 1 (*N* = 9) or potentially self-declared synaesthesia (by answering "Yes" *N* = 24, or "Don't Know" *N* = 57 to our synaesthesia question; see *Procedure*). All respondents were compensated $0.40 for their participation. This left our final sample of 146 participants.

Materials

The materials from this study were a subset of the words generated in Experiment 1, which had been elicited in that study using phrases such as "A is for ____; B is for ____...". In the current study we selected only the top three highest-agreement response words as index words to be tested here. This resulted in a final list of 78 words (3 words x 26 letters); these items are listed in full with their percentage agreement from Experiment 1 in Appendix D.

Procedure

We created our study using Qualtrics survey software (www.qualtrics.com) and posted its URL on MTurk. After a brief introduction and the collection of basic demographic information (age, gender, nationality, native and other languages spoken), the 78 target words were presented in a unique random order for each participant. For each word, participants were instructed to form a mental image and then provide the "strongest, most dominant colour." Colours were selected from a drop-down list of basic colour terms (black, white, red, orange, yellow, green, blue, purple, pink, brown, gray). Participants were also asked to provide a confidence rating for how sure they were that the colour they had chosen was the best colour for each item, on a Likert scale from 1 (not sure at all) to 7 (very sure). The test required participants to give both a colour and a confidence score for every item before it would allow them to advance.

Results

Response validation

We conducted an initial check that participants had completed the test to a sufficient standard and in good faith. We identified and removed 12 participants who were responding randomly or repeating the same colour-choice throughout on the following basis. First, we selected four items we believed should have a unique colour association*: banana, elephant, orange*, and *yellow* (coloured: yellow, grey, orange, and yellow, respectively). We then asked three independent raters to confirm our intuition, which they did in 100% in agreement, as did over 92% of our participants. We therefore removed any participants who differed from these independently-established responses for two (50%) or more of these standardised items. Next, we identified participants who had chosen the same colour repeatedly regardless of the item (e.g. green for most words). We calculated the mean number of times that each colour was chosen for each word across all participants, and established a first cutoff at 2.5 standard deviations above the mean, and a second cutoff at 3 standard deviations above the mean. For example, *orange* was selected an average of 4.5 times (SD = 3.0) by each participant, with a first cutoff of 12.0 and a second cutoff of 13.5. We removed any participants who selected more than two colours above the first cutoff (e.g. who selected orange 12 times or more), or one colour above the second cutoff (e.g. who selected orange 14 times or more). Using both these criteria, we identified 12 problematic participants and excluded them from the analysis. This left a final pool of 134 participants.

After removing inattentive participants, we next validated our dependent measure, which was the frequency with which any given colour was selected for our target items. For example, for *apple*, the most commonly selected colour was red, which was selected 122 times out of 134 responses, which gave a maximum colour frequency of 91.04%. As described above, participants also gave a confidence rating, which we used to validate their colour choice. To do this, we compared the mean confidence rating for each word to its max colour frequency, using a Spearman nonparametric correlation as our data were not normally distributed. This correlation showed that when participants were more consistent in their colour selection (i.e. items had higher max colour agreement), they were also more confident in their colour choice [Spearman's ρ (78) = .82, $p < .001$], thereby validating our dependent measure.

Do index words predict letter-colour trends?

We can now use our results from Experiments 1 and 2 (letter →index word → colour) to predict which colours would be most often associated with each letter if index words do indeed influence direct letter-colour pairings. To do this, we took into account how often each index word was chosen for its letter and how often a particular colour was chosen for each index word. We will illustrate our procedure using *A* and *apple*. We first calculated the percentage of each colour for each of the three index words, so for *apple* the responses were 91.04% red and 8.96% green, and so on for every colour. We then multiplied this result by the percentage of times that particular index word was selected for that letter. So, as *apple* accounted for 84.13% of all of the responses for *A*, this means that red accounts for 76.59% (.8413 x 91.04%) of the colour selections for *A* via *apple*. On the other hand, since *animal* was only given as a response for *A* 1.90% of the time and brown was selected for *animal* 58.21% of the time, the brown responses for *animal* only count as 1.11% (.190 x 58.21%) of the total proportional colour responses for *A*. We applied this procedure to all combinations of index words and colours, then we summed the resultant weighted colour responses for all three words for each letter – all the red responses, all the brown responses, etc. This gave a colour score for each colour within each letter. The highest colour score for each letter indicates the most *dominant colour* for that letter, and is therefore the colour we would predict if letter colour is mediated by index words. These predictions are detailed in full in Appendix E.

This colour score predicts the most likely dominant colour for each letter, while still taking into account the amount of index word agreement. That is, the colour scores for *A* are 76.61% red, 7.60% green, etc., but the sum of the eleven colour scores for *A* is less than 100%. This is because each score reflects the percentage agreement for index words for each letter. In the case of *A*, for example, the three index words *apple, animal* and *aardvark* accounted for 87.08% of all responses for *A*, so the eleven different colour scores for *A* sum to 87.08%. The remaining response words for *A* (e.g. *automobile, ant,* etc.) account for the remaining 13.02% of all responses. Since we only collected colours for the top three index words, the colours for non-index response words like *automobile* or *ant* are not represented in the colour score, so the colour distribution of the remaining 13.02% for *A* is unknown. In contrast to *A*, the three index words for *S*, which are *snake*, *sister,* and *stop,* together only accounted for 17.8% percent of the response words for *S*, so the remaining 82.22% of the colour score variation for *S* is unknown and the highest colour score for *S* is very low (dominant colour: green, 3.93%). This distribution is illustrated in Figure 4 below for both *A* and *S*.

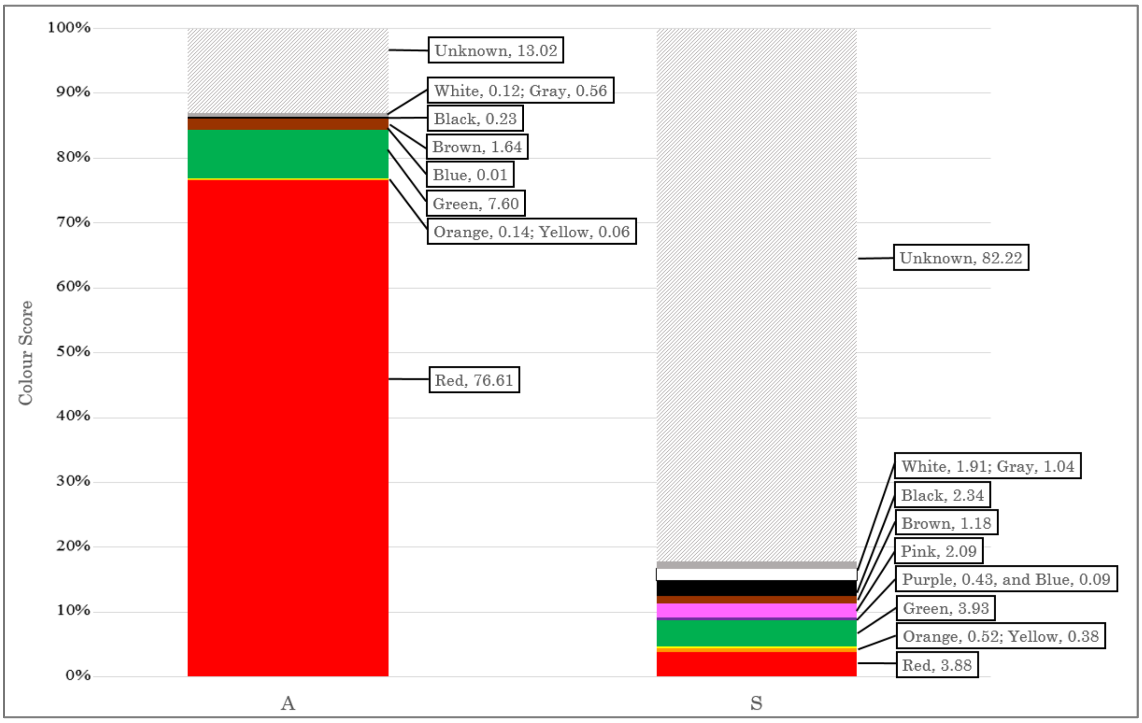


Figure 4. The weighted proportions of colour choices for A (left column) and S. In both cases, the diagonally barred "unknown" proportion shows how much of the overall colour distribution for each letter is accounted for by response words not included in our word → colour selection task (since we included only the top three index words for each letter). For *A* this is only 13.02%, but for *S* this is 82.22%. Each section is labelled with both the colour name and its colour score, in percent.

Comparison to previously reported letter-colour trends

Next, we evaluated whether the dominant colour that we predicted for each letter via index words, as calculated above, successfully matched with previously reported letter-colour associations in the literature. To do this, we return to the colour associations presented in the introduction, where we explored the grapheme colouring trends in both synaesthetes and non-synaesthetes. Table 2 compares the predictions from the current study with the previously published colour-letter associations across the studies reviewed in the literature.

The colour-score-predicted dominant colour matched with previously observed trends for nine letters, or 35% of the alphabet: *X, Z, I, A, O, Y, W, Q,* and *D*. However, two of these letters, *Q* and *W,* were associated by synaesthetes and not by non-synaesthetes, so we will focus here on the seven letters (*X, Z, I, A, O, Y,* and *D*, 27% of the alphabet) that matched between previous associations reported for non-synaesthetes and predicted dominant colours from our own non-synaesthete participants. The cumulative probability of obtaining seven or more matches out of 26 by chance, given an equal 1/11 = .09 probability of each of the eleven colour terms being selected for any given letter, is approximately one in 147.8, or p = .007. This is a promising result, but in order to quantify it further, we will gather our own direct letter-colour associations *from the same population* in Experiment 3, below, to match our results by linguistic and cultural background.

Table 2. An abbreviated summary of the previously reported trends in grapheme-colour associations (first three columns) for comparison with the current study (far right column) predicting associations via index words. The first column reports the letters that show associations across all three "significance studies", i.e. Simner et al. (2005), Rich et al. (2005), and Jonas (2010), who reported the pairings that were statistically significant. The second column reports only the most commonly chosen letter-colour pair (Witthoft et al., 2015). Colour-letter pairings from any of the first three columns that match with the current study are highlighted in **bold**.

Colour	Synaesthetes, significance studies	Synaesthetes, Witthoft et al.	Non-synaesthetes, all studies	Current Study
Black	**IXZ**	**XZ**	**XZ**	CHPU**XZ**
White	**I**O	**I**O	**I**W	G**I**
Red	**A**R	**A**R	**A**R	**A**JL
Orange	J	HJKN	**O**	**O**T
Yellow	**Y**	CLS**Y**	L**Y**	B**Y**
Green	-	EFG	FG	S
Blue	B	BDT**W**	BI	**W**
Purple	V	P**Q**V	PVM	**Q**
Pink	P	-	P	-
Brown	**D**	-	**D**HT	**D**KMNV
Grey	X	-	UX	EFR

Discussion

In the current experiment, we asked participants to tell us the prototypical colours that they thought of for index words identified in Experiment 1 (e.g. "apples are red"). We then calculated a colour score, combining the agreement within each letter from Experiment 1 with these colour associations, that allowed us to make predictions about which colours would be most dominantly paired with which letters, and how strong that association would be. By comparing our predictions with previously published results from studies of synaesthetes and non-synaesthetes, we showed that our predictions were able to account for nine letter-colour associations (*X, Z, A, I, O, Y, W, Q,* and *D*) that were not previously directly explicable. The current experiment provides a new explanation for previously reported trends in grapheme-colour pairs. We suggest that during the course of literacy acquisition, alphabet books and other classroom materials pair index words with letters, such as *apple* and *A*, and this orthographic association is internalised as an association between *A* and red via the prototypical semantic colour for *apple*. We further suggest that the same effect can explain *Q* with purple (via *queen*), *D* with brown (via *dog*), *W* with blue (via *water*), and *X* and *Z* with black (via *x-ray* and *zebra*).

We will now test the predictions of our dominant colours directly using American English-speaking participants. Thus far, all of the grapheme-colour trend studies have used populations and methods different from the current study. In order to match as closely as possible for the influences of cultural and sociolinguistic background, we will collect our own direct letter-colour pairings for further analysis from the same population.

Experiment 3: *A* is red

This third experiment will directly gather letter-colour associations from a similar population that provided index words (Experiment 1) and the colours of those index words (Experiment 2). As described in the introduction, some studies have already sought to establish colour trends in non-synaesthetes but used populations with cultural and linguistic backgrounds different from the current study. As the previous trends come from British (Simner et al., 2005; Jonas, 2010), Australian (Rich et al., 2005), and mixed nationality (Witthoft et al., 2015) participants, we will obtain our own letter-colour associations from the same population as the previous experiments (i.e. English-speaking American MTurk workers), which allows us to control for location and language. We still expect to find some of the same general patterns of letter-colour associations as have been previously reported. More importantly, we will be able to directly and numerically compare these letter-colour choices with the dominant-colour predictions made by our colour score from Experiment 2.

Method

Participants

We tested a final sample of 175 American English-speaking non-synaesthetes, 56% female (N = 98) with an average age of 33.30 (SD = 10.10 years, range = 20 – 82 years). Following the same procedure described in Experiments 1 and 2, additional participants were removed from the analysis if they declared a nationality or native language other than American and English. We also only used responses from participants who answered “No” to our screening question for synaesthesia, which asked, “Do you experience synesthesia? In other words, were the colors you gave in this task associations that you've known about all your life?” This question was different from Experiments 1 and 2 because of the difference between the tasks – those experiments asked for the prototypical colour of real-world objects, rather than synaesthetic colours. On the basis of nationality, language background, and/or self-reported synaesthesia, 110 participants were removed. Finally, five participants were excluded because they had already participated in a previous experiment. This gave us a final pool of 175 participants.

Materials and procedure

The recruitment, instructions, and survey apparatus were similar to that reported in Experiment 2 with the following adjustments. First, the list of words used in Experiment 2 were switched out for the 26 letters of the English alphabet. Participants were instructed to choose the colour for each letter that “seems to fit the letter best.” We also removed the confidence rating task. All participants provided colours freely for all letters (i.e. they could choose the same colour as many times as they liked) but were required to provide a colour for every letter. The order of the letters was randomised for each participant, and the order of colour options was also randomised for each letter. This was especially important as Simner et al. (2005) showed that non-synaesthetes tend to associate colours with letters in the order that the colours are easiest to generate.

Results

Data validation

We conducted a basic data validation procedure similar to that outlined in Experiment 2. For each colour, we calculated the distribution of colours for each participant, and the overall mean number of selections for each colour. Thirteen participants were excluded because they had more than one colour selected above 2.5 standard deviations from the mean, and/or they had chosen the same colour for more than 25% (in this case, 7 or more) of the letters. This resulted in a final pool of 162 participants.

Letter-colour associations

In this analysis, we found the most frequently selected (i.e. highest-agreement) colour for each letter by calculating the proportions of colour selections for each letter.

Table 3 below compares this *modal colour* to the letter-colour data from previous studies of letter-colour associations.

Table 3. An abbreviated summary of the previously reported trends in grapheme-colour associations for comparison with the current study directly associating letters and colours. Letter-colour pairings from any of the first three columns that match with the current study are highlighted in **bold**; letters that appear more than once were tied for modal colour.

	Synaesthetes			Non-synaesthetes	
Colour	Significance studies	Witthoft et al.	Current Study	All previous studies	Current Study
Black	I**XZ**	**XZ**	NSU**XZ**	**XZ**	**XZ**
White	**I**O	**I**O	**I**W	I**W**	H**W**
Red	**AR**	**AR**	**AR**	**AR**	**AFKR**
Orange	J	HJK**N**	**N**O	**O**	C**O**
Yellow	**Y**	C**L**S**Y**	E**LY**	**LY**	H**LY**
Green	-	**EFG**	**EFG**JN	FG	**G**S
Blue	**B**	**B**DTW	**BCT**	**BI**	**BI**JTU
Purple	**V**	P**QV**	**QV**	**PV**M	J**PV**
Pink	**P**	-	K**P**	P	Q
Brown	**D**	-	**D**HM	**D**HT	**D**MN
Grey	X	-	-	UX	E

To further establish the connections between letters and colours, we next conducted an analysis of *statistically significant* letter-colour pairings, following Simner et al. (2005). We did this to distinguish between colours that are simply selected often overall as opposed to associations that occur at a frequency significantly beyond chance. In this analysis, we first counted the total number of times each colour was selected for all participants and all words, from which we calculated the baseline probability of each colour being selected. We then used a binomial distribution to calculate the probability of each letter-colour pairing occurring at above chance levels. These *significant colours* are described in Table 4, below.

The results of the binomial analysis, first, confirm the modal (i.e. most selected) colour as statistically significant for every letter. However, a comparison between modal and significant colours reveals how the degree of agreement can vary for each letter. For example, the modal colour for *S* was green, but red was also significantly associated with *S*, with only 1.23% of agreement separating them. Eleven letters differed between their modal colour and second most frequently selected colour by less than 5%, and three (*H*, *J*, and *K*) had exact ties for their modal colour. For this reason, we will include both first and second most selected colours for each letter for the purposes of evaluating the predictions of the index word colour score, and refer to these as *first* and *second modal* colours. In this way, we can distinguish the colours that are selected most often (i.e. *modal* colours) from those that are associated at a statistically significant level (i.e. *significant* colours).

Table 4. For each colour term, the letters are listed which were significantly associated with that colour at least at $p < .05$ level in our non-synaesthete participants' responses. Letters appear more than once if they were significantly associated with more than one colour. For each associated colour, letters are listed in separate columns if they were significant at $p < .01$ (middle column) or $p < .001$ (right column).

Colour	$p < .05$	$p < .01$	$p < .001$
Black	B D	-	X Z
White	E F T	H I	W Z
Red	F S	-	A R
Orange	J N	-	C O
Yellow	H	-	L Y
Green	C D E F	S T	G
Blue	A I J	U	B T
Purple	I Q	J	P V
Pink	I L	K Q	P
Brown	H T	B K M	D N U
Grey	D U	G Q X	E

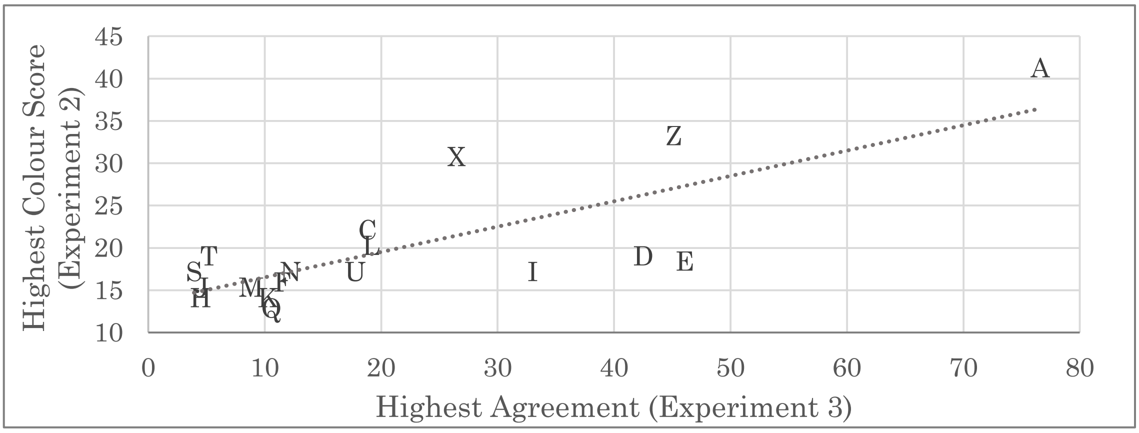
Index word predictions vs letter-colour associations

We can now directly compare the letter-colour combinations predicted by the colour score from Experiment 2 with the actual letter-colour associations collected in the current experiment. We will address this in two ways: first, by examining the relationship between levels of agreement; and second, by comparing the colours themselves.

