Emotion Word Processing in Aphasia

Catherine Mason (BAppSc (Speech Pathology))
Department of Cognitive Science
Macquarie University, Australia

A thesis submitted for the degree of Master of Research.

October 2019

Table of Contents

List of Appendices

Abstract

Nouns that label emotions, such as "happiness", "misery" and "enjoyment", have been shown to have a lexical processing advantage over other "non-emotion" abstract words in unimpaired speakers. However, it is unknown whether emotion words are more quickly and accurately processed by people with aphasia. This thesis explored this question by examining emotion word processing in nine participants with aphasia. Performance on emotion and non-emotion abstract words was compared on three tasks: list recall, lexical decision and single-word reading. A significant emotion word advantage was found, but limited to only some participants and tasks. The results are discussed with focus on the possible influences of bilingualism, severity and level of language breakdown on the participants' performance. Importantly, an emotion word advantage was evident when emotion and non-emotion words were matched for imageability, suggesting that the effect could not be explained by the heightened imageability ratings of emotion words. Possible alternate explanations for the emotion word processing advantage are explored, with a focus on context availability and valence.

Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

(Signed) Date: 7th October 2019

Acknowledgements

The completion of my studies was made possible through the invaluable support and assistance I have received from many sources. First and foremost, my thanks go to my supervisor Professor Lyndsey Nickels. My route to reaching this point has had a few stops and starts, unexpected sidesteps and meandering turns. My heartfelt and enormous thanks for your enduring support throughout this sometimes unconventional journey.

My thanks also go to my participants, BBK, DTR, MMR, JDC, SPN, TLK, OLG, JOG and MLK. I was touched by not only your willingness to participate in this study but also by your enthusiasm and support.

I would like to extend my gratitude to Kate Makin and Vanessa Arratia from the Speech Pathology Department at the Royal Rehabilitation Centre (Ryde) for their tremendous help in recruiting participants.

In completing this thesis, I received vital assistance forming experimental and analysis methods. My immense thanks go to Professor David Howard and Dr. Peter Humburg for sharing their extensive knowledge and insights to assist in the development of statistical approaches. My gratitude also goes to Dr. Lisi Beyersmann for her advice regarding the word lists used in the experimental tasks.

I would like to also extend my appreciation to my fellow Master's students for your camaraderie over the past year and to my family for your unending patience, encouragement and support. Special thanks go to my husband, Wilfrid Li, for his assistance in coding the PyschoPy experiments and for his amazing technical know-how.

1. Introduction

1.1: Aphasia, Word-Retrieval and Concreteness

Aphasia is an acquired language disorder that results from damage to the language networks of the brain with causes including cerebrovascular disease, brain tumour and traumatic brain injury. While exact incidence numbers are difficult to obtain (Code & Petheram, 2011), it is estimated that of the 56 000 strokes that occur in Australia each year (Stroke Foundation, 2017), approximately one third will result in aphasia (e.g., Ferreira, 2012). Communication is fundamental to the human experience: we communicate when we socialise, when we go about daily tasks, when we work and when we learn. Thus, aphasia can have a profound impact on the individual that may be felt across many facets of his or her life.

An integral first step in the treatment of aphasia is to identify the patterns of difficulty displayed by the person with aphasia. In the case of an individual who has impaired word-retrieval, this may involve determining the word types that present the greatest difficulty, for example: Does the individual display impaired production of verbs but not nouns? Are inanimate objects more difficult to name than animate objects? One of the parameters that is known to influence word retrieval is the "concreteness" of the concept to which a word refers. Studies with unimpaired populations have demonstrated that concrete nouns (physical entities/objects) are more accurately and quickly processed than abstract nouns (concepts without a physical form) (e.g., Christian, Bickley, Tarka & Clayton, 1978; James, 1975; Paivio, 1969; Paivio, Walsh & Bons, 1994). Effects of concreteness on word processing are also evident in people with aphasia. For example, Franklin, Howard and Patterson (1994) present the case of DRB, a man with aphasia who had difficulty in tasks that required a word to be linked with its meaning, for example word-to-picture matching and generation of word associations, despite intact access to its lexical form. Crucially, his performance was modulated by concreteness, with greater difficulty with abstract words than concrete words. In a later study, Franklin, Howard and Patterson (1995) investigated DRB’s word production, and again found an effect of concreteness with his word-retrieval difficulties being greater for abstract than concrete words. However, individual cases in which people with aphasia display a “reverse-concreteness effect” have also been reported, particularly in primary progressive aphasia. For example, Breedin, Saffran and Coslett (1994) described a case of a person with semantic variant primary progressive aphasia who was more accurate in providing definitions of abstract words than concrete words and also demonstrated an advantage for abstract words over concrete words on a word-to-picture matching task.

Research concerning the concreteness effect has usually treated abstract words as a monolithic group, bound together by what they lack - a physical form. However, some researchers have suggested that more attention needs to be given to how abstract words are encoded and whether there might be sub-categories of abstract words that may be more or less difficult to process (e.g., Ghio, Vaghi, Perani, & Tettamanti, 2016; Setti & Caramelli, 2005). One potential sub-category of abstract words that has received particular attention is “emotion words”, that is, words that describe emotion or feeling states. Currently, research suggests that unimpaired speakers process emotion words differently from abstract words that do not refer to emotions, with emotion words having a processing advantage (e.g., Altaribba & Bauer, 2004). However, little is known about the processing of emotion words by people with aphasia. Specifically, do people with aphasia show a processing advantage for emotion words compared with non-emotion abstract words as is reported for unimpaired speakers? This is the aim of the research presented here.

1.2: What are "Emotion Words"?

The term "emotion words" is subject to variable usage across different research traditions. Some authors use the term to include not only words that refer to emotions but also words with emotional connotations such as "freedom" (e.g., Kousta, Vinson & Vigliocco, 2009; Scott, O'Donnell, Leuthold & Sereno, 2009) or taboo words such as "damn" (e.g., Ayçiçegi-Dinn & Caldwell-Harris, 2009). In contrast, others have used a more restricted definition of "emotion words" to include only those words that specifically label emotions such as "love", "hate" or "embarrassment" (e.g., Altarriba & Bauer, 2004; Altarriba, Bauer & Benvenuto, 1999; Zhang, Teo & Wu, 2019). In this study, as the primary concern is to investigate retrieval of words that refer to specific emotions compared with other abstract words, this stricter definition of emotion words is applied as characterised by Knickerbocker, Johnson and Altarriba (2015): "Emotion words label a state of mind that can be experienced." (p. 787).

In order to understand why emotion word processing may differ from other abstract words it is necessary to first consider what makes concrete words "concrete" and abstract words "abstract", as well as the cognitive mechanisms that are believed to underpin the concreteness effect.

1.3: The Concreteness Effect

Crystal (2003) outlines the distinction between concrete and abstract nouns as the presence or absence of a physical form: "*concrete* [is] said of nouns which refer to physical entities [and] contrasts with *abstract* which applies to nouns lacking a physical reference" (p. 460). However, this division is not always clear-cut; while some words are always concrete and some are always abstract, there are a number of words to occupy a grey area in that they can be abstract or concrete depending on context. Consider, as an example, the word "property". If used in a sentence to refer to the notion of ownership e.g., *“They needed to protect their intellectual property”,* then it does not refer to a physical entity and is therefore abstract. However, in the sentence, *“They inspected the property on Saturday”,* the word is being used synonymously with “house” and thus is concrete. Given this ambiguity, it is important in linguistic research to have clear methods by which to quantify “concreteness”. This is usually achieved using ratings of concreteness (how concrete a concept is) and/or imageability (how easily a word conjures a mental image). Given that physical entities are typically easy to think of as images, unsurprisingly imageability and concreteness usually correlate, with high concreteness words also having high imageability, and low concreteness words having low imageability. However, emotion words are an exception to this rule. While emotion words are rated low in concreteness, they tend to have moderate to high imageability ratings (e.g., Altarriba et al., 1999; Dellantonio, Mulatti, Pastore & Job, 2014; Paivio, Yuille & Madigan, 1968). This mismatch may have implications for the lexical processing of emotion words as the correlation between imageability and concreteness underpins one of the dominant theories used to explain the concreteness effect: “Dual-Coding Theory” (e.g., Paivio, 1991, Paivio, Rogers & Smythe, 1968).