To begin, we can evaluate the relationship between the colour score that we calculated in Experiment 2 and the max colour agreement described above with a correlation. First, we removed the letters from the analysis for which the colour association could be explained another way – namely, the initial letters of colour terms (i.e. *B, W, R, O, Y, G, V, P*). As the introduction describes, the letter-colour agreement for these letters is very high because of their connection to colour terms (e.g. *R* for red); indeed, the binomial analysis above showed that all initial letters of colour terms were significantly associated with their colour at $p < .001$, which may mask an effect for non-colour-term letters. After removing these colour-term letters, the remaining 18 letters showed a significant correlation between colour score (from

Experiment 2) and colour agreement (from the current experiment; Spearman's ρ = 0.70, p = .001). In other words, the letters that had higher agreement in the colour obtained via their index words also had higher agreement for their directly chosen colour. The plot in Figure 5 illustrates this correlation.

Figure 5. Scatterplot showing the correlation between the highest colour score



obtained via index words in Experiment 2 and the highest agreement reported for direct colour association by non-synaesthete participants. Each of the 18 letters included in this correlation represents its respective point.

We will next explore *how* index words predict colours for letters. Figure 6 shows the first and second dominant colours predicted by the colour score from Experiment 2 and the first and second modal colour associations from Experiment 3. Altogether, there was a match between at least one of the dominant colours and one of the modal colours for 17 out of 26 letters (65%), and 22 out of a possible 52 combinations of first/second dominant colour and first/second modal colour, a hit rate of 38.5%. In order to understand how likely this pattern was to emerge by chance, we used a Monte Carlo simulation. For each iteration of the simulation, we randomly generated two pairs of colours, sample 1 and sample 2, 26 times, representing the top two colours that our participants selected for each letter via index words (dominant colours from Experiment 2) and directly (modal colours from the current experiment). Both sample 1 and sample 2 were composed of mutually exclusive colours – that is, each randomly generated pair of colours had to consist of two different colours, not the same colour twice. We then counted the number of times there was a match between the colours in sample 1 and in sample 2, reflecting the same colour-matching process between the dominant colours from Experiment 2 and the modal colours from the current experiment that we conducted on our data from our participants (see Figure 6 for an illustration). We then repeated this process of generating two random pairs of colours for each of 26 letters and counting the number of matches for one million (1000000) iterations, which gave us a simulation of colour matches at chance level.

Letter	Dominant Colour (Experiment 2)	Modal Colour (Experiment 3)	2nd Dominant Colour (Experiment 2)	2nd Modal Colour (Experiment 3)
A	Red	Red	Green	Blue
B	Yellow	Blue	Brown	Black
C	Black	Orange	Orange	Green
D	Brown	Brown	Black	Blue
E	Gray	Gray	White	Green
F	Gray	Red	Orange	Green
G	White	Green	Pink	Gray
H	Black	White/Yellow	Brown	White/Yellow
I	White	Blue	Blue	White
J	Red	Blue/Purple	White	Blue/Purple
K	Brown	Red	Red	Pink/Brown
L	Red	Yellow	Yellow	Green/Purple/Pink
M	Brown	Brown	White	Purple
N	Black	Brown	Red	Yellow
O	Orange	Orange	Gray	Black/White
P	Black	Purple	Brown	Pink
Q	Purple	Pink	White	Gray/Purple
R	Gray	Red	White	Purple
S	Green	Green	Red	Red
T	Orange	Blue	Green	Green
U	Black	Blue	Blue	Brown
V	Brown	Purple	Red	Black
W	Blue	White	White	Blue
X	Black	Black	Gray	Red
Y	Yellow	Yellow	Green	Gray
Z	Black	Black	White	White

Figure 6. Comparison between dominant colour predictions via colour score from Experiment 2 and modal letter-colour agreement from Experiment 3 for non-synaesthete participants. Coloured cells indicate a match between dominant and modal colour(s) for that colour. Matches for white have been coloured dark grey for visibility. Cells with more than one colour indicate a tie in agreement.

We then used this simulation to estimate of the probability of obtaining the patterns we observed. As noted above, we found at least one match for 17 out of 26 letters in our data. Our simulation found that 1,352 iterations out of 1000000 resulted in 17 or more letters with at least one colour match. Dividing this result by 1000000 gives a decimal probability of having matches for 17 or more letters by chance, comparable to a p-value, which in this case was $p = .001$. This by-letter analysis counted both a single match (e.g. sample 1: red, green; sample 2: blue, red) and a double match (e.g. sample 1: red, *blue*; sample 2: *blue*, red) as a single hit for a given letter. However, we also wanted to know the probability of obtaining the total number of matches we observed, so that a single match and a double match would count as two hits.

Counting double matches as two hits rather than one means that there were 52 total possible hits for each iteration, since there were two hits possible for each of 26 letters. The Monte Carlo simulation, again iterated 1000000 times, indicated that the probability of observing 22 out of 52 possible hits, as we did in our data, was .000017, or $p < .001$. Given this extreme unlikelihood that choosing colours by chance would lead to the pattern of matches we describe, we rather suggest that the colours for letters of the alphabet are systematically derived from index words.

Discussion

This experiment collected colour associations for letters from a large group of English-speaking American non-synaesthetes; previous studies have examined British (Simner et al., 2005), Australian (Rich et al., 2005), and mixed nationality (Witthoft et al., 2015) English speakers. We first calculated the top two *modal*, or most frequently selected, colours for each letter, and used a binomial analysis to show which colours were also *significantly* associated with each letter. We also compared our modal colours to other studies and found that our participants chose colours for letters similar to those previously reported. We then compared the *dominant* colours for each letter that we had predicted using index words (from Experiments 1 and 2) with the *modal* colours, and showed that index-word-based dominant colour matched with directly associated modal colours significantly beyond what chance would predict.

First, our data provide a strong initial indication that the index words associated with letters (*A* is for apple) may indeed influence letter-colour pairings through orthographic-semantic associations. We also saw that despite sharing a common language with the English speakers in previous studies, for our participants some letters differed in their modal colour associations from previously reported trends (e.g. *M* with brown, *E* with grey) that nonetheless matched the dominant colour predicted by index words. We will explore the implications of this in depth in the general discussion, but first we will examine whether the index word route described above is also a meaningful predictor of letter-colour associations for self-reported synaesthetes as for non-synaesthetes.

Experiment 4: Synaesthetes

To further evaluate the influence of the index word route, we turn now to the trends reported for grapheme-colour synaesthetes. It may be that while non-synaesthetes rely on overlearned index words to form associations in a task that is not particularly meaningful for them ("What colour is the letter *K*?"), synaesthetes might rather rely on other implicit or systematic processes to determine colour associations. Support for this idea comes from Simner et al. (2005), who found that while synaesthetes and non-synaesthetes shared some implicit 'rules' in associating letters with colours, they also showed certain differences as well. Therefore, index words may predict different letter-colour pairs for synaesthetes. To test this, we will repeat the index word → colour → letter analysis above, this time using the responses from participants who self-reported experiencing grapheme-colour synaesthesia.

Method

Participants

A total of 88 self-reported synaesthete participants took part across the two colour-gathering experiments. In Experiment 2 (index words → colours) we identified 24 self-reported synaesthetes as described below, 14 female (58%) with a mean age of 35.58 (SD = 12.42). In Experiment 3 (letters → colours) there were 64 self-reported synaesthetes, 37 female (60.94%), with a mean age of 35.56 (SD = 11.06 years). All participants were American English speakers.

In both experiments, we used a self-report question to evaluate synaesthesia. For Experiments 1 and 2, this question was, "Do you experience synesthesia (lifelong colors for words, letters, or digits)?" and in Experiment 3, "Do you experience synesthesia? In other words, were the colors you gave in this task associations that you've known about all your life?" As explained in Experiment 3, above, this question was different between the experiments because Experiments 1 and 2 did not specifically ask for synaesthesia-like associations, whereas Experiment 3 did. For this present analysis, we included participants who answered "Yes" to these questions for either question. We acknowledge that this self-report measure is far less stringent than the widely accepted objective validation method of testing synaesthetes' associations repeatedly over time (Baron-Cohen et al., 1993; Eagleman et al., 2007). For this reason, the results reported below should be considered an initial investigation pending further research.

Materials

All materials were the same as in Experiments 2 and 3. We used the same list of index words collected from non-synaesthetes in Experiment 1, and the testing apparatus was identical to Experiments 2 and 3[18].

Procedure

All participants were recruited and tested using MTurk and Qualtrics as reported in Experiment 2 and 3, above. There was no explicit indication in either task that synaesthesia was of interest until the last question, the self-report of synaesthesia. Therefore, self-reported synaesthetes and non-synaesthetes were tested together, and only separated into groups using their response to the synaesthesia question after they had completed the experiments.

Results

Self-reported prevalence of synaesthesia

The self-report of synaesthesia questions allowed us, first, to take an informal measure of the prevalence of self-reported grapheme-colour synaesthesia in a random sample of American English speakers. For both experiments, we compared the proportion of self-reported synaesthetes (i.e. "Yes" answers) to non-synaesthetes (i.e. "No" answers) within English-speaking Americans; participants responding "Don't Know" (Experiment 2 only) were excluded. For Experiment 2, there was a synaesthesia prevalence of 16.44% (24/170 total), and for Experiment 3 a prevalence of 27.20% (65/239 total). Combining both tasks, the overall prevalence of self-reported synaesthesia was 21.76% (89/409 total). This is an unexpectedly high proportion, more than four times the 4.9% prevalence of self-reported grapheme-colour synaesthesia found by Carmichael, Down, Shillcock, Eagleman, and Simner (2015), and well beyond their estimate of 1-2% of grapheme-colour synaesthesia in

---

[18] We decided to use the same index words collected from non-synaesthetes in Experiment 1 because we wanted to compare to the results for non-synaesthetes directly, as well as avoid any effects of bidirectionality (cf Weiss, Kalckert, & Fink, 2009). That is, if synaesthetes experience red for *A*, that automatic experience may make them more likely to choose a red item as an index word for *A*. Although this is also a possibility in assigning colours to index words, we believe this is less likely to pose a problem. First, the instructions explicitly asked participants to form a mental image and report the most dominant colour in that image (see Experiment 2, above, for details). This requires participants to focus their attention on the concept, not the word itself, and would likely lessen any impact of word-based synaesthetic colour on their index word-colour choice (Mattingley, 2009). As synaesthetes have been shown to have enhanced mental imagery (e.g. Barnett & Newell, 2008; Price, 2009), we believe they would be particularly *good* at this task, rather than influenced unduly by synaesthetic colours. Finally, we excluded any participants who consistently reported unexpected or unusual colours in this task (see *Data validation*), so we are confident that our self-reported synaesthete participants completed this task just as well as non-synaesthetes.

the general population. However, we note an important caveat with this measure beyond self-report. MTurk requires a description of the task before workers decide to accept. For Experiment 3, this was "You will provide color associations for a list of letters," which might have attracted people with synaesthetic experiences to take part. More informatively, we can compare the gender ratio of those who report synaesthesia to those who did not. Combining both experiments, the synaesthetes were 57.3% female, whereas the non-synaesthetes were 51.1% female. A chi-square test showed that there was no significant difference in gender between synaesthetes and non-synaesthetes [χ^2 (1) = 1.31, p = .253]. This result agrees with recent work indicating that there is no gender bias in synaesthesia (Simner & Carmichael, 2015). We will return to our recruitment method in the general discussion.

Data validation

Using the same data validation procedure as outlined in Experiment 2 and 3, we excluded participants who had not completed the task as instructed. This led to the exclusion of five participants from Experiment 2, for a final sample of 19; and five participants in Experiment 3, for a final sample of 59 and a final grand total of 78.

Do index words predict letter colours for self-reported synaesthetes?

Using the same index word agreement percentages from Experiment 1 and the colours from the first group of 19 self-reported synaesthetes, we calculated a colour score to predict the colour for each letter using the index word route. As described in detail in Experiment 2, above, this score represents the colour distribution for each letter that is accounted for by the colours of its top three index words, with the highest score indicating the most dominant colour for each letter. Appendix E details the highest colour score and associated dominant colour for each word.

The results demonstrate that self-reported synaesthetes and non-synaesthetes performed this task very similarly. This is likely due to the fact that index-word → colour associations are based in real prototypical colour (e.g. *dogs* are brown). However, predicted colour did differ for six letters: *G*, *H*, *J*, *K*, *S*, and *T*. Next, we calculated the highest agreement modal colour as well as significantly associated colours for each letter (see Experiment 2, above, for details). These results are compared to those in previous studies for synaesthetes in Table 5, below.

Table 5. Comparison of the results of letter-colour associations between the current study (far right columns) and previous studies of trends in letter-colour pairings. For

comparison, we also include modal colours for non-synaesthetes from Experiment 3 of the current study. Matches between our self-reported synaesthetes and previous studies are highlighted in **bold**.

Colour	*Synaesthetes*			*Non-synaesthetes*	
	Significance studies	Witthoft et al.	Current Study	All previous studies	Current Study
Black	I**XZ**	**XZ**	NSU**XZ**	**XZ**	**XZ**
White	IO	IO	**I**W	**I**W	HW
Red	**AR**	**AR**	**AR**	**AR**	**A**FK**R**
Orange	J	HJK**N**	**NO**	**O**	CO
Yellow	**Y**	C**L**S**Y**	E**LY**	**LY**	H**LY**
Green	-	**EFG**	**EFG**JN	F**G**	**G**S
Blue	**B**	**B**D**T**W	**B**C**T**	**B**I	**B**IJ**T**U
Purple	**V**	**PQV**	**QV**	**PV**M	J**PV**
Pink	P	-	K**P**	P	Q
Brown	**D**	-	**D**HM	**D**HT	**D**MN
Grey	X	-	-	UX	E

We also calculated the letter-colour pairs that had a significant binomial distribution. The results are summarised as before in Table 6, below.

Table 6. For each colour term, the letters are listed which were significantly associated with that colour at least at p < .05 level in our self-declared synaesthete participants' responses. Letters appear more than once if they were significantly associated with more than one colour. For each associated colour, letters are listed in separate columns if they were significant at p < .01 (middle column) or p < .001 (right column).

Colour	$p < .05$	$p < .01$	$p < .001$
Black	D S U X	Z	-
White	H	I	W
Red	-	-	A R
Orange	-	-	O
Yellow	E	L	Y
Green	E	-	G
Blue	D I	T	B
Purple	Q Z	-	P V
Pink	-	I K	P
Brown	B T	D H M	-
Grey	J	G	-

As with non-synaesthetes, the correlation between colour score and colour agreement was significant for the 18 letters that are not the first letters of colour terms (Spearman's ρ = .533, p = .023). This correlation is illustrated in Figure 7, below.

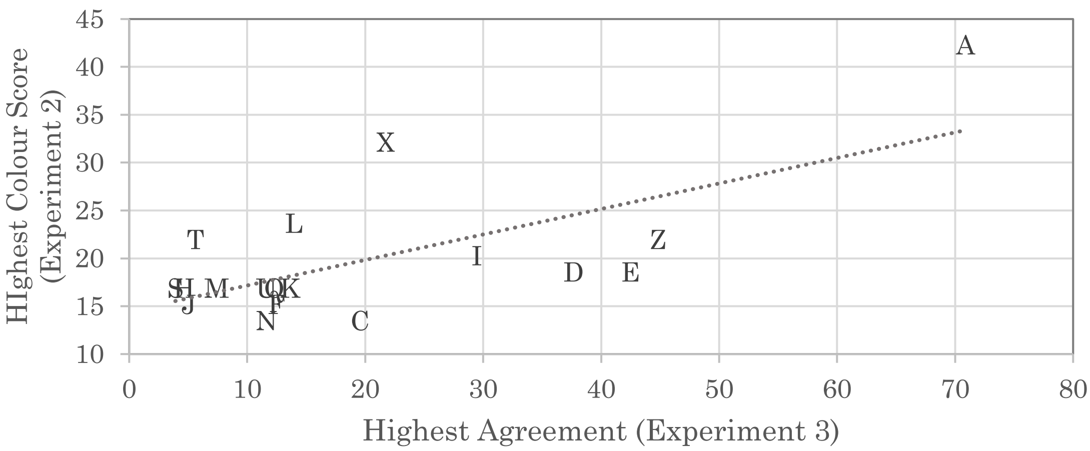


Figure 7. Scatterplot showing the correlation between the highest colour score obtained via index words and the highest agreement reported for direct colour association for self-reported synaesthete participants. Each of the 18 letters included in this correlation represents its respective point.

As summarised in Figure 8, below, index words successfully predicted letter colour associations for 20 letters (76.9% of all letters), with 21 matches overall between colour score and agreement out of a possible 52 (maximum two matches per letter); this is a hit rate of 40.38%. We evaluated the probability of obtaining this pattern of results as in Experiment 3, using a simulation of matching two pairs of randomly generated colours for 26 letters, iterated 1000000 times. We again conducted two of these Monte Carlo simulations. The first counted the number of iterations in which there was at least one match for 20 out of 26 letters as 8/1000000, or $p < .001$. The second counted the number of iterations in which there were overall 21 matches out of 52 possible matches across the entire alphabet and found 62/1000000, or $p < .001$. This indicates, as with synaesthetes, that the matches between index-word-predicted dominant colours and directly-associated modal colours are very unlikely to be coincidental. Rather, these consistent patterns suggest that letter-colour associations are influenced by semantic colours transferred to the letters via the index words associated with those letters during literacy acquisition.

Letter	Dominant Colour (Experiment 2)	Modal Colour (Experiment 3)	2nd Dominant Colour (Experiment 2)	2nd Modal Colour (Experiment 3)
A	Red	Red	Green	Black/Blue
B	Yellow	Blue	Blue	Brown
C	Black	Blue	Orange	Grey/Orange/ Yellow/Green
D	Brown	Brown	Black	Blue
E	Grey	Yellow/Green	Yellow	Yellow/Green
F	Grey	Green	Orange	Blue
G	Pink	Green	Grey	Grey
H	Brown	Brown	Black	White
I	White	White	Blue	Blue/Brown
J	Green	Green	Red	Grey/Orange/Blue
K	Red	Pink	Brown	Black
L	Red	Yellow	Yellow	Blue
M	Brown	Brown	White	Blue
N	Black	Black/Orange/Green	Red	Black/Orange/Green
O	Orange	Orange	White	Black/Yellow
P	Black/Green	Pink	Black/Green	Purple
Q	Purple	Purple	White	Orange
R	Grey	Red	White	Pink
S	Red	Black	Green	Red
T	Green	Blue	Orange	Brown
U	Black	Black	Yellow	Blue
V	Brown	Purple	Red	Green
W	Blue	White	White	Orange/Yellow
X	Black	Black	White	Red
Y	Yellow	Yellow	Green	Grey
Z	Black	Black	White	Purple

Figure 8. Comparison between dominant colour predictions via colour score from Experiment 2 and modal letter-colour agreement from Experiment 3 for self-reported synaesthete participants. Coloured cells indicate a match between dominant and modal colour(s) for that colour. Matches for white have been coloured dark grey for visibility. Cells with more than one colour indicate a tie in agreement.

Discussion

In this experiment, we 1) calculated a prevalence and gender ratio of self-reported synaesthesia in a sample of American English speakers, 2) calculated the colour scores for each letter and 3) gathered letter-colour associations for these self-reported grapheme-colour synaesthetes. This allows us to evaluate the index word route as an influence on trends in letter-colour pairings.

Although the total number of matches was lower for self-reported synaesthetes, the number of letters for which there was at least one match was higher. We note that while seven letters had a match between first dominant colour and first modal colour for non-synaesthetes, this first-order match accounted for over half of the matches for synaesthetes. Self-reported synaesthetes also had only one double match (for *I*), while non-synaesthetes had five (for *I, K, S, W,* and *Z*). This, as well as the much smaller number of significant letter-colour matches for synaesthetes – 34 for synaesthetes versus 59 for non-synaesthetes – may be due to the much lower numbers of synaesthete participants for both experiments, but particularly Experiment 2 (index words to colours). Even given the small number of synaesthetes, it is striking that index words were able to predict colour associations for the majority of the letters of the alphabet. The nuances and implications of these findings will be explored in more depth below.

General discussion

This study has quantified for the first time the influence of *index words* on the development of letter-colour associations in English (i.e. words starting with a particular letter that are strongly associated with that letter; e.g., *A* is for apple). We examined letter-colour associations that have been previously reported in the literature and also elicited our own responses from American self-reported synaesthetes and non-synaesthetes. We first gathered the set of words that were most commonly associated with each letter of the alphabet and identified the top three for each letter as *index words* (e.g., *A* is for *apple, animal,* and *aardvark*). We then asked what colour was prototypically associated with each index word (e.g., apples are red). Next we calculated a *colour score* predicting the dominant colours for each letter. This colour score combined the percentage agreement from letters to index words, and the percentage agreement from index words to colours. We then compared the colours predicted by index words to the colours reported in direct letter-colour associations. For both synaesthetes and non-synaesthetes, we found that the dominant and/or second-dominant colour (i.e. the colours with the highest and second-highest colour score for each letter; for *A,* red and green respectively) matched the most common direct letter-colour associations for 17 out of 26 letters for non-synaesthetes, and 20 out of 26 letters for self-reported synaesthetes, a rate much higher than chance. This is the first indication that orthographic-semantic associations (beyond red for *R*; e.g., Rich et al., 2005; Simner et al., 2005) have some measurable influence on the development of these population-wide letter-colour pairings. We will address the implications of these findings in order, beginning with a discussion of the nature of index words themselves and their implications for literacy acquisition. We follow this with a discussion of the impact of index words on letter-colour associations in synaesthetes and non-synaesthetes.

Index words in literacy learning

Despite the large body of work on children's acquisition of the alphabet, there is little data on the index words that accompany alphabet instruction. Worden and Boettcher (1990) found that children's success at naming a word beginning with a particular letter (i.e. naming what we have termed *index words*) followed the same pattern as other measures of alphabet literacy, increasing steadily with age, but it was the most difficult of their alphabet knowledge tasks. In that study it was not clear whether children did indeed acquire particular index words, despite training in these associations via alphabet books, because Worden and Boettcher (1990) only recorded the number of letters for which the children could name an index word (i.e. their rate of successful naming), but not which words they produced. The impact of index words in literacy acquisition is also not clear, since children often fail to connect the spelling, sounds, and meanings of words, even when explicitly coached by their parents using alphabet books (B. J. Davis, Evans, & Reynolds, 2010). However, there is abundant evidence that alphabet books do promote alphabet learning (e.g. Both-de Vries & Bus, 2014; Brabham, Murray, & Bowden, 2006; Evans, Saint-Aubin, & Landry, 2009; Murray, Stahl, & Ivey, 1996; Nowak, 2015). Despite this apparent conflict between the efficacy of alphabet-book-based literacy training and the actual ability of children to produce index words, little research has addressed whether index words as such have any direct influence on the acquisition of alphabet or general literacy, or indeed whether they have any enduring connection with individual letters. This study attempted to provide a first indication of this connection into adulthood by showing that some particular letters do have index words, with very high levels of agreement.

Index words in determining letter-colour associations

Our results suggest that index words are one influence (among many; see Introduction) on letter-colour pairings in synaesthetes and non-synaesthetes. Our results show that not only can index words predict the matching of letters with colours, but they can also explain *why* particular letter-colour pairs consistently recur across large groups of both synaesthetes and non-synaesthetes (e.g., why *A* tends to be red but not blue). We will here discuss the influence of index words on letter-colour associations for different groups of letters that appear to show similar patterns: the initial letters of colour terms; the letters for which index words provide a new explanation for consistent colour association trends; and the letters for which we have not yet been able to identify an index-word-based source.

First, we will look at the initial letters of colour terms, namely *B, W, R, O, Y, G, V,* and *P*. On the one hand, it is tempting to disregard the consistent colour association between *R* and red, *B* and blue, etc., since they are obviously the first letters of basic colour terms. However, the relationship between initial letter and colour term is not quite so straightforward – for instance, *B* is consistently associated with blue for both groups, even though *black* is a more frequent colour term (although we note that *black* was the second highest directly associated colour term for non-synaesthetes). The same conflict applies to *P* for *pink/purple* and *G* for *green/gray*. We suggest that *P* may be associated with *pink* because of the typical association of *V* with purple (via *violet*), and *G* with *green* due to higher frequency. More fundamentally, however, colour terms are a clear example of the first letter of a particular word becoming associated with that word's colour due to an orthographic-semantic connection. While this is not exactly the same type of index word influence as we have explored above, it is still a linguistic connotation that is fossilised into an automatic, explicit colour association for synaesthetes, and an intuitive colour association for non-synaesthetes. In other words, for synaesthetes, the colours denoted by colour terms have become indelibly associated with their initial letters past the point of conscious association and into automatic perception (e.g. Dixon, Smilek, & Merikle, 2004; Gray et al., 2006; Mattingley, Rich, Yelland, & Bradshaw, 2001; Mills, Boteler, & Oliver, 1999). However, synaesthetes who have these colour-term associations must form them *after* reaching the realisation that graphemes represent a series of independent sounds which, taken together, can create a word. As Nodelman (2001) summarises, understanding and appreciating the meaning of a phrase like *A is for apple* is a complex process that children struggle with, even when they are explicitly instructed in it. A synaesthete child must realise that the symbol *B* makes the phonetic sound /b/, that the sequence B + L + U + E represents the word *blue*, and connect this abstract representation of the word *blue* with semantic knowledge of the colour blue, before they can form a connection between *B* and blue synaesthetically. Therefore, even colour-term-based synaesthetic colours are necessarily fundamentally rooted in the literacy acquisition process.

The second group of letters to consider are those that have consistent colour associations not explicable by colour terms, but which could nonetheless be predicted from index words such as *apple*. For non-synaesthetes, index-word-predicted dominant colour matched directly-associated modal colour for seven letters, *A, D, E, M, S, X,* and *Z*, and for self-reported synaesthetes for eleven letters, *A, D, H, I, J, M,* *N, Q, U, X,* and *Z.* The number of matches in each participant-group show that index words have a particularly strong influence on synaesthetes, especially since synaesthetes had fewer significant letter-colour pairs overall: 34 for synaesthetes versus 59 for non-synaesthetes. This strong influence might stem from the nature of the development of synaesthetic associations. A child with a genetic predisposition to synaesthesia may find the explicit link between letters and words with strong colours, as alphabet books often present them, to be a compelling formative influence during literacy, as they are predisposed to develop such connections (see Brang & Ramachandran, 2011). Meanwhile, the non-synaesthete child may learn these letter-word-colour connections, but as the associations are less salient, the influence of index words on colour associations may be somewhat diminished as there is no explicit perceptual experience of colour with letters.

The final group of letters are those for which we could not explain the systematic colour trends: *F, K,* and *T* for synaesthetes, and *F, H, J,* and *N* for non-synaesthetes. For example, if the green of *F* is from, for example, *frog*, we have no evidence of it. The lack of matches could be due to several factors. First, some letters (e.g. *N*) had a lack of agreement in index words, indicating that there was no particular word favoured above all others. Another reason we could not explain certain trends is because there was low agreement in direct letter-to-colour mapping; if a letter has no strongly associated colour, any attempt to predict a colour for that letter will fail. Furthermore, there could also have been a lack of agreement in the colour of index words. Although we found that index words were highly imageable, this does not mean they necessarily have a prototypical colour. To use *F* as an example, although *fish* had high agreement in Experiment 1 (25.4%), it could not predict the direct colour associated with *F* because there was no strong prototypical colour to transfer, as blue, silver/grey, or yellow/gold could all be possibilities. On the other hand, a word like *frog* does have a strong prototypical colour (i.e. green), but it had low agreement among our response words for *F* and was therefore not included as an index word. Finally, we have suggested that index words given by adults likely reflect their early learning from alphabet books, but this may not be the case. Certain low-frequency, low-age-of-acquisition words, like *frog, dragon,* or *jungle*, may appear frequently in alphabet- or storybooks for children but are seldom encountered by adults. Therefore, it may be that the high-imageability, low-age-of-acquisition, strongly prototypically coloured words that are frequently encountered by children may be strong influences on the development of letter-colour associations during childhood, but are too infrequent for adults to have been elicited by our study. The next step, currently underway by our lab, is to collect index words from children during literacy acquisition for a clearer representation of the words that are influential during the period of synaesthetic colour association formation.

We must also briefly address the use of data from self-reported and unverified synaesthetes in our analyses. Carmichael, Down, Shillcock, Eagleman, & Simner (2015) randomly tested a large (N = 2847) population for synaesthesia and found that 4.9% self-reported having synaesthesia, while only 1.2% scored below the threshold for genuineness of <1 using an objective measure (see Carmichael et al., 2015; Eagleman et al., 2007). We can therefore estimate that approximately 25% of our self-referred synaesthetes would be confirmed as objectively genuine. We point out that our group of self-reported synaesthetes almost certainly did contain some genuine synaesthetes; while we were unable to test them to ascertain exactly how many, the verification of genuine synaesthesia is an ongoing question in the field, and the number of "true" synaesthetes in our sample would vary depending on which definition we used (see Carmichael et al., 2015; Rothen, Seth, Witzel, & Ward, 2013). Our results nevertheless show that the two groups exhibited commonalities in their letter-colour pairings in line with the trends in these associations that we set out to investigate, which we attribute in some part to the shared influence of index word associations in childhood literacy acquisition. We plan to expand these promising preliminary results with verified synaesthetes in the future.

Our study has provided evidence to support one influence on letter-colour associations that has often been suggested previously. Index words are particularly useful as they allow us to explain *why* particular letter-colour pairs recur. Besides the oft-cited *A is for apple*, we can also now suggest that, for example, *D* is brown because of *dog* and *Q* is purple because of *queen*. A clear way to test these semantic influences on grapheme-colour trends would be to conduct a similar study as that described above in non-English languages. Our favoured example in this study, *apple*, happens to sit at the intersection of index word, red colour, and first letter, so it is difficult to establish which way the direction of influence runs: is *apple* for *A* because *A* is already red, or does the influence of *apple* help lend *A* its redness? In Spanish, for example, *A* may be for *agua* "water" or *árbol* "tree" (*apple* is ineligible, as it is *manzana* in Spanish), in which case we might find that Spanish speakers may be more likely to attribute blue or green to *A* rather than red. Such an investigation could also ask whether the ease-of-colour-generation effects are stronger than these semantic influences.

In conclusion, this study has suggested that both self-reported synaesthetes and non-synaesthetes are influenced in the formation of letter-colour associations by the prototypical colour of index words for each letter. We propose that this semantic influence, rooted in the process of literacy acquisition, works in conjunction with other salient linguistic factors to form lifelong associations between letters and colours. The agreement across a population in index words can explain both why particular letters and colours are paired, and why these associations occur in both synaesthetes and non-synaesthetes. Ultimately, this indicates that grapheme-colour synaesthesia is not a random pairing across modalities, but is rooted in literacy and language-based systematicity.

Acknowledgements

The research leading to these results has received funding for author JS from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n. [617678]. The authors would like to thank Prof. Jamie Ward for his help with our research.

Chapter 6

General Discussion

This thesis set out to apply psycholinguistic methods and theories to the study of synaesthetic colours evoked by letters and words. I have argued that if synaesthesia is indeed a psycholinguistic phenomenon (Mankin, 2017; Simner, 2007), then the features that influence the processing of letters and words should be reflected in the synaesthetic colours for those items. This would establish grapheme-colour synaesthesia as essentially linguistic in nature. The preceding chapters present evidence that synaesthetic colours for linguistic items are indeed fundamentally rooted in features of language. This also means that grapheme-colour synaesthesia, through its systematic mapping of colour to linguistic features, can be used to investigate psycholinguistic questions about normal language processing in everyone. I briefly summarise this thesis' experiments and findings below, then offer recommendations for further expansion and development of the field.

Orthographic and morphological structure in synaesthesia

I began my investigation of the effects of word structure at the larger, morphological level in Chapter 2. The primary aim was to test whether the morphological structure of compound words was represented in their synaesthetic colours across synaesthetes. We first wanted to know whether the number of synaesthetic colours for a compound word like *rainbow* systematically corresponded to its representation in the mental lexicon. To do this, we obtained synaesthetes' colours for compounds varied in their frequency. We hypothesised that low-frequency compounds (e.g. *seahorse*) would be analysed by decomposing them into their two constituent morphemes (e.g. *sea* and *horse*), which would accordingly elicit two colours. However, high-frequency compounds (e.g. *rainbow*) may be lexicalised as wholes, and therefore more likely than low-frequency compounds to elicit a single colour rather than two. We also tested whether the semantic meaning of the compound influenced its synaesthetic colours in the same way that frequency did. Therefore, this experiment obtained synaesthetes' colour associations for two lists of compound words, one varying in frequency (e.g. high-frequency *rainbow* vs. low-frequency *seahorse*) and in semantic transparency (e.g. transparent *birdhouse* vs. opaque *hogwash*). Critically, we allowed synaesthetes to report up to two colours for each compound, and specifically instructed them to provide the "strongest, most dominant" colour in the word, and a second colour if they wished. Our results showed that more frequent compounds were indeed more likely to have one rather than two colours, but the semantic transparency of the compounds had no effect on the number of colours. This indicated that the frequency of the compound may influence its representation in the mental lexicon, and the lexicalisation of high-frequency compounds as single words was reflected by the number of synaesthetic colours. However, compound meaning may be accessed too late in processing to have a similar effect.

To further explore this structure-based synaesthetic colouring, we conducted a follow up to this study, reported in Chapter 3. We returned eight months later to the same synaesthetes and again asked for their colour associations, this time for the constituent morphemes of the compounds from Chapter 2. That is, whereas we had asked them for the colours of *rainbow* in Chapter 2, we next asked them for the colours of *rain* and *bow* separately. Here, we wanted to know what the source of a word's colour was. That is, does the colour of a word come from a particular letter in the word, and does the colour of a compound word (e.g. *rainbow*) come from the colours of its letters or its constituent morphemes? Chapter 3 described two experiments investigating the propagation of synaesthetic colour: letter to word (e.g. *R* to *rain*) and word to compound (e.g. *rain* to *rainbow*). In the first experiment, we found consistent and significant preferences for the colour of a word to match a particular letter in the word – typically the first letter or the first vowel, which we termed the *source letter*. We also found that, analogous to our finding in Chapter 2 that high-frequency compounds are more likely to have one colour, higher-frequency words were more closely and consistently coloured like their source letter than lower-frequency words. It seems that frequent exposure to a word has a consolidating and clarifying effect on the synaesthetic colour associated with it. In the second experiment, we showed that the first-reported, most dominant colour of a compound word comes from its first constituent morpheme, and the secondary colour typically from the second (e.g. if *rainbow* has two colours, its dominant colour matches *rain* and its secondary colour matches *bow*). This further supports the idea that synaesthetic colours for compound words are mapped onto the underlying morphological structure of those words. We also showed that whether the dominant compound colour derives more closely from its first letter or its first morpheme (e.g. whether *rainbow* is coloured more like *R* or *rain*) depends on the frequency of both the constituent morpheme and the whole compound. Specifically, we found that the dominant colour of the whole compound tended to derive from the first morpheme colour when that morpheme was low in frequency but the whole compound was high frequency. Finally, we used this systematic mapping of linguistic characteristics to provide evidence for theories of compound processing and for distinct consonant/vowel processing in word recognition. Overall, these two chapters provided strong evidence that implicit structural features of words are systematically represented in the synaesthetic colours associated with them. The experiments in this thesis have shown that the synaesthetic colours of compound words such as *rainbow* reflect morphological structure in two ways. As Chapter 2 showed, the number of synaesthetic colours reported for a compound word reflects its lexical structure. Chapter 3 elaborated on this by showing that those colours also strongly resemble the colours of the word's morphemes. That is, not only does lower-frequency *seahorse* tend to have more synaesthetic colours than higher-frequency *rainbow*, reflecting a decomposition versus direct lookup processing route, but the colours of *seahorse* and *rainbow* derive measurably from the colours of *sea, horse, rain,* and *bow*.

As mentioned in the general introduction, these studies focused on compound words because their constituent morphemes are also independent words, which allowed us to make comparisons to the colours for simplex words as well. Another type of word that has been frequently studied by psycholinguists, particularly to understand how complex words are processed, are affixed words. Many of the complex word processing theories that this thesis has addressed have tested their predictions using affixed words (e.g., Baayen et al., 1997; Taft, 1979; Taft & Forster, 1975; Taft & Nguyen-Hoan, 2010). Unlike compound words, which contain two independent morphemes, affixed words are composed of an independent morpheme and an affix, which is a meaning-bearing morpheme that cannot stand on its own. In English, these affixes typically take the form of prefixes (e.g. *re-* + *mind* = *remind*) and suffixes (e.g. *corn* + *-y* = *corny*). Therefore, these words can be used to investigate whether the correspondence between synaesthetic colour and morphological structure holds with affixed words, as the studies above have done with compound words. Our lab has already obtained colours from synaesthetes for prefixed and suffixed words (e.g. *remind, corny*). We will contrast the colours for these affixed words with colours from the same synaesthetes for words that appear to be affixed but are in fact monomorphemic, known as pseudoaffixed words (e.g. *relish, fairy*). We can test whether synaesthetic colours map onto morphemes that are not independent words, such as suffixes and affixes (e.g. *re-*, *-y*), by obtaining multiple colours per word.

Based on our findings with compound words, we hypothesise that genuinely affixed words (which have two morphemes, e.g. *remind*) would be more likely to have two colours than pseudoaffixed words (which have only one morpheme, e.g. *relish*). This would point to a decompositional model of word processing, in which a word like *remind* is parsed by first breaking it down into *re-* + *mind* (Stockall & Marantz, 2006; Taft & Nguyen-Hoan, 2010). However, we will also test whether higher-frequency affixed words might be more likely to have a single colour, which would indicate lexicalisation and thereby a multiple-route process in which both decomposition and lexicalisation processes are at work (Baayen et al., 1997; Kuperman et al., 2009; MacGregor & Shtyrov, 2013). If we do not find these morphological and frequency effects, this will help us establish the limits of the systematic correspondence between synaesthetic colour and morphological structure – that is, whether synaesthetes experience distinct colours only for independent words, or for any meaningful linguistic element, such as affixes within a word. In short, as this thesis has shown that morphological structure is systematically reflected in synaesthetic colours, other studies can therefore use synaesthesia to investigate the underlying structure of other words and elements of language, and thereby continue to investigate enduring questions of normal language processing in everyone.