Dual-Coding Theory proposes that human cognition has two symbolic representational systems: a verbal system known as “logogens” which includes both written and spoken words, and a nonverbal system known as “imagens” (e.g., Paivio, 1991). According to Dual-Coding Theory, all words have a “logogen” encoding, however concrete words, by virtue of their high imageability, can also be represented nonverbally as visual imagery. The dual representations of concrete words are thought to have an additive effect on activation of the word-form, which in turn explains the processing advantage for concrete words over abstract words.

An alternate theory of the concreteness effect was posited by Schwanenflugel and Shoben (1983) in which “context availability” is argued to account for the differences in behavioural outcomes for concrete and abstract words. According to this theory, because concrete words usually have well-understood and unchanging meanings, the word-form can be easily accessed even in situations where the word is removed from a meaningful context (such as single-word tasks). In contrast, abstract words tend to be more poorly understood and ambiguous in meaning and thus are more reliant on contextual information to distinguish their meaning. Other research has suggested that the differences between concrete and abstract words may not be limited to semantic features. Reilly and Kean (2007) showed that the word-form characteristics typical of concrete words differ from those of abstract words, for example, abstract words are usually longer and more morphologically complex than concrete words (see also Reilly, Westbury, Kean & Peelle, 2012). As a consequence, it is suggested that the concreteness effect may be additionally influenced by word-form properties.

Evidence from brain imaging also supports the argument that abstract and concrete words are processed differently in the cognitive system. For example, Binder, Westbury, McKiernan, Possing and Medler (2005) compared fMRI results for abstract and concrete words during a lexical decision task. While there was some overlap in the regions activated during processing of both word types, differences were also evident. In particular, concrete words were associated with greater activation of right hemisphere regions, which the authors suggest may be indicative of access to nonverbal “imagery” information, consistent with Dual-Coding Theory. When abstract words were presented, increased activation of the left inferior frontal regions was evident, which is variably described as either involved in the activation of semantic associations (e.g., Jessen et al., 2000) or processing of the phonological form (Binder et al., 2005).

1.4: Emotion Word Processing Advantage

As noted above, a number of studies have found emotion words can have a processing advantage compared to non-emotion abstract words. For example, Altarriba and Bauer (2004) directly compared the outcomes for emotion, non-emotion abstract and concrete words using a free recall task. Participants were presented with a list of 30 words of each word type and then were asked to recall as many words as possible. The participants recalled significantly more emotion words than non-emotion abstract words and also, perhaps surprisingly, more emotion words than concrete words. Differences between emotion and non-emotion word processing have been demonstrated on a number of other tasks including: visual lexical decision (Altarriba & Bauer, 2004), eye-fixations in reading (Knickerbocker et al., 2015) and a false recall task (Bauer, Olheiser, Altarriba & Landi, 2009) (See Table 1 for an overview of the literature). While the majority of studies have found that emotion words have a processing advantage, studies with bilingual participants have had mixed results. To date, two studies have examined processing of emotion and non-emotion abstract words in proficient bilinguals. El-Dakhs and Altarriba (2018) found that late Arabic-English bilinguals displayed an emotion word advantage on a free recall task in both their first and second languages, however, they did not show a difference between emotion and non-emotion words on an association generation task in either language. Zhang et al. (2019) examined Chinese-English bilinguals' performance on a modified flanker task in which participants indicated if the font colour of a target word was congruent with the colour of a non-target word. They found that there were no significant differences in accuracy or response latencies between emotion words and neutral words. However, as participants were only tested in their second language (English) it is not possible to determine if the same participants would have performed differently in their first language (Chinese). With the aim of examining emotion words in second language acquisition, Altarriba and Basnight-Brown (2012) taught Spanish words for emotion, concrete and non-emotion abstract concepts to monolingual English speakers. Following a learning phase, the participants completed a translation task, on which it was found that emotion words resulted in more errors and longer response latencies than both concrete and non-emotion abstract words. The authors suggested that this showed that emotion words, despite usually having a processing advantage in a speaker's first language, may be particularly difficult to acquire in second language learning.

The processing advantage for emotion words compared with non-emotion abstract words has usually been interpreted within Dual-Coding Theory and therefore as underpinned by the high imageability ratings of emotion words (e.g., Altarriba & Bauer, 2004). Yet, to date studies investigating words that label emotions have not included a manipulation of imageability to directly examine this theory. In other words, there has been no direct test of the Dual Coding Theory's prediction that an emotion-label word processing advantage should disappear when emotion and non-emotion abstract words are matched for imageability. Importantly, in Altarriba and Bauer (2004) emotion words were found to have an advantage over concrete words (which have high imageability). This suggests that other factors might influence emotion and non-emotion word processing beyond imageability. This issue has been explored by other authors who argue that imageability may not underlie the emotion word processing advantage and instead it may be driven by the inherent "emotionality" of the words (Kousta, Vigliocco, Vinson, Andrews & Del Campo, 2011; Kousta et al., 2009; Vigliocco, Kousta, Vinson, Andrews, & Del Campo, 2013; Vinson, Ponari & Vigliocco, 2014). The "emotionality" of words can be defined by two dimensions: valence (how positive or negative a word is) and arousal (how active or passive a concept is) (e.g. Russell, 1980). Regarding valence, emotion words are commonly rated at the extremes: they are either strongly positive (e.g., "happiness"), or strongly negative (e.g., "anger") while on ratings of arousal, emotion words can be placed on a continuum from active (e.g., “excitement”) to passive (e.g., “serenity”) (Cowie & Cornelius, 2003; Cowie et al., 1999).

Kousta et al. (2009) showed that valence, but not arousal, appeared to influence word processing. Using a lexical decision task, they found that both negatively valanced and positively valanced words had a processing advantage over neutral words, and that this was unrelated to arousal ratings. Similarly, Vinson et al. (2014) found that words with extreme valence had shorter latencies on a lexical decision task than words with neutral valence, however arousal ratings did not influence outcomes. Kousta et al. (2011) investigated the influences of concreteness, imageability and valence on the processing of words with emotional associations. A regression analysis using data from the English Lexicon Project database (Balota et al., 2007) indicated that valence significantly influenced outcomes on a lexical decision task, however concreteness and imageability did not.

Strong valence ratings are not limited to words that refer to emotions, but also include words that have emotional connotations (e.g., “dagger”) and Kousta et al. (2009; 2011) and Vinson et al. (2014) included a range of strongly valanced words in their studies. However, words that specifically label emotions have been shown to have an advantage on behavioural outcomes compared with emotion-laden words (i.e., words that have emotional connotations; e.g., Altarriba & Basnight-Brown, 2011; Kazanas & Altarriba, 2016; Knickerbocker & Altarriba, 2013). This suggests that valence alone cannot account for the emotion word processing advantage, with a possible explanation being that imageability and valence both play a role.

Table 1

Summary of Research Regarding Processing of Emotion Label Words.

Authors	Comparison Word/Set	Matching Criteria	Task/s	Emotion Word Advantage?
Altarriba & Bauer (2004)	Non-Emotion Abstract, Concrete	Word Length, Frequency	Free Recall; Visual Lexical Decision	Yes Yes
Bauer et al. (2009)	Non-Emotion Abstract	Word Length, Frequency	False Recall Paradigm	Yes
Knickerbocker & Altarriba (2013)	Emotion Laden, Neutral Words	Word Length, Frequency, Orthographic N	Repetition Blindness	Yes
Knickerbocker et al. (2015)	Emotion Laden, Neutral Words	Word Length, Frequency, Orthographic N	Eye Fixation in Sentence Reading	Yes
BILINGUAL STUDIES				
Altarriba & Basnight-Brown (2012)	Non-Emotion Abstract	Word Length, Frequency	Translation Task	No
El-Dakhs & Altarriba (2018)	Non-Emotion Abstract	Word Length, Frequency	Bilingual Free Recall; Word-Association Task	Yes No
Zhang et al. (2019)	Emotion Laden, Neutral Words	Word Length, Frequency, Valence, Arousal	Modified Flanker Task	No
CLINICAL STUDIES				
Hsieh et al. (2012)	Non-Emotion Abstract	Word Length, Frequency, Syllables, Valence, Arousal, Concreteness	Synonym Matching	No (PPA)
Joubert et al. (2017)	Non-Emotion Abstract	Frequency	Similarity Judgment	No (PPA)

Notes. Included are studies comparing emotion with non-emotion words. Studies classify "emotion words" as words that specifically label emotions. Clinical studies (Hsieh et al., 2012 and Joubert et al., 2017) included participants with primary progressive aphasia (PPA). Participants in all other studies were unimpaired speakers. Studies were conducted in English with the exceptions of El-Dakhs & Altarriba (2018): Arabic and English; Altarriba & Basnight-Brown (2012): Spanish; Joubert et al. (2017): French. Orthographic N = Orthographic Neighbourhood Size.