Meaning and imagery in synaesthesia

The second half of this thesis expanded on the psycholinguistic basis of grapheme-colour synaesthesia by investigating the inherent or canonical colour of words. The two experiments detailed in Chapter 4 tested whether the typical colour associated with particular words – e.g. the red colour evoked by the word *red,* or the fiery colour of *fire* – could influence the synaesthetic colour for those words. We compared the synaesthetic colours for words like *red* and *fire* with those of matched control words (e.g. *reed, fine*), which represented the "typical", letter-based synaesthetic colour. Our experiments examined whether the synaesthetic colours for *red* and *fire* were influenced by the canonical colours of those words (e.g. RED, FIRE) relative to the control words. We showed that this was indeed the case both for colour terms (e.g. *red*) and for high-imageability words (e.g. *fire*). For both types of words, the similarity to their canonical colour increased with their frequency – that is, high-frequency *red* and *fire* were closer in colourspace to their canonical colours RED and FIRE than low-frequency *azure* or *feast* were to the colours of AZURE or FEAST. We also showed that although this influence of canonical colours was important for both colour terms and high-imageability words, it was stronger for colour terms. That is, it seems that high-imageability words like *fire* do indeed evoke strong impressions of colours, but this canonical colour is less salient in word processing than it is for colour terms.

These results have several interesting implications. Both Chapters 2 and 3 showed that higher-frequency words were more likely to have a single colour, and that colour was more consistently derived from the colour of a particular letter. Based on these results, very frequent words like *red* should be even more strongly coloured by their letters. For example, if *R* is purple, any high-frequency word beginning with *R* should be consistently coloured the same purple. The tendency for higher-frequency colour terms and high-imageability words to match their *canonical* colour instead presents an apparent contradiction: instead of increased similarity to letter-based colour, our experiments found an increased similarity to canonical colour for these words. I suggest that this reflects an overall trend for higher exposure to a word to produce a more consistent synaesthetic colour. Every time a word is accessed, its associated synaesthetic colour(s) are also automatically evoked as well. For words with a clear canonical colour such as *red* and *fire*, the canonical colour is evoked alongside the synaesthetic colour. If these two colours are incompatible (e.g. letter-based purple and canonical RED for *red*), they are both consistently evoked, and both become strongly associated with the word by repeated exposure. The effect of frequency that we observed suggests that the canonical colour of words is not only explicitly evoked every time the word is accessed, but that the strength of this canonical colour may increase with exposure. That is, the canonical colour of a high-frequency word like *sun* may be more strongly evoked than the same colour in a low-frequency word like *lemon*. If so, in a colour judgment paradigm with an intervening colour word like that employed by Richter and Zwaan (2009; e.g. blue colour patch → the word *red* → blue patch), *sun* would have a more disruptive effect than *lemon*, as its associated colour would be stronger. If so, this in turn would suggest that canonical colours are subject to rehearsal effects – that is, the simulated experience of concepts can be strengthened and clarified by repeated exposure. In other words, the more you look at *sun*, the brighter it becomes. This would mean that studies of linguistically evoked colour, as well as other embodied concepts accessed via language, must control for the frequency of words.

It should also be considered, as Baayen et al. (2010) have suggested, that word frequency is often a shorthand for other, more influential linguistic characteristics.

In this case, it may not actually be higher frequency that increases the likelihood of choosing the canonical instead of the letter-based colour, but rather lower *age of acquisition* (AoA), which is strongly negatively correlated with frequency (i.e. higher-frequency words are typically learned earlier; e.g. Bird et al., 2001; Davies, Izura, Socas, & Dominguez, 2015; Morrison & Ellis, 1995). It may be the case that words with very low age of acquisition (such as *red*) are learned as spoken words[19] before literacy is acquired. In this case, there would be no letter-based colour to conflict with, so for a typically high-imagery synaesthete child, the word *red* would have a strong colour impression based on its canonical colour (i.e. it would be coloured RED). Once grapheme-colour associations develop alongside literacy acquisition, the word *red* gains a second colour based on its synaesthetic letter colours (e.g. purple based on *R*). This preference for the canonical colour of early-learned words would manifest in our analyses as an effect of frequency, due to the strong relationship between frequency and AoA. Indeed, in our wordlist there was a large correlation between Zipf frequency and AoA, so we were unable to test this hypothesis [across both studies in Chapter 4, $r = -0.70$, $p < .001$]. However, a carefully constructed wordlist could contrast frequency with AoA to test whether these two measures predict more consolidated colours for words independently of each other. This is also underscored by the widespread tendency among both synaesthetes and non-synaesthetes to associate the first letter of a colour term with that colour (e.g. *R* is red, *B* is blue, etc.; Jonas, 2010; Rich et al., 2005; Rouw et al., 2014; Simner et al., 2005; Witthoft et al., 2015). This suggests that young learners first learn that the word *red* signifies a particular colour, and then that red colour transfers to the first letter once the spelling is learned.

In Chapter 5, we provide evidence for this hypothesis that canonical colours are a strong influence on the development of synaesthesia in childhood. Here we tested the idea that the meaning of words plays a fundamental role in shaping colour associations for letters. Specifically, through alphabet books and other literacy learning materials, particular words are strongly and repeatedly associated with

---

[19] Children below the age of 4 – 5 are notoriously poor at naming colours, even when they have normal colour vision, have learned the names of colours, etc. Bornstein (1985) suggested that this apparent colour agnosia may indicate that neurological connections in V4 colour processing areas are still in development in childhood. As the V4 area has been implicated in synaesthesia as well (see e.g. Ramachandran & Hubbard, 2001a; Rouw, Scholte, & Colizoli, 2011), it is still unclear how early neurological development may influence both the acquisition of colour names and the development of synaesthetic associations.

their first letter, e.g. "A is for apple", "D is for dog", etc. When these *index words* have a strong canonical colour (e.g. apples are red, dogs are brown), the colour may become associated with the letter via the index word. Therefore, *A* is red because *A* is for *apple* and apples are red. Our experiments in Chapter 5 showed, first, that certain letters do have strongly preferred index words (*A* is indeed for *apple*). We then obtained canonical colours for those index words (apples are canonically red). We used these two connections (*A* → *apple* and apple → red) to predict the colour that each letter should have via its index word (*A* → red via *apple*). Our results showed that the index-word-predicted colour matched the directly associated colour (e.g. *A* → red) far more often than chance would predict, suggesting that index words are a meaningful and enduring influence on letter-colour associations.

Together, Chapters 4 and 5 demonstrate the crucial importance of meaning in synaesthetic associations. Beginning in childhood and enduring into adulthood, the conceptual and semantic information associated with both letters and words has a powerful effect on the colours that synaesthetes experience. The evidence in these chapters suggests that the evoked or mentally imaged colour that accompanies a word – the orange of fire, for instance – has a measurable influence on the colour that the synaesthete experiences for that word. This is not altogether surprising, as the primary purpose of language is to convey meaning (Pinker & Jackendoff, 2005). Much as a word is more than the sum of its letters, but can also convey conceptual information, the synaesthetic colour of a word is more than the sum of the colours of those letters, but also incorporates that conceptual information in the synaesthetic colour.

In conclusion, the studies reported above further show that synaesthetic colours provide a promising psycholinguistic tool for measuring the influence of these conceptually based colour influences. Altogether, this thesis has shown that both the structure and meaning of language play critical roles in synaesthetic experiences, and that the colours that synaesthetes experience for letters and words can be used to investigate normal language processing in everyone.

Sample size, power, and recruitment

An important point to raise to give these findings proper context is the role that sample sizes and individual differences may play in the identification of generalizable trends in grapheme- or word-colour associations. The modern body of research on synaesthesia is founded in single-case studies (e.g. Baron-Cohen et al., 2007, 1987; Dixon et al., 2000; Mills et al., 2002; Simner, Glover, et al., 2006; Simner & Ludwig, 2012; Watson, Akins, Spiker, Crawford, & Enns, 2014), in which extraordinary reports from particular individuals were the main topic of interest. However, the studies described in this thesis, as well as many other papers in the field, now seek to identify underlying commonalities among synaesthetes and extrapolate general patterns about language and cognition for the general population, and in this case, the small sample sizes reported in previous chapters are cause for concern. The small numbers may be due in part to the rarity of the condition and the difficulty of confidently identifying and recruiting synaesthetes to take part in research, but this does not change the fact that synaesthesia research, perhaps even more severely than many areas of psychology (Brysbaert & Stevens, 2018), is chronically underpowered, with generalisations about the nature of synaesthesia derived from work with very small sample sizes. As an illustration of this problem, in a forthcoming meta-analysis of research on synaesthesia and memory, the average sample size per experiment across 29 papers was N = 17.10, with only three of those papers reporting sample sizes greater than 30 and six with sample sizes smaller than 10 (Ward, 2018). These small sample sizes are directly linked to a lack of statistical power to detect important effects, which raises concerns about the reliability and replicability of statistically “significant” findings (see e.g. Benjamin et al., 2017; Maxwell, 2004).

Accordingly, caution must be applied as well to the experiments reported in this thesis, which also tended to have statistically small sample sizes (Chapter 2, N = 19; Chapter 3, N = 12; Chapter 4, N = 16). For this among other reasons, the experiments reported in these chapters have utilised linear mixed effects analyses, which have greater power in comparison to more traditional methods (e.g. separate by-participant and by-item F-tests or similar) due to the lack of prior averaging (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017). Using these techniques, I have been able to identify underlying influences that colour synaesthetes' experiences, but I would be remiss if I did not acknowledge the inordinate impact of the specific and highly individual experiences of my few dozen synaesthete participants on the overall picture of synaesthesia and language. As I will argue further on, large-scale investigation is needed, both qualitative and quantitative, to confirm and expand these initial exploratory findings, as well as clarify some of the unexpected findings reports in the previous chapters.

Chapter 5 presented data from self-reported synaesthetes, who were recruited via Mechanical Turk rather than selected from a pool of verified synaesthetes. This method of recruitment allowed us to have a larger sample size and compare their results to the non-synaesthetes within the same study. However, this sample was still objectively quite small (24 in Experiment 2, 64 in Experiment 3), and as discussed in the chapter, the number of synaesthetes within that number who would be identified as genuine by the standards of the consistency test the other chapters have employed (Eagleman et al., 2007; Rothen et al., 2013) would very likely be smaller still. This means it is nearly unavoidable that some of the data from our self-reported "synaesthetes" did not in fact come from synaesthetes, which limits the generalisation of our findings to the synaesthete population. In contrast, Chapters 2, 3, and 4, reported even smaller samples, but we could be confident that those included were in fact synaesthetes (at least by the current verification standards; see Simner, 2012a, 2012b). To summarise, the conclusions from the studies presented above represent an important initial step in investigating synaesthesia psycholinguistically, but they are limited by the difficulty inherent in identifying and recruiting a sufficient number of participants with a rare condition like synaesthesia. As I will next argue, obtaining data from a wider variety of synaesthetes and developing tests to better capture synaesthete experiences should be the focus of immediate action, particularly for studying synaesthesia and language.

Next steps in studying the psycholinguistics of synaesthesia

The preceding chapters have also shown that some of the assumptions frequently made about grapheme-colour synaesthesia may not be consistently true of all synaesthetes, and may in fact obscure the true nature of synaesthetes' experiences. Two of these assumptions, in particular, are challenged by the results presented in this thesis: one about the *number* of colours in a word, and the other about the *nature* of those colours. That is, synaesthesia researchers have frequently assumed that a single colour is sufficient to represent a synaesthete's colour association for a word (e.g., Asano & Yokosawa, 2012; Barnett et al., 2009; Rich et al., 2005), and that that single colour is straightforwardly derived from a single letter source within that word (e.g., Simner, Glover, et al., 2006; Ward et al., 2005). This assumption was indeed based on case studies and reports from synaesthetes themselves (Baron-Cohen et al., 1993, 1987; Mills et al., 2002; Rich et al., 2005). However, the combined evidence presented in this thesis suggests that these assumptions do not necessarily account for important elements of many synaesthetes' experiences. The studies reported above have shown, first, that synaesthetes frequently ask for the ability to report more than one colour per word, and indeed when given this opportunity, report two colours more often than only one. Second, the colour of a word is often a composite of multiple colours in a word, and attempting to attribute this colour to a single source letter may obscure important influences on whole-word colour[20]. The above chapters have further suggested that these problems are connected: restricting synaesthetes to one colour may distort the accuracy of that colour, as they try to capture multiple colour impressions in one shade (e.g. by attempting to represent both light pink and dark maroon with a dark pink). Therefore, the methods of documenting synaesthetes' colours must be expanded and improved if further research expects to make meaningful contributions to the understanding of synaesthesia and language.

First, it is imperative that the *number* of colours synaesthetes experience for words be systematically investigated on a large scale. Experimental reports often state that synaesthetes experience a word colour that matches the colour of the first letter (Baron-Cohen et al., 1993; Mills et al., 2002; Rich et al., 2005; Ward et al., 2005) or the first vowel (Simner, Glover, et al., 2006). In Chapter 3, we reported similar results, and suggested that this may reflect the underlying processes of visual word recognition, which place disproportionate weight on the identity of the first letter. However, it could be that when asked to give the colour of a word, synaesthetes most frequently report the colour of the first letter *because it is first* as well as important in lexical processing. Given the assumption that words have only one colour, no further colours are asked for or accepted, leading to the impression that that one colour comprises the whole-word colour. However, when given the chance to describe their word colours, case studies document multiple colours: for example, the city name *Catonsville* appealed to synaesthete MLS because of its "browns and greens and this nice shiny *N*" (Mills et al., 2002, p. 1376), and synaesthete MD mentioned that *banana* was yellow although it "should be dark blue and black" (Rich et al., 2005, p. 25). Simner, Glover, and Mowat (2006) also showed that letters downstream in a word could reinforce the synaesthete's colour experiences (e.g. the synaesthetic colour of *ether* named more quickly than that of *ethos*), implying that multiple colours

---

[20] Some accounts from synaesthetes, as detailed below, mention having multiple colours for individual *letters*, which further complicates the issue.

in a word are pertinent to the whole word's colour. Critically, though, the widespread assumption that all words have a colour based *only* or primarily on one particular letter, typically the first, has never been objectively verified. Only two studies to date have explicitly obtained multiple colours per word in an experimental setting: our compound words study, Mankin et al. (2016 [Chapter 2]), and a case study by Blazej and Cohen-Goldberg (2015), produced simultaneous to this thesis. In both these studies, the one- vs two-colour difference was of experimental interest, but to date no study has investigated how many colours synaesthetes *typically* experience or report per word, as a baseline or as a function of length, morphological structure, frequency, etc. Furthermore, it is not clear whether the typical number of colours synaesthetes experience for words is discernibly different from the number of colours non-synaesthetes would give for the same words. This is a fundamental gap in the scientific portrayal of grapheme-colour synaesthesia, but also represents a golden opportunity both to expand our understanding of synaesthesia itself and gain a clearer reference point from which to explore exactly how synaesthesia maps onto language. I note again that the single letter colour → single word colour assumption that has dominated grapheme-colour research thus far was based on synaesthetes' reports, and may indeed be true for many synaesthetes. However, it is clear that some synaesthetes have more complex colour experiences, and it is impossible to know what elements of those experiences are of psychological or scientific import until they are studied.

The second, related, assumption is that the colour of a word derives straightforwardly from the colour of a particular dominant grapheme – that is, if *R* is purple, so will be *rain* and *rhythm* and any other word beginning with *R*. There is now ample evidence to suggest that this explanation glosses over complex features of word colouring that could be informative for both psycholinguistics and synaesthesia. First, these conclusions were often reached with colour-term descriptions (e.g. synaesthetes asked to describe in words the colour they experienced for *R* and *rain*), where using the same colour category when describing the colour of the word and the colour of its initial letter counted as a match (e.g. Rich, Bradshaw, & Mattingley, 2005; Ward, Simner, & Auyeung, 2005). Even when more precise colours were collected using colour palettes, they were often interpreted in terms of colour categories (e.g. Asano & Yokosawa, 2012). The experiments we described using colourspace distances show that the colours of words can vary in subtle ways that cannot be adequately captured using only colour categories or verbal descriptions. That is, the synaesthetic colour of purple *rain* might be bluer to reflect the canonical colour associated with rain, while *rhythm* might be lighter due to yellows and oranges for *H, Y,* and *M,* but this would still count as "purple" (see Chapter 4). Although this distinction appears minor, it is extremely pertinent to accurately understanding subtle influences of spelling, meaning, and canonical colour, which have been of interest to us here. Online colour palettes, like that available through the Synesthesia Battery (www.synesthete.org) or the Synaesthesia Toolkit (www.syntoolkit.org) with a customised interface for synaesthesia research, make written descriptions obsolete. However, the primary difficulty involves both sensitivity and number of colours: namely, that in the typical research paradigm deployed by synaesthesia researchers – including the studies I have designed and run for this thesis – neither the number nor detail of the colours obtained from synaesthetes adequately reflects their experiences, and these two issues in combination may actually obscure the genuine experience of synaesthesia. Besides the examples documented in the preceding chapters, evidence that this is indeed the case comes from the synaesthetes themselves, who have reported this in so many words, as described below.

Letter dominance – an essential component of the complete picture

The synaesthetes who participated in the experiments reported in Chapter 4 answered two additional questions. The first asked them to describe where the colours of whole words came from, and the second was an open text box to write any thoughts or feedback they wanted us to know. Between these two spaces, 16 out of the 20 synaesthetes who completed the test provided written descriptions of their synaesthetic experiences. The voluntary descriptions of these colour experiences, many of them thoughtful and detailed, paint quite a different picture than a straightforward first-letter or first-vowel source categorisation. Almost all of the comments describe experiencing multiple colours in a word. Several participants mentioned that the first letter tended to be influential, but are also subject to other influences downstream: "In the majority of cases, the whole word colour comes from one of the first few letters in the word…Often, others letters in the word affect the overall colour of a word, such as the word bleed - b words are often bright blue (for b), but bleed is very dark blue because of the black of the 'e's…" These reflections on the synaesthete participants' perceptual experience highlights the diversity and complexity of synaesthetes' experiences, as well as evident introspective thoughtfulness about the systematic influences that form their colours.

Besides describing multiple colours in each word, a common theme in the responses was the inadequacy of the online colour palette for capturing synaesthetic experiences. Several synaesthetes mentioned that the colour selection task, and particularly the option of choosing only one colour, did not reflect their actual experience. One synaesthete said that they experience "several colors simultaneously – i.e., instead of mixing red and blue to get purple, imagine wearing those 3D glasses that have one blue lens and one red lens, so you're experiencing both colors at the same time. Only being able to choose one color for a word is difficult and sometimes inaccurate [*sic*]." Another characterised their colours as "'metallic' in nature sometimes or like the northern lights…my colours have grades and variants and shades. A photoshop like pallet [*sic*] is sometimes hard to choose a colour of a word with because my colours can be detailed and varying within a space." Tellingly, one synaesthete describes experiencing colours in a way that cannot be captured by a colour palette, and mentions doing their best to make a decision based on other factors: "[S]ome of the words I saw with two colours in a stripe effect based on dominant letter and meaning. I didn't know how to indicate this and so chose the strongest colour if it was dominant." This not only indicates that the simple single-colour test may not be adequate, but also introduces the idea of letter dominance.

The most striking recurrent theme across these responses was the repeated reference to "strong" or "dominant" letters, which have a greater influence on the colour of the word than "weaker" letters. This is sometimes the first letter, but not always. One synaesthete explains:

> "Some letters are stronger than others, and so in those cases when they are the first letter of a word, they color the entire word. Other letters are less strong, and so those tend to blend together to make a new color. If it's a blend of strong letters, then it's like a mural of very distinct, individual colors. And then every once in a while there are letters that have a pattern with multiple colors."

This characteristic of letter dominance has, to date, never been experimentally investigated by synaesthesia researchers[21], but in this sample, over half (N = 13,

[21] I have employed the term "dominant" several times in this thesis to describe the strongest colour in a compound word (Chapter 2) and the most influential letter in determining a whole word's colour (Chapter 3). This sense of "dominance" has received a good degree of attention from synaesthesia researchers, including the studies described in this thesis (see also Blazej & Cohen-Goldberg, 2015; Rich et al., 2005; Simner, Glover, et al., 2006; Ward et al., 2005; etc.). However, the quality of "letter dominance" described at present is a different characteristic, referring to an intrinsic quality of the letter itself. Despite the potential for

65%) of the synaesthetes independently and spontaneously mentioned letter dominance as a primary influence in determining word colour[22]. Many of them do note that the dominant letter is also often the first letter of the word, but explicitly point out that other dominant letters further on in the word have a disproportionate influence on word colour, for example: "For words, it can sometimes be the colour of the first letter, or a mixture of several dominant colours in the word." Another participant explains that letter dominance can be the primary source of whole-word colour: "The colour of the word is formed from the dominant letters. For example 'f' is a dominant letter and will give the word a green tinge. 'a' is dominant letter and will give the word a yellow colour." It is particularly interesting to note that only three of the synaesthetes mentioned a strong or dominant *colour*; this dominance was usually described as a property of the letter itself, not the colour it was associated with. Therefore, it seems that dominant letters have a disproportionate influence on the synaesthetic colours of whole words, regardless of the nature of the colour associated with that letter and possibly independent of, or despite, position in the word. This strongly implies that attempting to capture letter-to-word colouring patterns without taking letter dominance into account presents an incomplete picture of the true nature of whole-word colouring.

Recommendations for future work

As a matter of urgency, all future research into synaesthesia, and especially that interested in complex and nuanced influences of language and meaning, must consider carefully the assumptions that underlie current research design. It is past time to discard research practices that are retained not because they accurately capture synaesthetic experiences, but because they simplify the collection and analysis of data for researchers. As an example of the potential dangers of such practices, the evidence from synaesthetes presented above supports the argument that the single-colour-selection task, which is the primary, standard method of obtaining synaesthetic colour associations – and, thereby, verifying a synaesthete as

---

confusion, I have again used the term "dominant letter" here because it is the term most frequently used by synaesthetes themselves to describe this quality.

[22] I emphasise here that all of these studies were conducted online, without any contact between synaesthetes. However, as they were a part of a database specifically for synaesthesia research, some of them may have participated in studies before, and possibly had contact with each other. Nevertheless, it is unlikely that these participants obtained the idea from other sources.

"genuine" – may be based on incomplete or inaccurate diagnostic criteria. If a synaesthete sees a letter or word as equally red and blue simultaneously as described above, requiring that they choose between these two colours on different occasions may erroneously result in higher inconsistency or difficulty in discerning colour patterns when this colour complexity is not taken into account. Furthermore, this overreliance on a narrow diagnostic for synaesthesia may lead to a self-reinforcing but unenlighting feedback loop. That is, researchers select grapheme-colour synaesthetes who are consistent on the standard letter-colour test and ignore the ones who report synaesthetic experiences that do not fit this mould. These consistent synaesthetes are then included in studies and – unsurprisingly – show consistent patterns in letter and word colouring, which implies that this consistency is characteristic of synaesthetic experiences. However, if a synaesthete's whole-word colours are influenced or determined by the inherent dominance of the letters in the word, as described above – an as-yet-unresearched but apparently pervasive and influential quality – then attempting to describe their colour patterns based only on colour or linguistic factors will not capture a vital element of the underlying system. This may explain at least some instances of so-called "malingerers", who claim to have synaesthesia but do not meet the consistency threshold that is taken as indicative of genuine synaesthesia (see Carmichael et al., 2015; Eagleman, 2009; Novich et al., 2011; Simner, 2012b). That is, the current, widespread methods of collecting synaesthetic associations, evaluating consistency, and analysing systematicity may actually penalise or disqualify those with the richest and most complex synaesthetic experiences simply because the apparatus used to evaluate those experiences does not accurately capture such complexity.

Of course, researchers studying synaesthesia cannot account for all the individual nuances in synaesthetic experiences if they want to be able to identify underlying patterns and systematic relationships. At some point, the map becomes the territory, and the idiosyncratic quirks of letter and colour that, to the synaesthete, are a valued and cherished part of their perception of the world become indecipherably complex to the scientist attempting to understand them and discern underlying structures. Although there are limitations on the complexity of the data researchers can meaningfully analyse, this does not imply that the current understanding of synaesthetic experience is the best, or most accurate, representation that research can produce. Indeed, when synaesthetes spontaneously and consistently report the need for multiple colour options, or a measure of letter dominance, in order to accurately represent their synaesthetic experiences, it follows that studies on synaesthesia that disregard these considerations will miss opportunities to capture potentially fundamental aspects of the synaesthetic experience. Based on feedback from synaesthetes and the results of the analyses described in this thesis, I therefore propose the following recommendations for future work studying the synaesthetic colouring of letters and words.

First, before developing any specific hypotheses based on these reports, I must emphasise that synaesthetes themselves are an underutilised source of advice and inspiration for future studies. As the introduction details in depth, synaesthetic experiences can and have been objectively verified to be real, but the nature of those experiences – that is, what a given synaesthete actually sees, or hears, or feels, when presented with an inducing stimulus – cannot be objectively ascertained by technological or behavioural methods. In short, researchers desiring to learn what the experience of synaesthesia is actually like must *ask synaesthetes*; otherwise, they will be investigating synaesthesia as they assume it must be, rather than as it actually is experienced. Therefore, I propose a systematic and carefully considered qualitative investigation into the experiences of synaesthetes when viewing, hearing, or using language, without assuming what may or may not be important. The first step in developing this investigation would be to contact the hundreds of grapheme-colour synaesthetes already registered in the Sussex Synaesthesia Database with a few simple preliminary questions, such as: What do you think researchers should investigate about synaesthesia? What do you think influences your colours for letters and/or words? What do you think is missing in synaesthesia research that is important in your own everyday experience? Combining the answers from these questions with previous reports and my own findings, I would then develop a larger qualitative investigation, with two main goals: to paint a detailed picture of what synaesthetes actually experience when they use language, and to identify common patterns of experience that may inspire new research questions. To do this, I would develop sets of questions that target language at various levels of complexity, from letters and spoken phonemes to simple words, complex words, and sentences. I would also include exemplars of words that have particular properties – such as morphological complexity, as in Chapters 2 and 3, or the pre-existing colour connotations of *red* and *fire* as in Chapter 4 – and ask synaesthetes how they experience such words. Most importantly, I would draw on the answers from the preliminary study to identify elements of synaesthetic experience, such as letter dominance described above, that have not yet received research attention, and attempt to ascertain how widespread and influential such experiences are. This systematic overview of a wide variety of synaesthetic experiences would allow synaesthetes to provide input on what they perceive to be the most salient aspects, such as the need for multiple colours, which could – and ought to – inform the development of tests and research questions that more accurately reflect the lived experience of synaesthetes. Indeed, Smilek and Dixon (2002) similarly advocated for first-person reports from synaesthetes to be integrated into the research design and interpretation process, citing many early successes in small case studies that synergistically combined experimental and subjective data. I believe that an investigation of this type would not only produce a wealth of insights and new directions for synaesthesia research, but would also give a voice in directing and developing the research process to those who know best what synaesthesia is actually like.

In the interim, this thesis has already identified some specific ways that current research could be improved or expanded. For instance, I have shown in depth that the current methods of obtaining and analysing synaesthetes' colour experiences are often inadequate. However, it is not enough to offer more than one colour option, as the comments reported above indicate that synaesthetes are highly sensitive to the position of letters in the word, the meaning and nature of the words, and, most critically, the relative strength or dominance of those letters. As the comments above have shown, the quality of letter dominance is a widespread and influential characteristic of synaesthetic colour associations, if its spontaneous mention by a majority of synaesthetes is any indication. Accordingly, any test of synaesthetic colours for letters should include the option to indicate the dominance of the letter or colour. This will better account for patterns in colouring such as those identified in this thesis, but more importantly, it opens a wide variety of interesting possibilities for study. Of particular interest is whether this dominant quality is a type of personality associated with the letters, as in ordinal linguistic personification (i.e. a different subtype of synaesthesia which associates personality traits with graphemes, e.g. *C* is a timid male or *8* an perfectionistic female; Simner & Holenstein, 2007; Smilek, Malcolmson, et al., 2007). Alternatively, letter dominance may not be a full-fledged personality trait, but rather a conceptual property of each letter. Future research in this area could examine whether letter dominance is further related to properties of the associated colours (e.g. lightness and saturation), particular hues, letter frequency, position in the alphabet sequence, shape, and so on. It may be that there are trends in the dominance of certain letters as there are trends in their associated colours (see Chapter 5), which could be connected to the processes involved in recognising graphemes and words and therefore uncover a new tool for studying these psycholinguistics mechanisms.

Finally, this thesis has only begun to describe the influence of canonical colours (e.g. the fiery orange of "fire") on synaesthetic experiences, an area which warrants much further attention. In order to study the effects of canonical colour both on typical language processing and synaesthetic colours, I recommend a standardised method or database for obtaining the prevailing canonical colours of a large set of words, as well as a standardised measure of the consistency or degree of agreement for the canonical colour. To date, no such database or measure of "canonicity" exists, and related measures such as imageability do not capture the same quality (e.g. both *son* and *sun* are highly imageable but differ in the consistency of their canonical colour associations). We have shown here that canonical colouring is a powerful influence both in the development of letter-colour associations (Chapter 5) and in the colouring of whole words (Chapter 4), and any study of whole-word colouring must account for this influence. For instance, it may be that letter and canonical colours can interact in word colouring. That is, the influence of canonical colour may depend on the presence of letters in the word that reinforce this colour, e.g. the synaesthetic colour of *flame* may be more fiery than that of *fire* because *A* tends to be red, reinforcing the canonical colour of FLAME. This again underlines the need to offer more colours per word in order to understand these complex interactions properly. As these canonical colour effects mirror those identified in non-synaesthetes, a standardised database of canonical colours and further investigations of their influences would benefit studies in the general population as well.

Conclusions

I have argued here that in grapheme-colour synaesthesia, letters are not simply symbols with colours associated with them, and the colours of words are not straightforwardly calculable from the letters that compose them – that is, the whole-word colour is more than the sum of its parts. The meanings, ideas, and possibly personal experiences associated with words may all inform their synaesthetic colours. The same may be true of letters, as the mediation of index words between letter and colour shows. Therefore, a complete understanding of connections between letters and colours in synaesthesia must necessarily include the measures associated with those linguistic items, including frequency, meaning, and associated canonical colour. The identity and qualities of the individual word may have important implications for synaesthetic experiences, and therefore for the study of both synaesthesia and language, and cannot be ignored when investigating these experiences. That is, not only is grapheme-colour synaesthesia useful for researching psycholinguistics, it is an essentially and fundamentally psycholinguistic phenomenon, and must be researched and understood through that lens.

In conclusion, this thesis has investigated the nature of the "special" status of language in synaesthesia. I have shown that both structural and semantic properties of linguistic inducers map systematically onto features of synaesthetic colours. I have also argued that this systematic correspondence between features of language and synaesthetic colour makes synaesthesia a promising new tool to investigate complex and implicit processes in natural language processing. Finally, I have demonstrated this by testing enduring questions in the field of psycholinguistics – namely word reading, compound processing, and semantic influences on word and letter processing – using colour evidence from synaesthetes. This thesis has shown that the psycholinguistics of synaesthesia can provide a better understanding of the intricacies of synaesthetic experience, as well as many new insights into both language and cognition in everyone.

References

Adelman, J. S. (2011). Letters in time and retinotopic space. *Psychological Review*, *118*(4), 570–582. https://doi.org/10.1037/a0024811

Altarriba, J., Bauer, L. M., & Benvenuto, C. (1999). Concreteness, context availability, and imageability ratings and word associations for abstract, concrete, and emotion words. *Behavior Research Methods, Instruments, & Computers*, *31*(4), 578–602. https://doi.org/10.3758/BF03200738

Andrews, S., Miller, B., & Rayner, K. (2004). Eye movements and morphological segmentation of compound words: There is a mouse in mousetrap. *European Journal of Cognitive Psychology*, *16*(1–2), 285–311. https://doi.org/10.1080/09541440340000123

Asano, M., & Yokosawa, K. (2012). Synesthetic colors for Japanese late acquired graphemes. *Consciousness and Cognition*, *21*(2), 983–93. https://doi.org/10.1016/j.concog.2012.02.005

Aschenbrenner, A. J., Balota, D. A., Weigand, A. J., Scaltritti, M., & Besner, D. (2017). The first letter position effect in visual word recognition: The role of spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(4), 700–718. https://doi.org/10.1037/xhp0000342

Baayen, R. H. (2005). Storage and computation in the mental lexicon. In G. A. L. Jarema (Ed.), *The Mental Lexicon: Core Perspectives* (pp. 81–104). Elsevier.

Baayen, R. H. (2013). LanguageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics" R package version 1.4.1. Retrieved from https://cran.r-project.org/package=languageR

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subject and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and Plurals in Dutch: Evidence for a Parallel Dual-Route Model. *Journal of Memory and Language*, *37*(1), 94–117. https://doi.org/10.1006/jmla.1997.2509

Baayen, R. H., Kuperman, V., & Bertram, R. (2010). Frequency effects in compound processing. In S. Scalise & I. Vogel (Eds.), *Compounding*. Amsterdam/Philadelphia: Benjamins.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database {CD-ROM}.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology. General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*(3), 445–459. https://doi.org/10.3758/BF03193014

Banissy, M. J., Cassell, J. E., Fitzpatrick, S., Ward, J., Walsh, V. X., & Muggleton, N. G. (2012). Increased positive and disorganised schizotypy in synaesthetes who experience colour from letters and tones. *Cortex*, *48*(8), 1085–1087.

https://doi.org/10.1016/j.cortex.2011.06.009

Banissy, M. J., Holle, H., Cassell, J., Annett, L., Tsakanikos, E., Walsh, V., … Ward, J. (2013). Personality traits in people with synaesthesia: Do synaesthetes have an atypical personality profile? *Personality and Individual Differences*, *54*(7), 828–831. https://doi.org/10.1016/j.paid.2012.12.018

Bankieris, K., & Simner, J. (2015). What is the link between synaesthesia and sound symbolism? *Cognition*, *136*, 186–95. https://doi.org/10.1016/j.cognition.2014.11.013

Barnett, K. J., Feeney, J., Gormley, M., & Newell, F. N. (2009). An exploratory study of linguistic-colour associations across languages in multilingual synaesthetes. *Quarterly Journal of Experimental Psychology (2006)*, *62*(7), 1343-54–5. https://doi.org/10.1080/17470210802483461

Barnett, K. J., & Newell, F. N. (2008). Synaesthesia is associated with enhanced, self-rated visual imagery. *Consciousness and Cognition*, *17*(3), 1032–9. https://doi.org/10.1016/j.concog.2007.05.011

Baron-Cohen, S., Bor, D., Billington, J., Asher, J. E., Wheelwright, S., & Ashwin, C. (2007). Savant memory in a man with colour form-number synaesthesia and Asperger syndrome. *Journal of Consciousness Studies*, *14*(9–10), 237–251. Retrieved from http://www.ingentaconnect.com/content/imp/jcs/2007/00000014/F0020009/art00011

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia: prevalence and familiality. *Perception*, *25*(9), 1073–9.

Baron-Cohen, S., Harrison, J., Goldstein, L. H., & Wyke, M. A. (1993). Coloured speech perception: is synaesthesia what happens when modularity breaks down? *Perception*, *22*(4), 419–26.

Baron-Cohen, S., Wyke, M. A., & Binnie, C. (1987). Hearing words and seeing colours: an experimental investigation of a case of synaesthesia. *Perception*, *16*(6), 761–7.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioural and Brain Sciences*, *22*, 577–660. https://doi.org/10.1017/S0140525X99532147

Barsalou, L. W. (2008). Grounded Cognition. *Annu. Rev. Psychol*, *59*, 617–45. https://doi.org/10.1146/annurev.psych.59.103006.093639

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bedny, M., & Thompson-Schill, S. L. (2006). Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. *Brain and Language*, *98*(2), 127–39. https://doi.org/10.1016/j.bandl.2006.04.008

Beeli, G., Esslen, M., & Jäncke, L. (2007). Frequency correlates in grapheme-color synaesthesia. *Psychological Science*, *18*(9), 788–92. https://doi.org/10.1111/j.1467-9280.2007.01980.x

Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., … Johnson, V. E. (2017). Redefine statistical significance. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-017-0189-z

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution.* Berkeley, CA: University of California Press.

Bertram, R., & Hyönä, J. (2003). The length of a complex word modifies the role of morphological structure: Evidence from eye movements when reading short and long Finnish compounds. *Journal of Memory and Language*, *48*(3), 615–634. https://doi.org/10.1016/S0749-596X(02)00539-9

Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(2), 489–511. https://doi.org/10.1037//0278-7393.26.2.489

Bien, H., Levelt, W. J. M., & Baayen, R. H. (2005). Frequency effects in compound production. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(49), 17876–81. https://doi.org/10.1073/pnas.0508431102

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, *33*(1), 73–79. https://doi.org/10.3758/BF03195349

Blair, C. D., & Berryhill, M. E. (2013). Synesthetic grapheme-color percepts exist for newly encountered Hebrew, Devanagari, Armenian and Cyrillic graphemes. *Consciousness and Cognition*, *22*(3), 944–54. https://doi.org/10.1016/j.concog.2013.06.002

Blazej, L. J., & Cohen-Goldberg, A. M. (2015). Multicolored words: Uncovering the relationship between reading mechanisms and synesthesia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *75*, 160–179. https://doi.org/10.1016/j.cortex.2015.11.017

Bor, D., Billington, J., & Baron-Cohen, S. (2007). Savant memory for digits in a case of synaesthesia and Asperger syndrome is related to hyperactivity in the lateral prefrontal cortex. *Neurocase*, *13*(5), 311–9. https://doi.org/10.1080/13554790701844945

Bornstein, M. H. (1985). On the Development of Color Naming in Young Children: Data and Theory. *Brain and Language*, *26*, 72–93.

Both-de Vries, A. C., & Bus, A. G. (2014). Visual processing of pictures and letters in alphabet books and the implications for letter learning. *Contemporary Educational Psychology*, *39*(2), 156–163. https://doi.org/10.1016/j.cedpsych.2014.03.005

Brabham, E. G., Murray, B. a., & Bowden, S. H. (2006). Reading Alphabet Books in Kindergarten: Effects of Instructional Emphasis and Media Practice. *Journal of Research in Childhood Education*, *20*(3), 219–234. https://doi.org/10.1080/02568540609594563

Brang, D., Hubbard, E. M., Coulson, S., Huang, M., & Ramachandran, V. S. (2010). Magnetoencephalography reveals early activation of V4 in grapheme-color synesthesia. *NeuroImage*, *53*(1), 268–74. https://doi.org/10.1016/j.neuroimage.2010.06.008

Brang, D., Miller, L. E., McQuire, M., Ramachandran, V. S., & Coulson, S. (2013).