Taken together, the current literature indicates that while the mechanism underpinning an emotion word processing advantage remains a point of contention, the argument that emotion has an advantageous effect on word processing in unimpaired speakers is largely undisputed (with a possible exception of bilingual speakers). Word processing in aphasia is subject to many of the same influences from word variables as is found in unimpaired speakers, including imageability/concreteness (e.g., Alyahya, Halai, Conroy & Lambon Ralph, 2018b; Bird, Howard & Franklin, 2003). For this reason, it would be expected that people with aphasia should also display a processing advantage for emotion words compared to other non-emotion abstract words. In the following section, literature regarding the perception, comprehension and verbal expression of emotion in people with aphasia will be discussed.

1.5: Emotion Processing by People with Aphasia

A prerequisite for the appropriate use and understanding of emotion words is the ability to perceive and comprehend emotions. These functions have usually been found to be unimpaired in people with aphasia. For example, studies comparing people with aphasia and right-hemisphere brain damaged (RHD) participants consistently show that while RHD participants often display impaired recognition of emotion from facial expressions, this difficulty is rarely evident in people with aphasia (e.g., DeKosky, Heilman, Bowers & Valentstein, 1980; Lorch, Borod & Koff, 1998).

There is also evidence that many faculties relating to the expression of emotion remain intact in aphasia. For example, even in very early reports, it has been noted that people with aphasia typically retain the ability to emote through the use of interjections or expletives (e.g. Head, 1921). People with aphasia are also described as communicating emotional content effectively via nonverbal means such as gesture, facial expressions or pantomime (Blonder et al., 2005; Laakso, 2014; Lorch et al., 1998). In addition, Armstrong and Ulatowska (2007) showed that the personal narratives of people with aphasia usually include emotional evaluations and the participants in their study were reported to use range of verbal techniques, such as repetition, to convey intensity of emotion (e.g., *"It was just sad, sad, sad.")*.

The relative strength of people with aphasia in understanding and expressing of emotion is commonly attributed to the preservation of cortical regions that are thought to be activated in processing of both emotion and emotion words. Processing of emotion words has been associated right-hemisphere activation (e.g., Borod, Andelman, Obler, Tweedy & Wilkowitz, 1992; Landis, 2006) and with activation of distinct neural regions outside the language cortex, specifically limbic and motor cortices (Moseley et al., 2015; Pulvermüller, 2013). Outcomes from a lesion study also suggest that non-language regions may be involved in emotion word processing. Dreyer et al. (2015) report the case of a participant (CA) with a lesion to the left supplementary motor area, who had word recognition difficulties specific to emotion words. The argument that the neural networks involved in processing of emotion words may be relatively preserved in people with aphasia is also supported by electroencephalogram (EEG) findings. Ofek et al. (2013) showed that evoked potentials differed between neutral and emotion words for both unimpaired speakers and people with aphasia. Furthermore, while evoked potentials were delayed in people with aphasia, both participant groups displayed a large P3 amplitude for emotion words, which Ofek et al. suggest indicates that processing of emotion words remains at least partially intact in people with aphasia.

However, weaknesses in the linguistic expression of emotion have also been identified in people with aphasia. Lorch et al. (1998) compared the descriptions of emotional picture scenes by people with aphasia and people with right-hemisphere brain damage. It was found that despite being asked specifically to give their emotional response to the scenes, participants from both groups tended to describe the physical objects in the picture, with this being particularly evident in participants with Wernicke's and conduction aphasia. Blonder et al. (2005) also compared verbal expression of emotion by people with aphasia and people with right-hemisphere brain damage and showed that the participants with aphasia produced a smaller proportion of words with emotional connotations in conversation. Similarly, Armstrong, Mortensen, Ciccone and Godecke (2012) found that, compared with unimpaired speakers, people with aphasia used a smaller range of emotion words in conversation and that the emotion words produced were limited to low-intensity emotions (e.g., "like" rather than "adore").

Studies with participants with primary progressive aphasia (PPA) have made direct comparisons between emotion and non-emotion abstract word processing. Hsieh et al. (2012) showed that participants with semantic variant-PPA had no significant differences in accuracy for emotion and non-emotion abstract words on a synonym matching task. Joubert et al. (2017) found a similar pattern of results using a similarity judgment task, in which participants were presented with two written words and had to indicate whether they were semantically related or unrelated. Although the participants with semantic-variant PPA had impaired comprehension of abstract, concrete and emotion words overall, they showed higher accuracy for abstract words compared with concrete words. Judgement accuracy for emotion words was between that of concrete and abstract words, although not significantly different from either. In contrast, on the same task, unimpaired speakers were significantly more accurate for emotion words than concrete words. Hence, these findings do not confirm the presence of an emotion word advantage in semantic-variant PPA.

1.6: Purpose of Emotion Words

Determining whether emotion word processing can be selectively impaired or, conversely, show a processing advantage in aphasia is of particular interest given the role of emotion words in communication. Expression of emotion is known to have benefits for wellbeing and mental health. For example, the inclusion of emotional content in personal narratives is associated with the formation and maintenance of a sense of self (e.g., Bird & Reese, 2006; Fivush, 2007). In the case of traumatic memories, there is evidence that describing emotional reactions and consequences can help people process distressing events and reach a state of acceptance (e.g., Frattaroli, 2006). Expression of emotional information also has a social function, with people who discuss emotional events with their friends reporting high levels of social integration (Pennebaker & Graybeal, 2001) and long-lasting relationships (Slatcher & Pennebaker, 2006). It is notable that each of these areas can present specific problems for people with aphasia. Many people with aphasia describe a sense of disconnection with their pre-injury self (Brumfitt, 1993) and people with aphasia are more likely to experience distress than people who do not develop aphasia following a stroke (Hilari et al., 2010; Thomas & Lincoln, 2008). In addition, the impact of aphasia on social participation is well-known and has been extensively reported in the literature. People with aphasia on average have fewer friendships than same-aged peers (Davidson, Howe, Worrall, Hickson & Togher, 2008) and many people with aphasia report intense feelings of loneliness (e.g., Davidson et al, 2008; Nyström, 2006; Parr, 2007). Therefore, given the vital communicative functions of emotion words, further investigation of emotion word processing in aphasia may have important implications for assessment and treatment.

2. Aims and Hypotheses

The primary aim of this study is to investigate processing of (strongly valanced) emotion words by people with aphasia compared to processing of abstract words that do not refer to emotions (neutral valanced, non-emotion abstract words). It is hypothesised that under conditions where emotion words have a processing advantage for unimpaired speakers, then this advantage should also be evident in people with aphasia. Hence, participants with aphasia would be expected to show significantly greater accuracy and shorter response latencies for emotion words on single-word tasks.

In addition, this study aims to shed further light on the cognitive mechanisms that may drive any observed emotion word advantage. As discussed, in the current literature there are two main accounts of why emotion words may have a processing advantage: 1) The emotion word processing advantage is driven by the same forces that underpin the concreteness effect (with primary focus given to the influences of imageability), or 2) The emotion word processing advantage is unrelated to the concreteness effect and is instead caused by valence. Here we concentrate on the first of these claims, and in particular, the potential influence of imageability. Surprisingly, despite the argument that imageability underlies emotion word advantage, currently none of the studies that have examined emotion-label words manipulated imageability in order to further investigate this claim. Thus, with this intent, we compare two sets of matched emotion and non-emotion words. One set matched for imageability (and a number of other factors, to be outlined in a later section) and the other unmatched for imageability but matched for concreteness.

If imageability differences underpin a processing advantage for emotion words (as suggested by Altarriba & Bauer, 2004), then a significant difference in accuracy and response latencies between the emotion and non-emotion abstract words should be found for the word set that is *not* matched for imageability, but a difference between the words types should not be evident for the word set that is matched for imageability. On the other hand, if a different mechanism, such as valence, explains the better processing of emotion words, then an emotion word processing advantage will be evident for both word sets.

The specific aims of this study are:

1) Do people with aphasia show an emotion word processing advantage on single-word tasks as measured by accuracy and response latencies?

2) Is an emotion word processing advantage evident when a) concreteness and/or b) imageability is matched between emotion words and non-emotion abstract words?