Enhanced mental rotation ability in time-space synesthesia. *Cognitive Processing*, *14*(4), 429–34. https://doi.org/10.1007/s10339-013-0561-5

Brang, D., & Ramachandran, V. S. (2011). Survival of the synesthesia gene: why do people hear colors and taste words? *PLoS Biology*, *9*(11), e1001205. https://doi.org/10.1371/journal.pbio.1001205

Brang, D., Rouw, R., Ramachandran, V. S., & Coulson, S. (2011). Similarly shaped letters evoke similar colors in grapheme-color synesthesia. *Neuropsychologia*, *49*(5), 1355–8. https://doi.org/10.1016/j.neuropsychologia.2011.01.002

Breheny, P., & Burchett, W. (2016). visreg: Visualization of Regression Models. R package version 2.3-0. Retrieved from https://cran.r-project.org/package=visreg

Bronk, M., Zwitserlood, P., & Bölte, J. (2013). Manipulations of word frequency reveal differences in the processing of morphologically complex and simple words in German. *Frontiers in Psychology*, *4*(August), 546. https://doi.org/10.3389/fpsyg.2013.00546

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–90. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, *1*(1), 1–20. https://doi.org/10.5334/joc.10

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, *6*(1), 3–5. https://doi.org/10.1177/1745691610393980

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production, Vol II: Development, writing and other language processes* (pp. 257–294). London: Academic Press.

Caramazza, A., Miceli, G., Chialant, D., & Capasso, R. (2000). Separable processing of consonants and vowels. *Nature*, *403*(6768), 428–430. https://doi.org/10.1038/35000206

Carmichael, D. A., Down, M. P., Shillcock, R. C., Eagleman, D. M., & Simner, J. (2015). Validating a standardised test battery for synesthesia: Does the Synesthesia Battery reliably detect synesthesia? *Consciousness and Cognition*, *33*, 375–385. https://doi.org/10.1016/j.concog.2015.02.001

Carreiras, M., Duñabeitia, J. A., & Molinaro, N. (2009). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. *Cerebral Cortex*, *19*(11), 2659–2670. https://doi.org/10.1093/cercor/bhp019

Carreiras, M., Gillon-Dowens, M., Vergara, M., & Perea, M. (2008). Are Vowels and Consonants Processed Differently? Event-related Potential Evidence with a Delayed Letter Paradigm. *Journal of Cognitive Neuroscience*, *21*(2), 275–288.

Carreiras, M., & Price, C. J. (2008). Brain activation for consonants and vowels. *Cerebral Cortex*, *18*(7), 1727–1735. https://doi.org/10.1093/cercor/bhm202

Carreiras, M., Vergara, M., & Perea, M. (2007). ERP correlates of transposed-letter similarity effects: Are consonants processed differently from vowels? *Neuroscience Letters*, *419*(3), 219–224. https://doi.org/10.1016/j.neulet.2007.04.053

Carreiras, M., Vergara, M., & Perea, M. (2009). ERP correlates of transposed-letter priming effects: The role of vowels versus consonants. *Psychophysiology*, *46*(1), 34–42. https://doi.org/10.1111/j.1469-8986.2008.00725.x

Chambers, S. M. (1979). Letter and order information in lexical access. *Journal of Verbal Learning and Verbal Behavior*, *18*(2), 225–241.

Chanceaux, M., & Grainger, J. (2012). Serial position effects in the identification of letters, digits, symbols, and shapes in peripheral vision. *Acta Psychologica*, *141*(2), 149–158. https://doi.org/10.1016/j.actpsy.2012.08.001

Chun, C. A., & Hupé, J. M. (2016). Are synesthetes exceptional beyond their synesthetic associations? A systematic comparison of creativity, personality, cognition, and mental imagery in synesthetes and controls. *British Journal of Psychology*, *107*(3), 397–418. https://doi.org/10.1111/bjop.12146

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: Lawrence Earlbaum Associates.

Cohen Kadosh, R., & Henik, A. (2007). Can synaesthesia research inform cognitive science? *Trends in Cognitive Sciences*, *11*(4), 177–84. https://doi.org/10.1016/j.tics.2007.01.003

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology Section A*, *33*(4), 497–505. https://doi.org/10.1080/14640748108400805

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Connell, L. (2007). Representing object colour in language comprehension. *Cognition*, *102*(3), 476–485. https://doi.org/10.1016/j.cognition.2006.02.009

Connell, L., & Lynott, D. (2009). Is a bear white in the woods? Parallel representation of implied object color during language comprehension. *Psychonomic Bulletin & Review*, *16*(3), 573–577. https://doi.org/10.1016/j.cognition.2004.06.005

Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, *125*(3), 452–465. https://doi.org/10.1016/j.cognition.2012.07.010

Cytowic, R. E. (1989). *Synaesthesia: A Union of the Senses*. New York: Springer Verlag.

Cytowic, R. E., & Wood, F. B. (1982). Synesthesia. II. Psychophysical relations in the synesthesia of geometrically shaped taste and colored hearing. *Brain and Cognition*, *1*(1), 36–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6927553

Davies, S. K., Izura, C., Socas, R., & Dominguez, A. (2015). Age of acquisition and imageability norms for base and morphologically complex words in English and in Spanish. *Behavior Research Methods*. https://doi.org/10.3758/s13428-015-

0579-y

Davis, B. J., Evans, M. A., & Reynolds, K. P. (2010). Child miscues and parental feedback during shared alphabet book reading and relations with child literacy skills. *Scientific Studies of Reading*, *14*(4), 341–364. https://doi.org/10.1080/10888431003623504

Davis, C. J. (2005). N-watch: a program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*(1), 65–70. https://doi.org/10.3758/BF03206399

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713–58. https://doi.org/10.1037/a0019738

Deroy, O., & Spence, C. (2013). Why we are not all synesthetes (not even weakly so). *Psychonomic Bulletin & Review*, *20*(4), 643–664. https://doi.org/10.3758/s13423-013-0387-2

Deroy, O., & Spence, C. (2015). Lessons of synaesthesia for consciousness: Learning from the exception, rather than the general. *Neuropsychologia*, NSYD1500328. https://doi.org/10.1016/j.neuropsychologia.2015.08.005

DiLorenzo, K. E., Rody, C. a., Bucholz, J. L., & Brady, M. P. (2011). Teaching Letter–Sound Connections With Picture Mnemonics: Itchy's Alphabet and Early Decoding. *Preventing School Failure: Alternative Education for Children and Youth*, *55*(1), 28–34. https://doi.org/10.1080/10459880903286763

Dixon, M. J., Smilek, D., Cudahy, C., & Merikle, P. M. (2000). Five plus two equals yellow. *Nature*, *406*(6794), 365. https://doi.org/10.1038/35019148

Dixon, M. J., Smilek, D., Duffy, P., Zanna, M., & Merikle, P. (2006). The role of meaning in grapheme-colour synaesthesia. *Cortex*, *42*, 243–252.

Dixon, M. J., Smilek, D., & Merikle, P. M. (2004). Not all synaesthetes are created equal: projector versus associator synaesthetes. *Cognitive, Affective & Behavioral Neuroscience*, *4*(3), 335–43.

Duñabeitia, J. A., & Carreiras, M. (2011). The relative position priming effect depends on whether letters are vowels or consonants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1143–63. https://doi.org/10.1037/a0023577

Duñabeitia, J. A., Laka, I., Perea, M., & Carreiras, M. (2009). Is Milkman a superhero like Batman? Constituent morphological priming in compound words. *European Journal of Cognitive Psychology*, *21*(4), 615–640.

Duñabeitia, J. A., Perea, M., & Carreiras, M. (2007). The role of the frequency of constituents in compound words: Evidence from Basque and Spanish. *Psychonomic Bulletin & Review*, *14*(6), 1171–1176. https://doi.org/10.3758/BF03193108

Eagleman, D. M. (2009). The objectification of overlearned sequences: A new view of spatial sequence synesthesia. *Cortex*, *45*(10), 1266–1277. https://doi.org/10.1016/j.cortex.2009.06.012

Eagleman, D. M., Kagan, A. D., Nelson, S. S., Sagaram, D., & Sarma, A. K. (2007). A standardized test battery for the study of synesthesia. *Journal of Neuroscience Methods*, *159*(1), 139–45.

https://doi.org/10.1016/j.jneumeth.2006.07.012

Ehri, L. C. (1987). Learning to read and spell words. *Journal of Literacy Research*, *19*(1), 5–31. https://doi.org/10.1080/10862968709547585

Ehri, L. C. (2005). Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading*, *9*(2), 167–188. https://doi.org/10.1207/s1532799xssr0902

Ehri, L. C., & Wilce, L. S. (1983). Development of word identification speed in skilled and less skilled beginning readers. *Journal of Educational Psychology*, *75*(1), 3–18. https://doi.org/10.1037/0022-0663.75.1.3

Ehri, L. C., & Wilce, L. S. (1985). Movement into Reading: Is the First Stage of Printed Word Learning Visual or Phonetic? *Reading Research Quarterly*, *20*(2), 163–179.

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *2*, 143–188.

Emrich, H. M., Schneider, U., & Zedler, M. (2002). *Welche Farbe hat der Montag?* Stuttgart: Hirzel.

Evans, M. A., Saint-Aubin, J., & Landry, N. (2009). Letter names and alphabet book reading by senior kindergarteners: An eye movement study. *Child Development*, *80*(6), 1824–1841. https://doi.org/10.1111/j.1467-8624.2009.01370.x

Fiorentino, R., & Poeppel, D. (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes*, *22*(7), 953–1000. https://doi.org/10.1080/01690960701190215

Frisson, S., Niswander-Klement, E., & Pollatsek, A. (2008). The role of semantic transparency in the processing of English compound words. *British Journal of Psychology*, *99*, 87–107. https://doi.org/10.1348/000712607X181304

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, *12*(4), 395–427. https://doi.org/10.3758/BF03201693

Giraudo, H., & Grainger, J. (2001). Priming complex words : Evidence for supralexical representation of morphology. *Psychonomic Bulletin & Review*, *8*(1), 127–131.

Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: A model of letter position coding. *Psychological Review*, *115*(3), 577–600. https://doi.org/10.1037/a0012667

Goodhew, S. C., & Kidd, E. (2017). Language use statistics and prototypical grapheme colours predict synaesthetes' and non-synaesthetes' word-colour associations. *Acta Psychologica*, *173*, 73–86. https://doi.org/10.1016/j.actpsy.2016.12.008

Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data Collection in a Flat World: The Strengths and Weaknesses of Mechanical Turk Samples. *Journal of Behavioral Decision Making*, *26*(3), 213–224. https://doi.org/10.1002/bdm.1753

Grainger, J., Bertrand, D., Lété, B., Beyersmann, E., & Ziegler, J. C. (2016). A developmental investigation of the first-letter advantage. *Journal of*

Experimental Child Psychology, *152*, 161–172. https://doi.org/10.1016/j.jecp.2016.07.016

Gray, J. A., Chopping, S., Nunn, J., Parslow, D., Gregory, L., Williams, S., … Baron-Cohen, S. (2002). Implications of Synaesthesia for Functionalism. *Journal of Consciousness Studies*, *9*(12), 5–31.

Gray, J. A., Parslow, D., Brammer, M., Chopping, S., Vythelingum, G., & Fytche, D. (2006). Evidence Against Functionalism from Neuroimaging of the Alien Colour Effect in Synaesthesia. *Cortex*, *42*(2), 309–318. https://doi.org/10.1016/S0010-9452(08)70357-5

Hamada, D., Yamamoto, H., & Saiki, J. (2017). Multilevel analysis of individual differences in regularities of grapheme–color associations in synesthesia. *Consciousness and Cognition*, *53*, 122–135. https://doi.org/10.1016/j.concog.2017.05.007

Hancock, P. (2013). Synesthesia, alphabet books, and fridge magnets. In J. Simner & E. M. Hubbard (Eds.), *The Oxford Handbook of Synesthesia* (pp. 83–102). Oxford University Press.

Havlik, A. M., Carmichael, D. A., & Simner, J. (2015). Do sequence-space synaesthetes have better spatial imagery skills? Yes, but there are individual differences. *Cognitive Processing*. https://doi.org/10.1007/s10339-015-0657-1

Heyer, V., & Kornishova, D. (2017). Semantic transparency affects morphological priming … eventually. *The Quarterly Journal of Experimental Psychology*, 1–42. https://doi.org/10.1080/17470218.2017.1310915

Hubbard, E. M., Ambrosio, P., Azoulai, S., & Ramachandran, V. S. (2005). Grapheme and letter name based patterns in synesthetic colors. In *Cognitive Neuroscience Society Meeting*.

Hubbard, E. M., & Ramachandran, V. S. (2003). The Phenomenology of Synaesthesia. *Journal of Consiousness Studies*, *10*(8), 49–57.

Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron*, *48*(3), 509–20. https://doi.org/10.1016/j.neuron.2005.10.012

Hung, W.-Y., Simner, J., Shillcock, R., & Eagleman, D. M. (2014). Synaesthesia in Chinese characters: the role of radical function and position. *Consciousness and Cognition*, *24*, 38–48. https://doi.org/10.1016/j.concog.2013.12.009

Inhoff, A. W., Starr, M. S., Solomon, M., & Placke, L. (2008). Eye movements during the reading of compound words and the influence of lexeme meaning. *Memory & Cognition*, *36*(3), 675–687. https://doi.org/10.3758/MC.36.3.675

Janik-McErlean, A. B., & Banissy, M. J. (2016). Examining the Relationship Between Schizotypy and Self-Reported Visual Imagery Vividness in Grapheme-Color Synaesthesia. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00131

Janssen, N., Pajtas, P. E., & Caramazza, A. (2011). A set of 150 pictures with morphologically complex English compound names: norms for name agreement, familiarity, image agreement, and visual complexity. *Behavior Research Methods*, *43*(2), 478–90. https://doi.org/10.3758/s13428-011-0065-0

Jensen, A. R., & Rohwer, W. D. (1966). The stroop color-word test: A review. *Acta Psychologica*, *25*, 36–93. https://doi.org/10.1016/0001-6918(66)90004-7

Ji, H., Gagné, C. L., & Spalding, T. L. (2011). Benefits and costs of lexical decomposition and semantic integration during the processing of transparent and opaque English compounds. *Journal of Memory and Language*, *65*(4), 406–430. https://doi.org/10.1016/j.jml.2011.07.003

Johnson, R. L., & Eisler, M. E. (2012). The importance of the first and last letter in words during sentence reading. *Acta Psychologica*, *141*(3), 336–351. https://doi.org/10.1016/j.actpsy.2012.09.013

Jonas, C. N. (2010). *Effects of synaesthetic colour and space on cognition*. University of Sussex.

Jonas, C. N., Spiller, M. J., Jansari, A., & Ward, J. (2014). Comparing implicit and synaesthetic number-space associations: Visuospatial and verbal spatial-numerical associations of response codes. *Quarterly Journal of Experimental Psychology (2006)*, *67*(7), 1262–73. https://doi.org/10.1080/17470218.2013.856928

Juhasz, B. J. (2007). The influence of semantic transparency on eye movements during English compound word recognition. In *Eye movements: A window on mind and brain* (pp. 373–389).

Juhasz, B. J., Starr, M. S., Inhoff, A. W., & Placke, L. (2003). The effects of morphology on the processing of compound words : Evidence from naming , lexical decisions and eye xations. *British Journal of Psychology*, *94*, 223–244.

Klein, G. S. (1964). Semantic Power Measured through the Interference of Words with Color-Naming. *Source: The American Journal of Psychology*, *77*(4), 576–588. Retrieved from http://www.jstor.org/stable/1420768

Kubitza, A. (2006). Word frequency effect for synaesthetic word colour in German NN-compounds (case study). In *Paper presented at the UKSA Meeting*. London.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes*, *23*(7–8), 1089–1132. https://doi.org/10.1080/01690960802193688

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 876.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). lmerTest: Tests in Linear Mixed Effects Models. R package version 2.0-33. Retrieved from https://cran.r-project.org/package=lmerTest

Lacey, S., Martinez, M., McCormick, K., & Sathian, K. (2016). Synesthesia strengthens sound-symbolic cross-modal correspondences. *European Journal of Neuroscience*, *44*(9), 2716–2721. https://doi.org/10.1111/ejn.13381

Laeng, B., Hugdahl, K., & Specht, K. (2011). The neural correlate of colour distances revealed with competing synaesthetic and real colours. *CORTEX*, *47*, 320–331.

https://doi.org/10.1016/j.cortex.2009.09.004

Lavrynenko, O. (2014). A study of the cognitive synesthetic mechanism as a supplement to the sound imitation theory of lingual genesis. *East Eur. J. Psycholinguist*, *1*, 62–70.

Lee, H.-W., Rayner, K., & Pollatsek, A. (2002). The processing of consonants and vowels in reading: Evidence from the fast priming paradigm. *Psychonomic Bulletin & Review*, *9*(4), 766–772. https://doi.org/10.3758/BF03196333

Libben, G. (1998). Semantic transparency in the processing of compounds: consequences for representation, processing, and impairment. *Brain and Language*, *61*, 30–44. https://doi.org/10.1006/brln.1997.1876

Libben, G. (2006). Why Study Compound Processing? An overview of the issues. In G. Libben & G. Jarema (Eds.), *The Representation and Processing of Compound Words* (pp. 1–22). Oxford, UK: OUP Oxford.

Libben, G. (2010). Compound words, semantic transparency, and morphological transcendence. *Linguistische Berichte Sonderheft*.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: the role of semantic transparency and morphological headedness. *Brain and Language*, *84*(1), 50–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12537951

Libben, G., & Jarema, G. (2006). *The representation and processing of compound words*. (G. Libben & G. Jarema, Eds.). Oxford: Oxford University Press.

Lorenz, A., & Zwitserlood, P. (2014). Processing of nominal compounds and gender-marked determiners in aphasia: evidence from German. *Cognitive Neuropsychology*, *31*(1–2), 40–74. https://doi.org/10.1080/02643294.2013.874335

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*, 203–208.

MacGregor, L. J., & Shtyrov, Y. (2013). Multiple routes for compound word processing in the brain: evidence from EEG. *Brain and Language*, *126*(2), 217–29. https://doi.org/10.1016/j.bandl.2013.04.002

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin*, *109*(2), 163–203. https://doi.org/Doi 10.1037//0033-2909.109.2.163

Mankin, J. L. (2017). The psycholinguistics of synaesthesia: What can we learn from studying synaesthetes? *PsyPAG Quarterly*, *6*.

Mankin, J. L., Thompson, C., Branigan, H. P., & Simner, J. (2016). Processing compound words: Evidence from synaesthesia. *Cognition*, *150*, 1–9. https://doi.org/10.1016/j.cognition.2016.01.007

Mannaert, L. N. H., Dijkstra, K., & Zwaan, R. A. (2017). Is color an integral part of a rich mental simulation? *Memory & Cognition*.