3. Methods

3.1: Participants

Participants were a convenience sample recruited from the Macquarie University Aphasia Group database and the Royal Rehabilitation Centre (Ryde). Selection criteria were: ability to provide informed consent, diagnosis of aphasia and premorbid fluent English. Participants were excluded from the study who had severe comprehension difficulties that may have impacted on their ability to provide consent and/or their understanding of the study tasks, moderate-severe speech impairments (e.g., dysarthria or apraxia of speech), severe reading impairments or severe word-finding difficulties. All participants had normal or corrected-to-normal hearing and eyesight. The participants received an aphasia-friendly information sheet as well as a verbal description of the study prior to providing written consent. The study consisted of five sessions of 1.5 hours conducted once per week. Participants were paid $15 per session for their participation. This study was conducted with ethics approval from the Macquarie University Human Ethics Committee.

Nine (five female/four male) participants took part in the study. The participants were a minimum of 6 months post injury and had developed aphasia as a result of a stroke, with the exception of JDC who had a cerebral aneurysm. Participants ranged in age from 42 to 82 with a mean age of 66 years. Four participants were bilingual (TLK, OLG, JOG, MLK), all of whom had lived in Australia and spoken English for a minimum of 40 years. One participant (MLK) was receiving individual fortnightly treatment at the time of the study. Three other participants (TLK, OLG and JOG) took part in weekly aphasia discussion groups. The remaining five participants were not receiving any speech pathology intervention at the time of the study. As emotional state is thought to impact on processing of words that label emotions (e.g., Niedenthal, Halberstadt & Setterlund, 1997), participants were screened for depression using the Short-Form Geriatric Depression Scale (GDS; Sheikh & Yesavage, 1986), a 15-item test with a score of either 1 or 0 given for each question. All participants scored five or below, indicating no evidence of depression. A summary of the participants' demographic information is provided in Table 2.

Table 2

Demographic Information for Study Participants.

	BBK	**DTR**	**MMR**	**JDC**	**SPN**	**TLK**	**OLG**	**JOG**	**MLK**
Age	81	60	82	42	53	61	79	71	68
Sex	F	F	F	F	M	M	M	M	F
Years Post Injury	1	3	3	27	3	5	7	9	1
GDS	3	3	2	4	5	0	3	3	5
Bilingual	No	No	No	No	No	Yes (Cantonese/ English)	Yes (Spanish/ English)	Yes (Italian/ English)	Yes (Croatian/ English)
Education (years)	21	13	15	15	13	19	17	18	13
(Previous) Occupation	Nurse	Secretary	Nurse	Student	Electrician	GP	Accountant	Accountant	Tailor

Note. GDS = Geriatric Depression Scale. Scores above five are indicative of depression. GP = General Practitioner.

Participants undertook a battery of background assessments to provide an overview of cognitive and language skills. A summary of participant outcomes is provided in Table 3, with a discussion of each assessment in the section following.

Table 3

Background Assessment Results.

PARTICIPANT	**n=**	**Unimpaired Range Cut-Off**	**BBK**	**DTR**	**MMR**	**JDC**	**SPN**	**TLK**	**OLG**	**JOG**	**MLK**
COMPREHENSIVE APHASIA TEST (CAT)											
Cognitive Screen	38	68%	95%	100%	97%	97%	92%	97%	89%	100%	79%
Comprehension of Spoken Language	66	85%	97%	91%	**82%***	91%	86%	100%	**82%***	95%	**77%***
Comprehension of Written Language	66	80%	94%	91%	82%	**76%***	**74%***	91%	**73%***	86%	**70%***
Repetition	74	91%	99%	**84%***	**81%***	**68%***	**86%***	**82%***	**66%***	**73%***	92%
Repetition of Digit Strings		span of 4	7	4	5	5	5	6	4	4	6
Naming											
Naming Objects	48	90%	**88%***	**85%***	**83%***	**73%***	**58%***	**88%***	**71%***	**85%***	**67%***
Naming Actions	10	80%	100%	80%	**20%***	**50%***	**40%***	**70%***	**40%***	**30%***	**20%***
Word Fluency		13	15	20	19	15	**12***	19	16	**11***	**8***
Spoken Picture Description		33	**26***	33	**21***	**22***	**24***	**25***	**28***	**18***	**16***
Reading											
Reading Words	48	94%	**92%***	100%	**90%***	**79%***	**67%***	94%	**83%***	**90%***	**88%***
Reading Complex Words	6	67%	100%	100%	100%	**33%***	**33%**	83%	**50%**	**33%**	83%
Reading Function Words	6	50%	100%	100%	100%	100%	100%	100%	100%	100%	67%
Reading Nonwords	10	60%	60%	60%	**20%***	**30%***	**40%**	**40%**	**40%***	**0%***	**40%***
PYRAMIDS AND PALM TREES	52	94%	100%	96%	100%	**90%***	**92%***	94%	98%	96%	**85%***
PALPA 50: WRITTEN SYNONYM JUDGEMENTS											
High Imageability	30	94%	100%	100%	100%	97%	**90%***	**93%***	**93%***	**87%***	**57%***
Low Imageability	30	87%	100%	100%	100%	97%	90%	87%	93%	**83%***	**50%***

SHALLICE AND MCGILL WORD-PICTURE MATCHING											
Spoken Version											
Concrete	30	90%	100%	97%	100%	**80%***	**87%***	90%	**77%***	**77%***	**60%***
Abstract	30	70%	90%	87%	70%	83%	**57%***	93%	77%	**67%***	**53%***
Emotion	15	73%	80%	73%	87%	73%	73%	87%	**53%***	**67%***	73%
Written Version											
Concrete	30		97%	90%	80%	83%	97%	73%	86%	80%	60%
Abstract	30		93%	77%	63%	83%	53%	83%	87%	73%	53%
Emotion	15		87%	60%	73%	80%	47%	73%	73%	60%	40%
PALPA 35: ORAL READING REGULARITY											
Regular	30	99%	**97%***	**97%***	**87%***	**83%***	**60%***	**97%***	**93%***	**83%***	**60%***
Irregular	30	99%	**83%***	**90%***	**80%***	**90%***	**63%***	**93%***	**93%***	**77%***	**53%***

Notes. PALPA = Psycholinguistic Assessment of Language Processing in Aphasia (Kay, Lesser & Coltheart, 1992). Accuracy indicated as a percent of the total number of items. Significantly impaired outcomes are indicated in bold and with an asterisk. Norm data for PALPA Subtest 50 (Written Synonym Judgements) obtained from Nickels and Cole-Virtue (2004). Shallice and McGill norms from Lambon Ralph, Sage and Roberts (2000).

3.1.1: Background Assessments

Comprehensive Aphasia Test (CAT; Swinburn, Porter & Howard, 2004). This is a standardised assessment of language ability in people with aphasia. The CAT provides an overview of language comprehension and expression in both spoken and written modalities. A brief cognitive screening assessment is also included. Importantly, given that this study includes a recall task, this assessment includes a digit span subtest.

Pyramids and Palm Trees (3 Pictures Version; Howard & Patterson, 1992). Non-lexical semantics was assessed using the three-picture version of the Pyramids and Palm Trees Test in which participants are presented with a target picture at the top of a page (e.g., a pyramid) and two pictures below (e.g., a palm tree and a fir tree). The participant must choose which of the two pictures is associated with the target picture (i.e., the palm tree).

PALPA Subtest 50: Written Synonym Judgements (Kay et al., 1992). This test requires participants to make a decision whether two written words have similar meanings (e.g., marriage-wedding; idea-notion) or different meanings (e.g., marriage-lamp; idea-security). Half the words are "high imageability" and half "low imageability", which are matched on frequency.

Shallice and McGill Word-Picture Matching (unpublished). This assesses comprehension of concrete, abstract and emotion words, which are approximately matched for familiarity and visual complexity (Warrington & Shallice, 1984). Participants are presented with either a written or spoken word and must choose one of four pictures to match with the target word. The two versions (written and spoken) of the test were presented in different sessions (spoken version first).

PALPA Subtest 35, Oral Reading: Regularity (Kay et al., 1992). This PALPA subtest assesses reading aloud of mono- and bi-syllabic words with regular spelling to sound mappings (e.g. luck) or irregular spelling-sound mappings (e.g., yacht). Word sets are matched for frequency, imageability, grammatical class, and length (number of letters, phonemes and syllables).