Marantz, A. (2013). No escape from morphemes in morphological processing. *Language and Cognitive Processes*, *28*(7), 905–916.

https://doi.org/10.1080/01690965.2013.779385

Marelli, M., Aggujaro, S., Molteni, F., & Luzzatti, C. (2012). The multiple-lemma representation of Italian compound nouns: a single case study of deep dyslexia. *Neuropsychologia*, *50*(5), 852–61. https://doi.org/10.1016/j.neuropsychologia.2012.01.021

Marelli, M., & Luzzatti, C. (2012). Frequency effects in the processing of Italian nominal compounds: Modulation of headedness and semantic transparency. *Journal of Memory and Language*, *66*(4), 644–664. https://doi.org/10.1016/j.jml.2012.01.003

Marks, L. E. (1975). On colored-hearing synesthesia: Cross-modal translations of sensory dimensions. *Psychological Bulletin*, *82*(3), 303–331.

Masterson, J., & Druks, J. (1998). Description of a set of 164 nounsand 102 verbs matched for printed word frequency, familiarityand age-of-acquisition. *Journal of Neurolinguistics*, *11*(4), 331–354. https://doi.org/10.1016/S0911-6044(98)00023-2

Masterson, J., Druks, J., & Gallienne, D. (2008). Object and action picture naming in three- and five-year-old children. *Journal of Child Language*, *35*(2), 373–402. https://doi.org/10.1017/S0305000907008549

Mattingley, J. B. (2009). Attention, automaticity, and awareness in synesthesia. *Annals of the New York Academy of Sciences*, *1156*, 141–67. https://doi.org/10.1111/j.1749-6632.2009.04422.x

Mattingley, J. B., Rich, A. N., Yelland, G., & Bradshaw, J. L. (2001). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature*, *410*(6828), 580–2. https://doi.org/10.1038/35069062

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Maxwell, S. E. (2004). The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychological Methods*, *9*(2), 147–163. https://doi.org/10.1037/1082-989X.9.2.147

McDonough, C., Song, L., Hirsh-Pasek, K., Golinkoff, R. M., & Lannon, R. (2011). An image is worth a thousand words: why nouns tend to dominate verbs in early word learning. *Developmental Science*, *14*(2), 181–9. https://doi.org/10.1111/j.1467-7687.2010.00968.x

Mealor, A., Simner, J., Rothen, N., Carmichael, D. A., & Ward, J. (2015). Different Dimensions of Cognitive Style in Typical and Atypical Cognition: New Evidence and a New Measurement Tool. *PloS One*.

Meier, B., & Rothen, N. (2013). Grapheme-color synaesthesia is associated with a distinct cognitive style. *Frontiers in Psychology*, *4*, 632. https://doi.org/10.3389/fpsyg.2013.00632

Mencken, H. L. (1917, November 16). The Divine Afflatus. *New York Evening Mail*.

Mills, C. B. (1999). Digit Synaesthesia: A Case Study Using A Stroop-Type Test. *Cognitive Neuropsychology*, *16*(2), 181–191. https://doi.org/10.1080/026432999380951

Mills, C. B., Viguers, M. L., Edelson, S. K., Thomas, A. T., Simon-Dack, S. L., & Innis, J. a. (2002). The color of two alphabets for a multilingual synesthete. *Perception*, *31*(11), 1371–1394. https://doi.org/10.1068/p3429

Mokrzycki, W. S., & Tatol, M. (2012). Color difference ΔE - A survey. *Machine Graphics and Vision*, *8*. Retrieved from http://www.researchgate.net/publication/236023905_Color_difference_Delta_E_-_A_survey

Morrison, C. M., & Ellis, A. W. (1995). Roles of Word-Frequency and Age of Acquisition in Word Naming and Lexical Decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 116–133. https://doi.org/10.1037//0278-7393.21.1.116

Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology*, *21*(1), 167–180.

Mroczko, A., & Metzinger, T. (2009). Immediate transfer of synesthesia to a novel inducer. *Journal of Vision*, *9*(12:25), 1–8. https://doi.org/10.1167/9.12.25.Introduction

Mulvenna, C., & Hubbard, E. (2003). The relationship between synaesthesia and creativity (Vol. 8, p. 2001).

Murray, B. A., Stahl, S. A., & Ivey, M. G. (1996). Developing phoneme awareness through alphabet books. *Reading and Writing*, *8*(4), 307–322. https://doi.org/10.1007/BF00395111

Myles, K. M., Dixon, M. J., Smilek, D., & Merikle, P. M. (2003). Seeing double: The role of meaning in alphanumeric-colour synaesthesia. *Brain and Cognition*, *53*(2), 342–345. https://doi.org/10.1016/S0278-2626(03)00139-8

Nagai, J., Yokosawa, K., & Asano, M. (2015). Biases and regularities of grapheme–colour associations in Japanese nonsynaesthetic population. *Quarterly Journal of Experimental Psychology (2006)*, *69*(1), 1–13. https://doi.org/10.1080/17470218.2015.1018835

Nikolić, D. (2009). Is synaesthesia actually ideaesthesia? An inquiry into the nature of the phenomenon. *Proceedings of the Third International Congress on Synaesthesia, Science & Art*, 26–29.

Nodelman, P. (2001). A is for... What? The Function of Alphabet Books. *Journal of Early Childhood Literacy*, *1*(3), 235–253. https://doi.org/10.1177/14687984010013001

Novich, S., Cheng, S., & Eagleman, D. M. (2011). Is synaesthesia one condition or many? A large-scale analysis reveals subgroups. *Journal of Neuropsychology*, *5*(2), 353–71. https://doi.org/10.1111/j.1748-6653.2011.02015.x

Nowak, S. N. (2015). *ABCs by Design: The Role of Alphabet Book Design and Children's Alphabetic Behaviours in Emergent Literacy Skill Acquisition*. Unversity of Guelph, ON, Canada. https://doi.org/10.1017/CBO9781107415324.004

Nowak, S. N., & Evans, M. A. (2012). Parents' goals for and perceptions of alphabet books. *Reading and Writing*, *26*(8), 1265–1287. https://doi.org/10.1007/s11145-012-9417-0

Nunn, J. a, Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J., … Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nature Neuroscience*, *5*(4), 371–5. https://doi.org/10.1038/nn818

Oldfield, R. C., & Wingfield, a. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*(4), 273–281. https://doi.org/10.1080/17470216508416445

Paivio, A. (1969). Mental Imagery in Associative Learning and Memory. *Psychological Review*, *76*(3).

Paivio, A. (1971). Imagery and Language. In *Imagery: Current Cognitive Approaches* (pp. 7–32). New York: Academic Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal Of Psychology*, *45*(3), 255–287. https://doi.org/10.1037/h0084295

Paivio, A., Walsh, M., & Bons, T. (1994). Concreteness effects on memory: When and why? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1196–1204. https://doi.org/10.1037/0278-7393.20.5.1196

Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, Imagery, and Meaningfulness Values for 925 Nouns. *Journal of Experimental Psychology*, *76*(1), 1–25. https://doi.org/10.1037/h0025327

Perea, M., & Lupker, S. J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory & Cognition*, *31*(6), 829–841. https://doi.org/10.3758/BF03196438

Pfeifer, G., Rothen, N., Ward, J., Chan, D., & Sigala, N. (2014). Associative memory advantage in grapheme-colour synaesthetes compared to older, but not younger adults. *Frontiers in Psychology*, *5*(696). https://doi.org/10.3389/fpsyg.2014.00696

Pinker, S. (1991). Rules of language. *Science*, *253*, 530–535.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it? *Cognition*, *95*(2), 201–36. https://doi.org/10.1016/j.cognition.2004.08.004

Pollatsek, A., & Hyönä, J. (2005). The role of semantic transparency in the processing of Finnish compound words. *Language and Cognitive Processes2*, *20*(1–2), 261–290. https://doi.org/10.1080/01690960444000098

Pollatsek, A., Hyönä, J., & Bertram, R. (2000). The Role of Morphological Constituents in Reading Finnish Compound Words. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 820–833.

Prado, E. L., & Ullman, M. T. (2009). Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(4), 849–866. https://doi.org/10.1037/a0015286

Price, M. C. (2009). Spatial forms and mental imagery. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *45*(10), 1229–45. https://doi.org/10.1016/j.cortex.2009.06.013

Prinzmetal, W., Hoffman, H., & Vest, K. (1991). Automatic processes in word perception: an analysis from illusory conjunctions. *Journal of Experimental*

Psychology. Human Perception and Performance, *17*(4), 902–923. https://doi.org/10.1037/0096-1523.17.4.902

Proctor, R. W. (1978). Sources of color-word interference in the Stroop color-naming task. *Perception & Psychophysics*, *23*(5), 413–419.

R Core Team. (2016). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Ramachandran, V. S., & Hubbard, E. M. (2001a). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings. Biological Sciences / The Royal Society*, *268*(1470), 979–83. https://doi.org/10.1098/rspb.2000.1576

Ramachandran, V. S., & Hubbard, E. M. (2001b). Synaesthesia--A Window Into Perception, Thought and Language. *Journal of Consciousness Studies*, *8*(12), 3–34.

Rayner, K. (1979). Eye guidance in reading: fixation locations within words. *Perception*, *8*, 21–30.

Rayner, K., White, S. J., Johnson, R. L., & Liversedge, S. P. (2006). Raeding wrods with jubmled lettres: There is a cost. *Psychological Science*, *17*(3), 192–193. https://doi.org/10.1111/j.1467-9280.2006.01684.x

Rich, A. N., Bradshaw, J. L., & Mattingley, J. B. (2005). A systematic, large-scale study of synaesthesia: implications for the role of early experience in lexical-colour associations. *Cognition*, *98*(1), 53–84. https://doi.org/10.1016/j.cognition.2004.11.003

Richter, T., & Zwaan, R. A. Processing of color words activates color representations, 111 Cognition § (2009). https://doi.org/10.1016/j.cognition.2009.02.011

Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin and Review*, *13*(2), 310–315. https://doi.org/10.3758/BF03193849

Roach, P. (2009). *English Phonetics and Phonology* (4th ed.). Cambridge University Press.

Roelofs, A. (2003). Goal-Referenced Selection of Verbal Action: Modeling Attentional Control in the Stroop Task. *Psychological Review*, *110*(1), 88–125. https://doi.org/10.1037/0033-295X.110.1.88

Rothen, N., & Meier, B. (2010). Grapheme-colour synaesthesia yields an ordinary rather than extraordinary memory advantage: evidence from a group study. *Memory (Hove, England)*, *18*(3), 258–64. https://doi.org/10.1080/09658210903527308

Rothen, N., Meier, B., & Ward, J. (2012). Enhanced memory ability: Insights from synaesthesia. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1952–63. https://doi.org/10.1016/j.neubiorev.2012.05.004

Rothen, N., Seth, A. K., Witzel, C., & Ward, J. (2013). Diagnosing synaesthesia with online colour pickers: maximising sensitivity and specificity. *Journal of Neuroscience Methods*, *215*(1), 156–60. https://doi.org/10.1016/j.jneumeth.2013.02.009

Rouw, R., Case, L., Gosavi, R., & Ramachandran, V. S. (2014). Color associations for days and letters across different languages. *Frontiers in Psychology*, *5*(MAY), 369. https://doi.org/10.3389/fpsyg.2014.00369

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, *10*(6), 792–7. https://doi.org/10.1038/nn1906

Rouw, R., & Scholte, H. S. (2010). Neural basis of individual differences in synesthetic experiences. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *30*(18), 6205–13. https://doi.org/10.1523/JNEUROSCI.3444-09.2010

Rouw, R., & Scholte, H. S. (2016). Personality and cognitive profiles of a general synesthetic trait. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2016.01.006

Rouw, R., Scholte, H. S., & Colizoli, O. (2011). Brain areas involved in synaesthesia: a review. *Journal of Neuropsychology*, *5*(2), 214–42. https://doi.org/10.1111/j.1748-6653.2011.02006.x

Rouw, R., van Driel, J., Knip, K., & Ridderinkhof, K. R. (2013). Executive functions in synesthesia. *Consciousness and Cognition*, *22*(1), 184–202. https://doi.org/10.1016/j.concog.2012.11.008

Sabsevitz, D. S., Medler, D. A., Seidenberg, M., & Binder, J. R. (2005). Modulation of the semantic system by word imageability. *NeuroImage*, *27*(1), 188–200. https://doi.org/10.1016/j.neuroimage.2005.04.012

Sagiv, N., & Ward, J. (2006). Crossmodal interactions: lessons from synesthesia. *Progress in Brain Research*, *155*(2001), 259–71. https://doi.org/10.1016/S0079-6123(06)55015-0

Sandra, D. (1990). On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *42*(3), 529–567. https://doi.org/10.1080/14640749008401236

Sawilowsky, S. S. (2009). New Effect Size Rules of Thumb. *Journal of Modern Applied Statistical Methods*, *8*(2), 597–599. Retrieved from http://digitalcommons.wayne.edu/coe_tbf/4

Scaltritti, M., & Balota, D. A. (2013). Are all letters really processed equally and in parallel? Further evidence of a robust first letter advantage. *Acta Psychologica*, *144*(2), 397–410. https://doi.org/10.1016/j.actpsy.2013.07.018

Schanda, J. (Ed.). (2007). *Colorimetry: understanding the CIE system*. John Wiley & Sons.

Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology*, *59*(2), 132–138. https://doi.org/10.1037/h0087468

Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*(3), 333–367. https://doi.org/10.1080/01690960344000198

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B.

Feldman (Ed.), *Morphological Aspects of Language Processing* (pp. 131–154). Hillsdale, New Jersey: Lawrence Erlbaum.

Schwanenflugel, P. J., & Akin, C. E. (1994). Developmental Trends in Lexical Decisions for Abstract and Concrete Words. *Reading Research Quarterly1*, *29*(3), 250–264.

Shillcock, R., Ellison, T. M., & Monaghan, P. (2000). Eye-fixation behavior, lexical storage, and visual word recognition in a split processing model. *Psychological Review*, *107*(4), 824–851. https://doi.org/10.1037/0033-295X.107.4.824

Shin, E., & Kim, C.-Y. (2014). Both "나" and "な" are yellow: Cross-linguistic investigation in search of the determinants of synesthetic color. *Neuropsychologia*, 1–12. https://doi.org/10.1016/j.neuropsychologia.2014.09.032

Siegler, R., Eisenberg, N., DeLoache, J., & Saffran, J. (2014). *How children develop* (Fourth). New York, NY: Worth Publishers.

Simmons, W. K., Ramjee, V., Beauchamp, M. S., McRae, K., Martin, A., & Barsalou, L. W. (2007). A common neural substrate for perceiving and knowing about color. *Neuropsychologia*, *45*(12), 2802–2810. https://doi.org/10.1016/j.neuropsychologia.2007.05.002

Simner, J. (2007). Beyond perception: synaesthesia as a psycholinguistic phenomenon. *Trends in Cognitive Sciences*, *11*(1), 23–9. https://doi.org/10.1016/j.tics.2006.10.010

Simner, J. (2012a). Defining synaesthesia: a response to two excellent commentaries. *British Journal of Psychology (London, England : 1953)*, *103*(1), 24–7. https://doi.org/10.1111/j.2044-8295.2011.02059.x

Simner, J. (2012b). Defining synaesthesia. *British Journal of Psychology (London, England : 1953)*, *103*(1), 1–15. https://doi.org/10.1348/000712610X528305

Simner, J. (2013). The "rules" of synaesthesia. In J. Simner & E. M. Hubbard (Eds.), *The Oxford Handbook of Synesthesia* (pp. 149–164). Oxford, UK: Oxford University Press.

Simner, J., & Bain, A. E. (2013). A longitudinal study of grapheme-color synesthesia in childhood: 6/7 years to 10/11 years. *Frontiers in Human Neuroscience*, *7*, 603. https://doi.org/10.3389/fnhum.2013.00603

Simner, J., & Bain, A. E. (2017). Do children with grapheme-colour synaesthesia show cognitive benefits? *British Journal of Psychology*. https://doi.org/10.1111/bjop.12248

Simner, J., & Carmichael, D. A. (2015). Is synaesthesia a dominantly female trait? *Cognitive Neuroscience*, 1–9. https://doi.org/10.1080/17588928.2015.1019441

Simner, J., Glover, L., & Mowat, A. (2006). Linguistic determinants of word colouring in grapheme-colour synaesthesia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *42*(2), 281–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16683502

Simner, J., Harrold, J., Creed, H., Monro, L., & Foulkes, L. (2009). Early detection of markers for synaesthesia in childhood populations. *Brain : A Journal of Neurology*, *132*(Pt 1), 57–64. https://doi.org/10.1093/brain/awn292

Simner, J., & Holenstein, E. (2007). Ordinal linguistic personification as a variant of synesthesia. *Journal of Cognitive Neuroscience*, *19*(4), 694–703. https://doi.org/10.1162/jocn.2007.19.4.694

Simner, J., & Ludwig, V. U. (2012). The color of touch: A case of tactile–visual synaesthesia. *Neurocase*, *18*(2), 167–180. https://doi.org/10.1080/13554794.2011.568503

Simner, J., Mayo, N., & Spiller, M.-J. (2009). A foundation for savantism? Visuo-spatial synaesthetes present with cognitive benefits. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *45*(10), 1246–60. https://doi.org/10.1016/j.cortex.2009.07.007

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. a, Fraser, C., … Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, *35*(8), 1024–1033. https://doi.org/10.1068/p5469

Simner, J., Treffert, D. A., Hughes, J. E. A., Baron-Cohen, S., & Ward, J. (2017). Is Synaesthesia More Prevalent in Autism Spectrum Conditions? Only Where There Is Prodigious Talent. *Multisensory Research*. https://doi.org/10.1163/22134808-00002558

Simner, J., & Ward, J. (2008). Synaesthesia, color terms, and color space: Color claims came from color names in Beeli, Esslen, and Jäncke (2007). *Psychological Science*, *19*(4), 412–414. https://doi.org/10.1111/j.1467-9280.2008.02101.x.General

Simner, J., Ward, J., Lanz, M., Jansari, A., Noonan, K., Glover, L., & Oakley, D. a. (2005). Non-random associations of graphemes to colours in synaesthetic and non-synaesthetic populations. *Cognitive Neuropsychology*, *22*(8), 1069–85. https://doi.org/10.1080/02643290500200122

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*(9), 793–5. https://doi.org/10.1111/j.1467-9280.2007.01981.x

Smilek, D., & Dixon, M. J. (2002). Towards a synergistic understanding of synaesthesia: Combining current experimental findings with synaesthetes' subjective descriptions. *Psyche*, *8*(1).

Smilek, D., Malcolmson, K. a, Carriere, J. S. a, Eller, M., Kwan, D., & Reynolds, M. (2007). When "3" is a jerk and "E" is a king: personifying inanimate objects in synesthesia. *Journal of Cognitive Neuroscience*, *19*(6), 981–992. https://doi.org/10.1162/jocn.2007.19.6.981

Specht, K., & Laeng, B. (2011). An independent component analysis of fMRI data of grapheme-colour synaesthesia. *Journal of Neuropsychology*, *5*(2), 203–213. https://doi.org/10.1111/j.1748-6653.2011.02008.x

Spector, F., & Maurer, D. (2008). The colour of Os: Naturally biased associations between shape and colour. *Perception*, *37*(6), 841–847. https://doi.org/10.1068/p5830

Spector, F., & Maurer, D. (2009). Synesthesia: a new approach to understanding the development of perception. *Developmental Psychology*, *45*(1), 175–89. https://doi.org/10.1037/a0014171

Spector, F., & Maurer, D. (2011). The colors of the alphabet: naturally-biased associations between shape and color. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(2), 484–95. https://doi.org/10.1037/a0021437

Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Attention, Perception & Psychophysics*, *73*(4), 971–95. https://doi.org/10.3758/s13414-010-0073-7

Sperling, J., Prvulovic, D., Linden, D. E., Singer, W., & Stirn, A. (2006). Neuronal correlates of colour-graphemic synaesthesia: A fMRI study. *Cortex*, *2*(42), 295–303.