3.1.2: Summary of Background Test Results. One role of the background testing was to assess working memory and reading to determine whether participants would be able to complete the experimental tasks. Participants were able to complete the digit span task up to a minimum of 4 items suggesting that working memory should be sufficient to support the recall task. Most of the participants scored relatively well, but below ceiling, on the oral reading subtests of PALPA and CAT (except MMR who was at ceiling for real word reading on the CAT). On the PALPA Oral Reading (Regularity) subtest, participants showed a similar number of errors for regular and irregular spelled words, and no participant showed a significant difference between regular and irregular word accuracy (Fisher Exact Test, $p > .700$ two-tailed for all participants). The majority of errors for all participants were phonological/orthographic (e.g., context → concept) with a small number of no responses (SPN: 2 errors (9% of total errors); MMR: 1 error (10%)). Only BBK and MLK made regularisation errors (BBK: 1 error (17%), MLK: two errors (8%); e.g., iron → /ɪrən/) and MLK also made one unrelated error (mortgage → teacher). As all participants were above floor on irregularly spelled words this suggested that they retained an ability to read via access to the lexicon.

In terms of their general language processing ability, the CAT subtests indicated that the participants were more impaired in language expression compared with comprehension in both written and spoken modalities. All participants had impaired spoken naming.

Three participants (JDC, SPN and MLK) were below cut-off for the Pyramids and Palm Trees Test, indicating a possible semantic impairment. MLK also showed difficulty on written synonym matching with accuracy not significantly above chance (53% accuracy; Binomial Test, $p = 0.699$ two-tailed). The remaining participants were at ceiling or displayed only a mild impairment on this test. On the Shallice and McGill Word-Picture Matching Test, MMR, TLK and OLG had significant differences in accuracy between the test versions; OLG being more accurate for written presentation (McNemar's Test: $p = 0.004$ two-tailed) and MMR and TLK both being more accurate for the spoken version (both $p=0.021$). For the spoken version of the test, JOG had impaired performance for both emotion and abstract words. Three other participants showed impaired performance on *either* emotion *or* abstract words (OLG emotion words; SPN and MLK abstract words). Notably, all of these participants (JOG, OLG, SPN and MLK) were also below the unimpaired range for concrete words, suggesting that their difficulties were not specific to comprehension of abstract or emotion concepts.

3.2: Experimental Stimuli

Experimental stimuli consisted of a total of 52 emotion abstract words and 52 non-emotion abstract words. These stimuli were used to develop two word sets i) 50 emotion and non-emotion abstract words matched for imageability (but not concreteness) and ii) 36 words matched for concreteness (but not imageability). Due to the markedly low concreteness ratings and high imageability ratings for emotion words compared with the low imageability/concreteness ratings for non-emotion abstract words, it was not possible to match on both of these dimensions. 52 concrete words were also selected which act as fillers and are not analysed further.

All of the word stimuli were nouns. Words that could be classified as different word classes depending on context were only included if their dominant use was as a noun (as determined by their definition in the Macquarie English Dictionary, 2019). For example, both “love” and “hate” can be nouns or verbs. For “love”, the dominant meaning is the noun form and thus it was included in the study, while “hate” was excluded as its dominant meaning is the verb form. Homophones were excluded (e.g., feat/feet), as were homographs (e.g., ‘entrance (place of entry) vs. en’trance (to fill with delight)) and compounds (e.g., sunflower).

Emotion and non-emotion abstract words had concreteness ratings of less than 4.5 (on a seven-point scale; Glasgow Norms, Scott, Keitel, Becirspahic, Yao, & Sereno, 2019) and were significantly lower than the concrete words (all comparisons two-sample t-tests $p < .001$ (two-tailed), see Appendix B), which had ratings greater than five. The emotion word set included only words that specifically labelled an emotion or feeling state and were drawn from previous studies: Altaribba and Bauer (2004), Altarriba et al., (1999) and Cowie et al. (1999). Valence ratings for the emotion words were in the range of 1 - 3.25 (negative) and 6.75 - 9 (positive) on a nine-point scale (ratings from the Glasgow norms; Scott et al., 2019). There was a slight imbalance of negative words (n=32) to positive words (n=20), which is consistent with previous findings that have shown a predominance of negative emotion words in English (e.g. Schrauf & Sanchez, 2004). Non-emotion abstract and concrete words were drawn from the neutral range, thereby excluding strongly valanced words. We used two-sample t-tests to determine that arousal and valence ratings of the emotion stimuli were significantly different from those of the non-emotion abstract and concrete words (*emotion vs. non-emotion abstract* concreteness-matched set: arousal t(70) = -2.41, p = .019, valence t(70) = 2.08, p = .041; imageability-matched set: arousal t(98) = -3.21, p = .002; valence t(98) = 2.20, p = .030; *emotion vs. concrete*: concreteness-matched set arousal t(70) = -2.97, p = .004; valence t(70) = 2.51, p = .015; imageability-matched set arousal t(98) = -3.81, p < .001, valence t(98) = 2.56, p = .012).

The word categories were matched across a range of factors relating to frequency and form for both the concreteness-matched and imageability-matched sets: Log Frequency-HAL (written frequency), Log Frequency-Subtlex (spoken frequency), familiarity, length (number of letters, phonemes, syllables), orthographic neighbourhood, phonological neighbourhood, orthographic Levenshtein distance (OLD), phonological Levenshtein distance (OLD), number of morphemes and mean bigram frequency. Ratings were obtained from the English Lexicon Project (Balota et al., 2007) with the exceptions of familiarity, concreteness, imageability, arousal and valence which were obtained from the Glasgow Norms (Scott et al., 2019). The subset of emotion and non-emotion abstract words that were matched for concreteness were significantly different in imageability (concreteness: t(70) = 0.70; p = .486, imageability: t(70) = -2.51, p = .010) and the set that was matched for imageability was significantly different for concreteness (concreteness: t(98) = 4.08, p = < .001; imageability: t(98) = -0.81, p = .419). For those variables where we wished to ensure there was no difference between the word types, we used Bayesian t-tests. Moderate evidence supporting no difference between word sets (Bayes Factor of 0.100-0.333) was found for all variables (see Appendix A for a full list of stimuli and variables, and Appendix B for statistical analyses and means).

The lexical decision task also included 156 nonwords, which were selected from the English Lexicon Project and matched with the real-word stimuli on length (number of letters), mean bigram frequency and orthographic neighbourhood (Bayesian t-test: number of letters:

Bayes Factor (BF) = 0.141; orthographic neighbourhood: BF = 0.135; mean bigram frequency: BF = 0.298). As the real word list included words with multiple morphemes, the nonwords were additionally matched by the number of pseudo-morphemes (through the inclusion of common English affixes e.g. "pleer***ful***") (Bayes Factor = 0.191).

The stimuli were randomly assigned into 3 blocks each comprising approximately equal numbers of each word type (17-18 words/type).

3.3: Experimental Tasks and Procedures

Processing of emotion words was compared to that of non-emotion abstract words across three tasks: list recall, lexical decision and single-word reading. An emotion word processing advantage has been demonstrated on recall and lexical decision tasks in the literature with unimpaired speakers (Recall: e.g., Altarriba & Bauer, 2004; El-Dakhs & Altarriba, 2018; Lexical Decision: e.g., Altarriba & Bauer, 2004; Kousta et al., 2009; Vinson et al., 2014). We additionally included single word reading to examine word production. Reading was chosen rather than picture naming due to the low pictureability of abstract and emotion words.

All three tasks used the same stimuli (52 emotion abstract, 52 non-emotion abstract; 52 concrete) split into three blocks (as described above). Each task took place over three sessions, one block per session (Table 4 provides an example session schedule). No block was repeated within a session.

Background assessments and experimental tasks were completed over five sessions of approximately 1.5 hours each, one week apart. The first session always comprised background testing, and sessions 1 - 4 included both experimental tasks and further background testing. Each block appeared first in a different task, counterbalanced across participants. If a session included a list recall task, this was always completed first to reduce the possibility of accidental recall of words from other blocks used in other tasks. Session order was also counterbalanced across participants. Table 4 presents an example outline of the study sessions.

Table 4

Example Schedule of Tasks.

Session 1	Session 2	Session 3	Session 4	Session 5
Comprehensive Aphasia Test	**List Recall Block 2**	**List Recall Block 3**	**Reading Block 2**	**List Recall Block 1**
Shallice and McGill (Spoken)	**Lexical Decision Block 3**	**Reading Block 1**	**Lexical Decision Block 1**	**Reading Block 3**
Geriatric Depression Scale	PALPA Oral Reading	**Lexical Decision Block 2**	Shallice and McGill (Written)	Pyramids and Palm Trees Test
		PALPA Synonym Judgements		

Notes. Order of task presentation was counterbalanced across participants. Experimental tasks in bold font, background assessments in normal font. Sessions were separated by one week.

Dependent variables were accuracy data for all tasks and response times (for correct responses) for lexical decision and reading. An outline of the procedures for each study task is provided below.

3.3.1: List Recall. The list recall task followed the procedures described by Altarriba and Bauer (2004) with some adaptations in order make the task more accessible for people with aphasia. Similar to Altarriba and Bauer, word types were presented blocked (i.e., recall for emotion words was separate from recall for non-emotion abstract). Although our study had more items (52 per set, compared with 20 per set) as the stimuli were presented in three blocks, participants received a similar number of items as in Altarriba and Bauer in each session (three 17-18 word lists per word type). Within a single session, one list of each word type (emotion, non-emotion abstract and concrete) was presented with the order of presentation counterbalanced between sessions.

The participants were instructed that they would hear a list of words which they would need to remember. The words were presented in the same random order at a rate of five seconds/word, however unlike Altarriba and Bauer (2004) who presented the words as auditory recordings, the word lists were presented as videos of the researcher saying the words in order to allow the participants to benefit from lipreading cues. The participants were asked to verbally recall list items in any order ("free recall") immediately following completion of the list presentation.

As we predicted that some participants may be impaired in free recall due to word production difficulties, immediately following free recall, we additionally used an identification task (hereafter referred to as "identification recall"). For this task, participants were presented with written target words intermixed with distractor words (at a ratio of 2 target words to 1 distractor word). Participants were asked to identify words that they recognised from the list.

3.3.2: Lexical Decision. The lexical decision task required the participants to decide whether the presented written target was a real English word or not. The task was presented using PsychoPy 3.0 (Peirce et al., 2019) which allows accuracy and response time data to be collected from a button-press response. An equal number of nonword and real word (156 each) test items were presented split across the three testing blocks.

Each presentation began with four filler words (two real words, two nonwords) to familiarise the participants with the procedure. During this time, the researcher provided feedback and reiterated the task instructions if necessary. The stimuli were presented in lowercase size 60 black Arial font in the centre of a light grey screen. Participants responded using one hand to press the left arrow key on the keyboard for a “yes” (indicating a real word) and the right arrow key for a “no” response (indicating a nonword). In order for the keys to be easily identifiable by the participants, they were labelled with a green sticker with the word “yes” and a red sticker with the word “no”. The stimulus word disappeared after two seconds and response latencies beyond this period were not collected. The next trial was initiated by the participant pressing the spacebar button on the computer.

3.3.3: Single-Word Reading Aloud. Single written words were presented in randomised order within each block using PsychoPy 3.0. The words were presented in lowercase size 60 black Arial font on a light grey screen, as per the lexical decision task. The participants were provided with written and verbal instructions indicating that they would see a single word on the screen which they needed to read aloud as quickly as possible.

Each presentation began with three practice items during which time the researcher provided feedback if needed. Each trial lasted for five-seconds after which the stimulus word would disappear. The PyschoPy program recorded a single Waveform audio file for each trial using a condenser microphone, with the recording initiating simultaneously with the presentation of the stimulus word and ending after 5 seconds. Thus, only responses completed within that time period were included in the analysis. The participant initiated the next trial by pressing the spacebar.

3.4: Response Scoring and Data Pre-Processing.

3.4.1: Accuracy Data. Accurate responses for the free recall task were correct productions of target stimuli from the list presented, or productions with minor phonological errors (defined as nonwords that had only one phoneme difference and that were unambiguous attempts at a target word e.g., “bappiness” for “happiness”). Words from other lists (e.g., a word from the concrete list when assessing the emotion word list) or non-target words were counted as incorrect. Oral reading accuracy was based on the first full response, which was defined as the first response that could comprise a phonotactically legal word (defined as a combination of English-language phonemes containing at least one vowel, not including a short vowel in word final position). False starts which contained phonemes in the same order as a subsequent response (e.g. co.. co.. corn) were not considered a full response and the subsequent response (in this example, “corn”) would be coded. Accuracy for the lexical retrieval task was based on an accurate button press response with no time limit.

3.4.2: Response Latency Data. Response latencies were analysed for correct responses. Response latencies for the lexical decision task were recorded by the PyschoPy programme. For oral reading, response latencies were manually adjusted for correct responses only by inspection of the waveform using Audacity audio editor software (Version 3.2.1). Response times were determined by the initiation of the first full response measured in milliseconds. Responses that included false starts before the full response were excluded from response time analysis (e.g., “ha, happiness”) as were responses with filled pauses (e.g., “um”) where the initiation point of the full response was difficult to determine from the waveform. In addition, response latencies were not obtainable for 11 words for MMR (8% of total correct responses; 4 emotion words, 4 non-emotion abstract words, 3 concrete words) due to background noise making the initiation point unclear. Prior to analysis, response latency data were inspected for responses under 200 milliseconds or more than three standard deviations above each participant's mean, for removal from the data set. The percentage of outliers excluded was between 0.9% and 4.8% of total responses (BBK: Lexical Decision: 1%; Reading: 1.1%; DTR: Lexical Decision: 1.9%, Reading: 0.9%; MMR: Reading: 0.9%; JDC: Lexical Decision: 2%; SPN: Lexical Decision: 4.8%, Reading: 1%; TLK: Lexical Decision: 1%, Reading: 2%; OLG: Lexical Decision: 0.9%, Reading: 0.9%; JOG: Lexical Decision: 1%; MLK: Lexical Decision: 2.9%).

4. Results

4.1: List Recall

In free recall, participants were able to recall only a very small number of words, nevertheless, significant differences between emotion and non-emotion word recall were evident for two participants. JDC recalled significantly more emotion words than non-emotion abstract words for the concreteness-matched word set (Fisher Exact Test; exact $p = .036$, two-tailed). SPN also showed better recall for the emotion words, however this was significant for the imageability-matched word set only ($p = .031$). All other comparisons were non-significant ($p > .05$, see Appendix D for full details of statistical results).

On the identification recall task, all participants produced significantly more 'yes' responses to targets than distractors (Fisher Exact Test, all $p < .001$ two-tailed), indicating performance greater than chance. JDC once again showed significantly better performance for emotion words and this was evident for both the concreteness-matched ($p = .001$) and imageability-matched ($p = .006$) word sets. Although SPN was numerically more accurate at identifying emotion words in both word sets, this was not significant (concreteness matched set p

= .563; imageability-matched set, $p = .329$). Other participants all showed non-significant effects ($p > .05$). Figure 1 shows total number words accurately recalled and identified in the two word sets (concreteness-matched and imageability-matched).

We additionally considered whether a there was a difference between performance for the emotion and non-emotion abstract words at the group level using Stouffer's Z-test (Zaykin, 2011). This method combines the participants' z-scores to form an overall group Z. In addition, a test of homogeneity was conducted to determine if there was a significant difference in the pattern of outcomes across participants. A significant effect combined across participants was found for the identification recall task for the concreteness-matched set and a trend for the free recall task (free recall: $z = -1.70$, $p = .089$ two-tailed; identification recall: $z = -2.87$, $p = .004$) and for both free recall and identification recall on the imageability-matched set (free recall: $z = -2.37$, $p = .018$; identification recall: $z = -2.25$, $p = .024$). Across the group, outcomes were not significantly heterogeneous for either set (*concreteness-matched set*, free recall: $H(8) = 6.63$, $p=.577$; identification recall: $H(8) = 9.14$, $p = .331$; *imageability-matched set*, free recall: $H(8) = 4.68$, $p=.792$; identification recall: $H(8) = 9.11$, $p = .333$).

As four of our participants were bilingual, we also examined whether the pattern of outcomes for the bilingual participants differed from the monolingual participants, albeit with little power given the small number in each subgroup: there were no significant differences between the monolingual and bilingual groups on this task (two sample t-test comparing z-scores: *concreteness-matched set*: free recall $t(7) = 1.35$, $p = .215$;, identification recall $t(7) = 0.61$, $p = .556$; *imageability-matched set*: free recall $t(7) = 1.24$, $p = .249$, identification recall $t(7) = 1.21$, $p = .260$).