Spiller, M. J., Jonas, C. N., Simner, J., & Jansari, A. (2015). Beyond visual imagery: How modality-specific is enhanced mental imagery in synesthesia? *Consciousness and Cognition*, *31*, 73–85. https://doi.org/10.1016/j.concog.2014.10.010

Stadthagen-Gonzale, H., & Davis, C. J. (2006). The Bristol norms for Age of Acquisition, Imageability, and Familiarity. *Behavior Research Methods*, *38*(4), 598–605. https://doi.org/10.1017/CBO9781107415324.004

Stockall, L., & Marantz, A. (2006). A single route, full decomposition model of morphological complexity: MEG evidence. *The Mental Lexicon*, *1*(1), 85–123. https://doi.org/10.1075/ml.1.1.07sto

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662.

Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, *7*(4), 263–272. https://doi.org/10.3758/BF03197599

Taft, M. (1988). A morphological-decomposition model of lexical representation. *Linguistics*, *26*, 657–667.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, *9*(3), 271–294. https://doi.org/10.1080/01690969408402120

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology, A: Human Experimental Psychology*, *57*(4), 745–765.

Taft, M., & Ardasinski, S. (2006). Obligatory decomposition in reading prefixed words. *The Mental Lexicon*, *1*(2), 183–199.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 638–647. https://doi.org/10.1016/S0022-5371(75)80051-X

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, *15*(6), 607–620. https://doi.org/10.1016/0022-5371(76)90054-2

Taft, M., & Nguyen-Hoan, M. (2010). A sticky stick? The locus of morphological representation in the lexicon. *Language and Cognitive Processes*, *25*(2), 277–296. https://doi.org/10.1080/01690960903043261

Taft, M., Xu, J., & Li, S. (2017). Letter coding in visual word recognition: The impact of embedded words. *Journal of Memory and Language*, *92*, 14–25. https://doi.org/10.1016/j.jml.2016.05.002

Taft, M., Zhu, X., & Peng, D. (1999). Positional Specificity of Radicals in Chinese Character Recognition. *Journal of Memory and Language*, *40*(4), 498–519. https://doi.org/10.1006/jmla.1998.2625

Teichmann, A. L., Nieuwenstein, M., & Rich, A. N. (2017). Digit-colour synaesthesia only enhances memory for colours in a specific context. *Journal of Experimental Psychology : Human Perception and Performance.*

Tremblay, A., & Ransijin, J. (2015). LMERConvenienceFunctions: Model Selection and Post-hoc Analysis for (G)LMER Models. Retrieved from https://cran.r-project.org/package=LMERConvenienceFunctions

Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 480–498. https://doi.org/10.1016/j.actpsy.2012.08.001

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6). https://doi.org/10.1080/17470218.2013.850521

Wan, X., Woods, A. T., van den Bosch, J., McKenzie, K. J., Velasco, C., & Spence, C. (2014). Cross-cultural differences in crossmodal correspondences between basic tastes and visual features. *Frontiers in Psychology*, *5*(December), 1–13. https://doi.org/10.3389/fpsyg.2014.01365

Ward, J. (2018). Meta-analysis of synaesthesia research. Manuscript in preparation.

Ward, J., Banissy, M. J., & Jonas, C. N. (2008). Haptic perception and synaesthesia. In M. Grunwald (Ed.), *Human Haptic Perception: Basics and Applications* (pp. 259–265). Basel: Birkhäuser Basel. https://doi.org/10.1007/978-3-7643-7612-3

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*(2), 264–280. https://doi.org/10.1016/S0010-9452(08)70352-6

Ward, J., Li, R., Salih, S., & Sagiv, N. (2007). Varieties of grapheme-colour synaesthesia: a new theory of phenomenological and behavioural differences. *Consciousness and Cognition*, *16*(4), 913–31. https://doi.org/10.1016/j.concog.2006.09.012

Ward, J., & Mattingley, J. B. (2006). Synaesthesia: an overview of contemporary findings and controversies. *Cortex*, *42*, 129–136.

Ward, J., & Simner, J. (2003). Lexical-gustatory synaesthesia: linguistic and conceptual factors. *Cognition*, *89*(3), 237–261. https://doi.org/10.1016/S0010-0277(03)00122-7

Ward, J., & Simner, J. (2005). Is synaesthesia an X-linked dominant trait with lethality in males? *Perception*, *34*(5), 611–623. https://doi.org/10.1068/p5250

Ward, J., Simner, J., & Auyeung, V. (2005). A comparison of lexical-gustatory and grapheme-colour synaesthesia. *Cognitive Neuropsychology*, *22*(1), 28–41. https://doi.org/10.1080/02643290442000022

Ward, J., Thompson-Lake, D., Ely, R., & Kaminski, F. (2008). Synaesthesia, creativity and art: what is the link? *British Journal of Psychology (London, England : 1953)*, *99*(Pt 1), 127–41. https://doi.org/10.1348/000712607X204164

Watson, M. R., Akins, K. A., & Enns, J. T. (2012). Second-order mappings in grapheme-color synesthesia. *Psychonomic Bulletin & Review*, *19*(2), 211–7. https://doi.org/10.3758/s13423-011-0208-4

Watson, M. R., Akins, K. A., Spiker, C., Crawford, L., & Enns, J. T. (2014). Synesthesia and learning: a critical review and novel theory. *Frontiers in Human Neuroscience*, *8*, 98. https://doi.org/10.3389/fnhum.2014.00098

Weiss, P. H., & Fink, G. R. (2009). Grapheme-colour synaesthetes show increased grey matter volumes of parietal and fusiform cortex. *Brain*, *132*(1), 65–70. https://doi.org/10.1093/brain/awn304

Weiss, P. H., Kalckert, A., & Fink, G. R. (2009). Priming letters by colors: evidence for the bidirectionality of grapheme-color synesthesia. *Journal of Cognitive Neuroscience*, *21*(10), 2019–26. https://doi.org/10.1162/jocn.2008.21166

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: the importance of word-beginning letters. *Journal of Experimental Psychology. Human Perception and Performance*, *34*(5), 1261–76. https://doi.org/10.1037/0096-1523.34.5.1261

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*(2), 221–243. https://doi.org/10.3758/BF03196158

Witthoft, N., & Winawer, J. (2006). Synesthetic colors determined by having colored refrigerator magnets in childhood. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, *42*(2), 175–183.

Witthoft, N., & Winawer, J. (2013). Learning, Memory, and Synesthesia. *Psychological Science*, *24*(3), 258–265. https://doi.org/10.1177/0956797612452573

Witthoft, N., Winawer, J., & Eagleman, D. M. (2015). Prevalence of Learned Grapheme-Color Pairings in a Large Online Sample of Synesthetes. *PLOS ONE*, *10*(3), e0118996. https://doi.org/10.1371/journal.pone.0118996

Worden, P., & Boettcher, W. (1990). Young children's acquisition of alphabet knowledge. *Journal of Literacy Research*, *22*(3), 277–295. https://doi.org/10.1080/10862969009547711

Xiao, X., Zhao, D., Zhang, Q., & Guo, C. (2012). Retrieval of concrete words involves more contextual information than abstract words: multiple components for the concreteness effect. *Brain and Language*, *120*(3), 251–8. https://doi.org/10.1016/j.bandl.2011.09.006

Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, Orthography, and Phonology Reading Chinese Compound Words. *Language and Cognitive Processes*, *14*(5–6), 525–565. https://doi.org/10.1080/016909699386185

Zwaan, R. A., & Pecher, D. (2012). Revisiting Mental Simulation in Language Comprehension: Six Replication Attempts. *PLoS ONE*, *7*(12), e51382. https://doi.org/10.1371/journal.pone.0051382

Zwitserlood, P. (1994). The role of semantic transparency in the processing and representation of Dutch compounds. *Language and Cognitive Processes*, *9*(3), 341–368. https://doi.org/10.1080/01690969408402123

Appendices

Appendix A

Items and frequency ratings for all 52 compound words and their first and second constituent morphemes, used in the analyses in Chapter 3. Frequency measures for all items are Zipf frequency ratings derived from the SUBTLEX-UK British English subtitle corpus (Van Heuven, Mandera, Keuleers, & Brysbaert, 2014). Note that four compound words had no rating in this corpus (*clothespin, firehose, lifevest, nailfile*) and therefore their frequency is listed as *NA*.

Compound		*1st Constituent Morpheme*		*2nd Constituent Morpheme*	
Word	**Frequency**	**Word**	**Frequency**	**Word**	**Frequency**
bagpipe	2.51	bag	4.89	pipe	4.26
bathtub	2.87	bath	4.65	tub	3.70
beehive	3.09	bee	4.19	hive	3.50
birdcage	2.56	bird	4.85	cage	4.03
birdhouse	1.84	bird	4.85	house	5.83
bobsled	2.28	bob	4.86	sled	3.09
bookshelf	2.60	book	5.21	shelf	4.02
bowtie	2.19	bow	4.22	tie	4.58
briefcase	3.14	brief	4.26	case	5.43
chessboard	2.44	chess	3.91	board	5.10
clothespin	*NA*	clothes	4.74	pin	4.26
crowbar	2.81	crow	3.65	bar	4.75
cupboard	4.27	cup	5.09	board	5.10
dartboard	2.79	dart	3.91	board	5.10
doughnut	3.50	dough	4.05	nut	3.99
drumstick	2.55	drum	4.33	stick	5.02
firehose	*NA*	fire	5.18	hose	3.55
fireplace	4.21	fire	5.18	place	5.78
fishbone	1.54	fish	5.19	bone	4.51
football	5.12	foot	4.92	ball	5.33
footprint	3.62	foot	4.92	print	4.29
hairbrush	2.84	hair	5.03	brush	4.29
handcuffs	3.29	hand	5.44	cuffs	3.02
headlight	2.49	head	5.61	light	5.28
headphones	3.54	head	5.61	phones	4.21
hourglass	2.74	hour	5.17	glass	4.92
keyboard	3.64	key	5.07	board	5.10
keyhole	3.01	key	5.07	hole	4.76
lamppost	2.98	lamp	4.09	post	4.87
lifevest	*NA*	life	5.81	vest	3.54
lightbulb	2.68	light	5.28	bulb	3.68
lighthouse	3.77	light	5.28	house	5.83
mailbox	2.64	mail	4.63	box	5.12
matchbox	3.11	match	5.12	box	5.12

mousetrap	2.81	mouse	4.41	trap	4.22
nailfile	*NA*	nail	4.18	file	4.08
necklace	3.80	neck	4.65	lace	3.73
notepad	2.55	note	4.62	pad	3.97
padlock	2.91	pad	3.97	lock	4.42
peanut	3.65	pea	3.91	nut	3.99
rainbow	4.18	rain	5.08	bow	4.22
raindrop	2.44	rain	5.08	drop	4.90
seagull	3.30	sea	5.20	gull	3.11
seahorse	2.80	sea	5.20	horse	4.99
snowflake	2.91	snow	4.79	flake	3.21
snowman	3.52	snow	4.79	man	5.86
suitcase	3.78	suit	4.59	case	5.43
teabag	2.68	tea	5.02	bag	4.89
teapot	3.78	tea	5.02	pot	4.76
toothbrush	3.48	tooth	4.09	brush	4.29
toothpaste	3.35	tooth	4.09	paste	3.92
wheelchair	4.01	wheel	4.51	chair	4.66

Appendix B

All items and psycholinguistic measures for Experiment 1 of Chapter 4. Colour terms (left columns) are paired in the same row with their orthographically and psycholinguistically matched control words (right columns). Age of acquisition (labelled "AoA") and imageability ("Img") ratings were obtained from two different norming studies (Bird et al., 2001; Gilhooly & Logie, 1980). Where these studies both gave a rating for a particular word, we used the mean of the two ratings, denoted below in *italics*. As described in Chapter 4, we obtained our own imageability rating for the control word *aztec*, but no AoA rating was available for this word, so it is given as *NA* below. Frequency measures for all items are Zipf frequency ratings derived from the SUBTLEX-UK British English subtitle corpus (Van Heuven et al., 2014). Items are sorted by the frequency of the colour terms.

Colour Terms				*Control Words*			
Word	**AoA**	**Img**	**Frequency**	**Word**	**AoA**	**Img**	**Frequency**
red	213	585	5.41	reed	369	520	3.79
black	208	589	5.26	blade	344	568	3.90
white	241	566	5.25	whip	397	579	4.08
green	225	609	5.22	greed	294	420	3.63
blue	206	569	5.16	blush	*381.5*	*564.5*	3.19
brown	243	564	5.01	broom	313	595	3.56
yellow	256	*593.5*	4.84	yelp	378	499	2.38
pink	277	577	4.80	pin	261	576	4.26
orange	203	626	4.64	organ	356	576	3.99
grey	*276*	487	4.51	grease	394	463	3.68
purple	338	584	4.29	purse	247	640	3.92
tan	425	503	4.06	tack	363	546	3.50
violet	344	560	3.70	violence	441	546	4.60
scarlet	409	587	3.56	scarf	269	610	3.76
indigo	542	469	3.25	inch	*340.5*	472	4.18
beige	452	476	3.10	beech	459	430	3.34
crimson	389	615	3.10	cricket	337	603	4.45
maroon	510	*504.5*	2.90	matron	541	495	3.19
mauve	482	500	2.71	mail	291	565	4.63
azure	549	452	2.38	aztec	*NA*	492	3.03

Appendix C

All items and psycholinguistic measures for Experiment 2 of Chapter 4. High-imageability target words (left columns) are paired in the same row with their orthographically and psycholinguistically matched low-imageability control words (right columns). Age of acquisition (labelled "AoA" below) and imageability ("Img") ratings were obtained from three norming studies (Bird et al., 2001; Gilhooly & Logie, 1980; Stadthagen-Gonzale & Davis, 2006) available through the NeighbourWatch psycholinguistics programme (Davis, 2005). Where more than one rating was available for a particular word, we used the mean of the ratings, denoted below in *italics*. No AoA rating was available for *boar*, so it is given as *NA* below. Frequency measures for all items are Zipf frequency ratings derived from the SUBTLEX-UK British English subtitle corpus (Van Heuven et al., 2014). Items are sorted alphabetically by the target words.

Target Words				*Control Words*			
Word	**AoA**	**Img**	**Frequency**	**Word**	**AoA**	**Img**	**Frequency**
boar	*NA*	524	3.57	bore	428	214	3.76
bone	261	625	4.51	bond	491	380	4.46
bull	286	639	4.28	bulk	407	372	3.81
chain	311	559	4.49	change	338	315	5.49
dirt	220	475	4.00	dirge	606	262	2.16
face	172	*600.5*	5.44	facet	589	277	2.71
feast	351	610	4.10	feat	531	228	3.65
fire	189	*638.5*	5.18	fine	321	*388.5*	5.44
flame	*267*	*632*	4.05	flank	586	275	3.17
hare	281	577	3.98	haze	407	360	3.16
hive	319	554	3.50	hint	*415*	*299*	3.95
iron	*298.5*	*618*	4.52	irony	606	293	3.73
mint	298	485	4.15	mind	336	*344.5*	5.47
pig	233	635	4.51	pick	*264.5*	331	5.16
prong	439	499	2.31	prow	583	307	2.52
rat	279	588	4.23	rate	456	311	4.85
seat	243	604	4.78	sect	611	250	2.98
stump	372	490	3.52	stuff	372	305	5.39
wax	331	547	3.87	wad	525	370	2.88
wine	403	624	4.84	whine	390	424	2.64

Appendix D

For each letter, the top three most common index words and the percentage agreement for those index words. Index words with higher than 20% agreement are *italicised.*

Letter	Index	Agreement (%)	Letter	Index	Agreement (%)
	apple	*84.13*		no	11.43
A	animal	1.90	**N**	night	6.67
	aardvark	0.95		nose	6.03
	banana	15.87		orange	16.51
B	boy	13.33	**O**	open	12.06
	bear	12.38		octopus	11.75
	cat	*49.21*		pie	3.17
C	car	10.48	**P**	pear	2.86
	cookie	2.86		penguin	2.86
	dog	*57.14*		*queen*	*31.11*
D	dad	4.76	**Q**	quiet	8.89
	door	2.54		question	7.30
	elephant	*46.35*		run	6.35
E	egg	8.57	**R**	rest	5.08
	eagle	4.76		rat	4.76
	fish	*25.4*		snake	8.25
F	fox	8.25	**S**	sister	5.71
	food	5.71		stop	3.81
	goat	9.84		time	8.57
G	girl	8.25	**T**	tiger	5.71
	giraffe	7.30		turtle	5.40
	hat	11.43		*umbrella*	*22.86*
H	house	6.67	**U**	under	16.19
	happy	6.35		up	5.40
	igloo	*20.95*		victory	11.75
I	ice	11.75	**V**	vendetta	8.89
	ice cream	5.71		violin	8.57
	jump	14.92		water	17.46
J	joke	9.21	**W**	wax	3.81
	jack	7.94		whale	3.81
	kite	*23.81*		*xylophone*	*51.43*
K	kangaroo	11.75	**X**	xray	27.94
	king	6.98		xerox	3.81
	love	*23.17*		*yellow*	*24.13*
L	lion	11.43	**Y**	yes	13.65
	laugh	3.49		you	8.25
	mom	10.79		*zebra*	*60.63*
M	man	8.89	**Z**	zoo	19.68
	monkey	8.25		zero	3.49

Appendix E

Summary of the dominant colours predicted for each letter using the letter → index word → colour route. For each letter, the dominant colour (defined as the colour within each letter with the highest colour score) is reported along with its colour score for non-synaesthetes (Experiment 3) and self-reported synaesthetes (Experiment 4). Where the dominant colour differs between the two groups, this is highlighted in the synaesthete column in *italics*.

	Non-synaesthetes		*Synaesthetes*	
	Dominant colour	**Colour score**	**Dominant colour**	**Colour score**
A	Red	76.61	Red	70.89
B	Yellow	15.93	Yellow	15.74
C	Black	18.82	Black	19.55
D	Brown	42.52	Brown	37.69
E	Grey	46.06	Grey	42.47
F	Grey	11.47	Grey	12.46
G	White	5.82	*Pink*	6.47
H	Black	4.49	*Brown*	4.66
I	White	33.03	White	29.47
J	Red	4.59	*Green*	5.06
K	Brown	10.26	*Red*	13.63
L	Red	19.12	Red	13.97
M	Brown	8.85	Brown	7.42
N	Black	12.21	Black	11.63
O	Orange	17.83	Orange	17.78
P	Black	2.56	Black	2.11
Q	Purple	10.57	Purple	12.31
R	Grey	4.33	Grey	5.28
S	Green	3.93	*Red*	3.91
T	Orange	5.2	*Green*	5.63
U	Black	17.78	Black	11.63
V	Brown	7.74	Brown	8.29
W	Blue	16.11	Blue	13.15
X	Black	26.46	Black	21.72
Y	Yellow	25.99	Yellow	26.43
Z	Black	45.13	Black	44.85