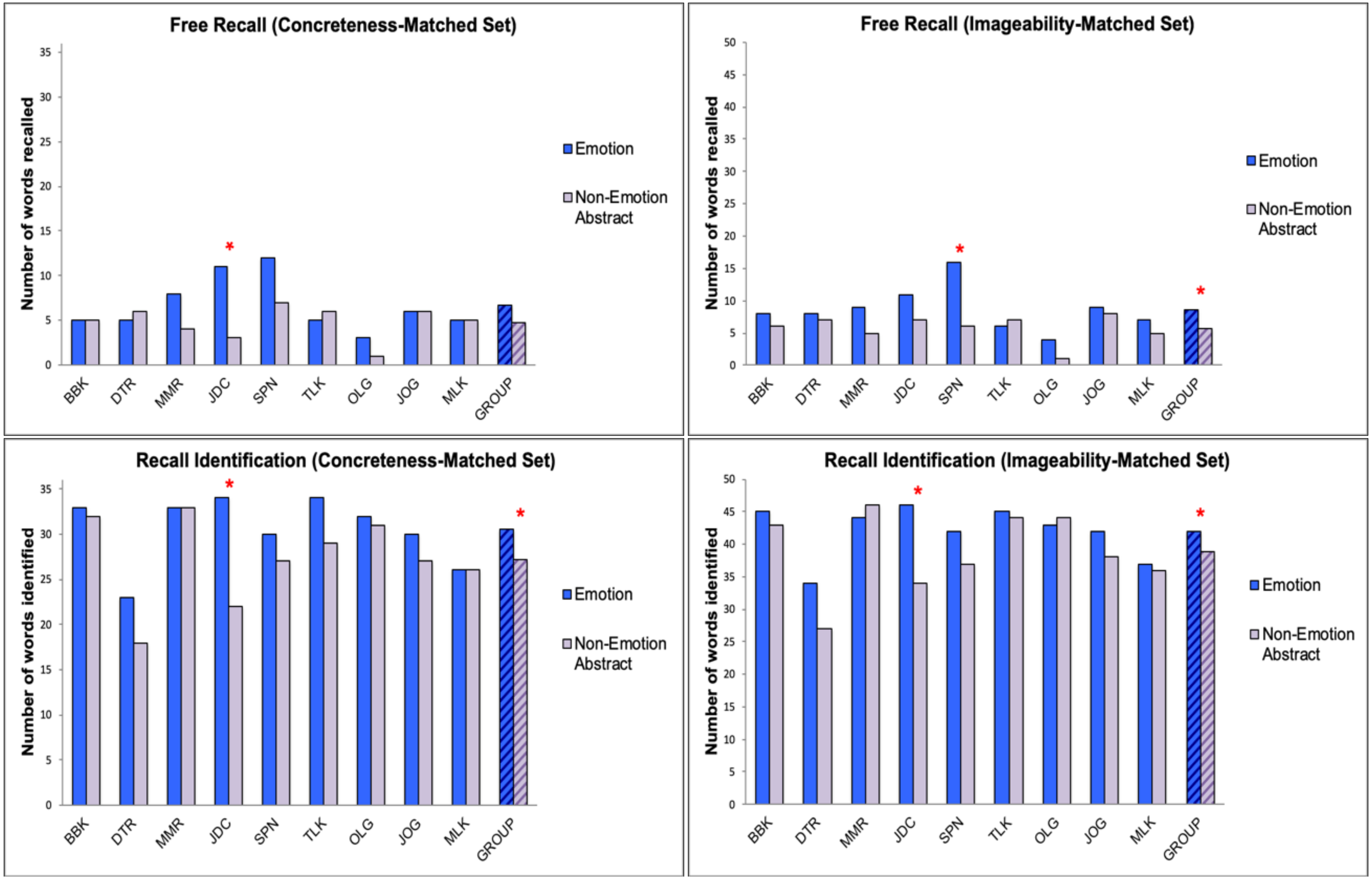


Figure 1: List recall results. Upper panels show free recall accuracy, lower panels show identification recall accuracy. * indicates p < .05 on a Fisher Exact (two-tailed) for individual participants and Stouffer's Test for group data. Initials refer to participants, grouped into monolingual (BBK, DTR, MMR, JDC, SPN) and bilingual (TLK, OLG, JOG, MLK) and ordered by severity of reading difficulties.

4.2: Lexical Decision

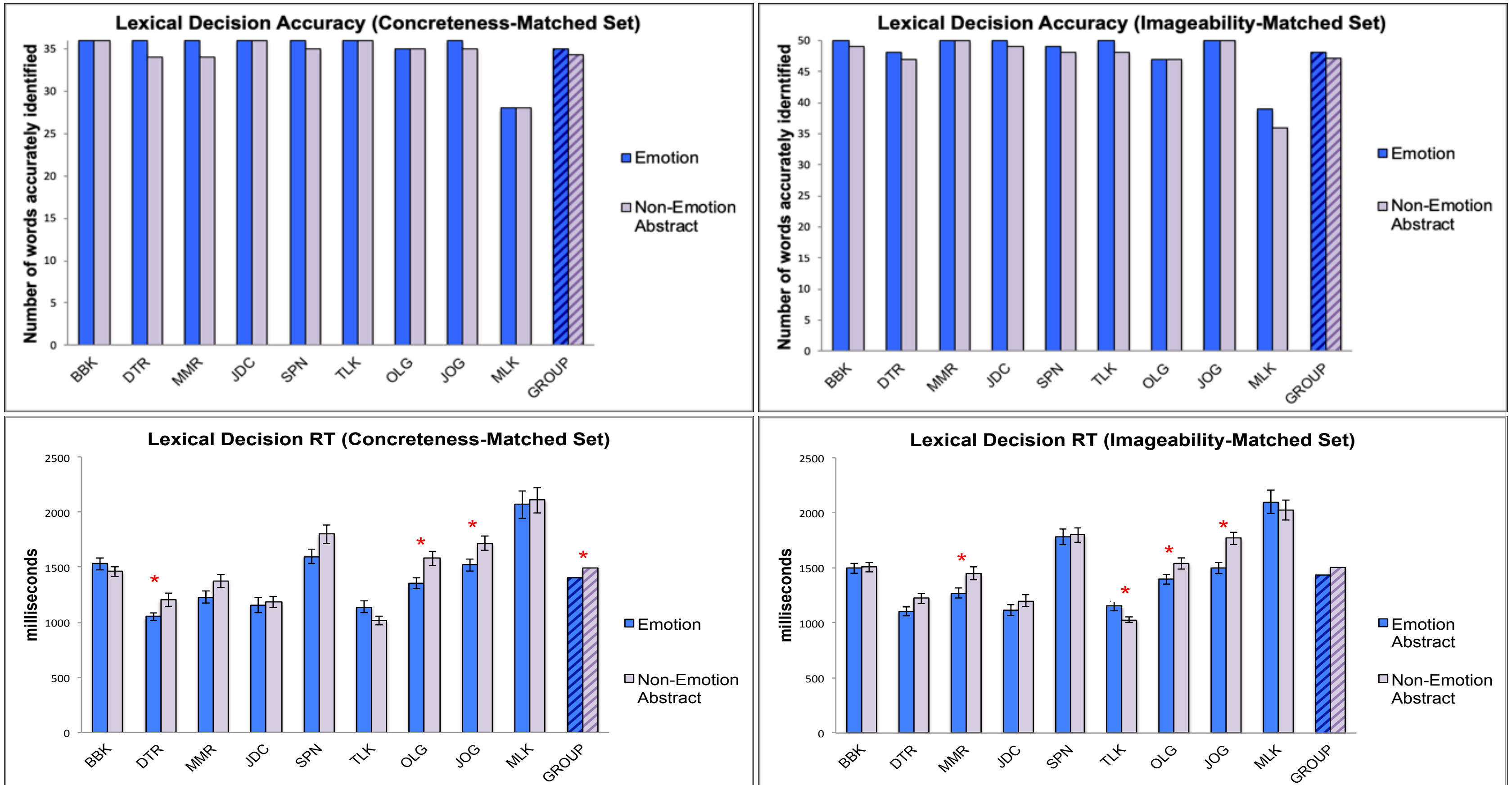
All participants performed significantly above chance on the lexical decision task (all participants, Binomial test, $p < 0.001$, two-tailed). With the exception of MLK, participants performed at or close to ceiling for both the concreteness-matched and imageability-matched word sets (see Figure 2). No significant differences in accuracy between the word types were evident for any participant (all participants, Fisher Exact Test; exact $p > 0.400$ two-tailed).

However, for response latency, significant differences were found for five participants. On the concreteness-matched set, DTR (two sample t-test; $t(61) = 2.19$, $p = 0.033$ two-tailed), OLG ($t(64) = 2.66$, $p = 0.010$) and JOG ($t(65) = 2.28$, $p = 0.026$) showed significantly shorter latencies for emotion than non-emotion abstract words. OLG ($t(90) = 2.09$, $p = 0.039$) and JOG ($t(89) = 3.35$, $p = 0.001$) showed the same pattern on the imageability-matched set, along with MMR ($t(95) = 2.44$, $p = 0.017$). However, TLK showed the reverse pattern and was significantly faster for the non-emotion abstract words ($t(80) = 2.49$, $p = 0.015$) for the imageability-matched set.

Group level analyses showed a significant effect for response latencies in the concreteness-matched set ($z = -2.87$, $p = .004$) with a trend in the imageability matched set ($z = -1.73$, $p = .084$). No significant outcomes were found for accuracy (concreteness-matched set: $z = -0.399$, $p = .697$; imageability-matched set: $z = -0.81$, $p = .420$). The homogeneity test was significant for response latency data, indicating that there was significant variance in performance between the participants (concreteness-matched set: $H(8) = 19.73$, $p = .041$;

Note. Degrees of freedom for individual-level t-tests refer to number of total items analysed and for group-level analysis indicate the total number of participants.

imageability-matched set: H(8) = 28.71, $p < .001$), but there was no significant heterogeneity for accuracy (concreteness-matched set accuracy: H(8) = 5.68, $p = .683$; imageability-matched set accuracy: H(8) = 6.04, $p = .792$) . There were no significant differences in performance between bilingual and monolingual participants for either accuracy or response latency (*concreteness-matched set:* accuracy $t(7) = 0.41$, $p = .694$, response latency $t(7) = 0.19$, $p = .852$; *imageability-matched set:* accuracy $t(7) = 0.08$, $p = .940$; response latency $t(7) = 0.34$, $p = .739$).



	BBK	DTR	MMR	JDC	SPN	TLK	OLG	JOG	MLK
N (Emotion)	36	34	36	33	28	36	32	33	18
N (Non-emotion abstract)	36	34	34	34	32	35	34	34	15

	BBK	DTR	MMR	JDC	SPN	TLK	OLG	JOG	MLK
N (Emotion)	50	47	50	47	46	49	45	44	25
N (Non-emotion abstract)	49	47	47	48	44	47	47	47	20

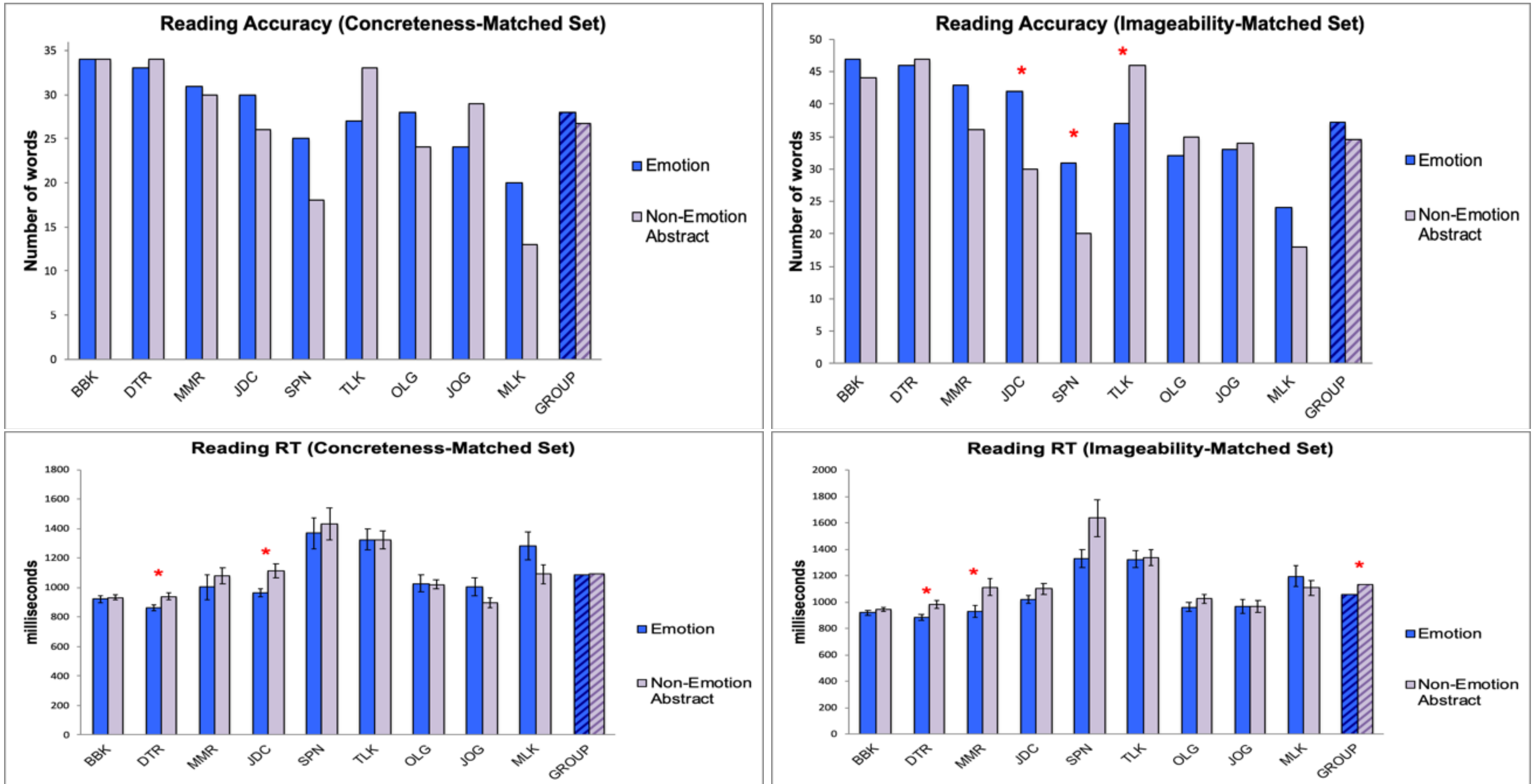
Figure 2: Lexical Decision results. Upper panels show accuracy data; lower panels show response time data. * indicates significant result, $p < .05$. Individual analysis: Fisher Exact (two-tailed) (accuracy data) or two sample t-tests (response latency data), group analysis: Stouffer's Test. Tables indicate number of analysed items for response latency tasks. Initials refer to participants, grouped into monolingual (BBK, DTR, MMR, JDC, SPN) and bilingual (TLK, OLG, JOG, MLK) and ordered by severity of reading difficulties. Error bars represent one standard error. RT = response time.

4.3: Single-Word Reading Aloud

On the oral reading task, three participants (JDC, SPN and TLK) showed significant differences in accuracy of emotion and non-emotion abstract words, but only for the imageability-matched, and not the concreteness-matched, set. JDC ($p = .015$) and SPN ($p = .046$) showed significantly higher accuracy for emotion words, while TLK ($p = .034$) displayed the opposite effect with significantly higher accuracy for non-emotion words (see Figure 3).

In the response latency analysis, JDC again showed a significant advantage for emotion words, along with DTR and MMR. JDC's latencies were significantly shorter for emotion words compared with non-emotion abstract for the concreteness-matched set only ($p = .008$, imageability-matched set $p = .113$), while MMR showed this effect for the imageability-matched word set ($p = .022$, concreteness-matched set, $p = .389$). DTR was significantly faster when reading emotion words for both word sets (concreteness-matched word set: $p = .028$; imageability-matched set: $p = .014$).

At the group level, there was only one significant measure: emotion words had significantly shorter response latencies (and marginally significantly greater accuracy) for the imageability-matched set ($z = -3.76$, $p < .001$; *concreteness-matched set* accuracy: $z = -0.71$, $p=.477$, response latency: $z = -1.27$, $p = .205$; *imageability-matched set* accuracy: $z=-1.70$, $p = .088$). For two of the non-significant outcomes, response latencies for the concreteness-matched set and accuracy for the imageability-matched set, the participants showed significant variability in performance (*concreteness-matched set:* accuracy $H(8) = 5.68$, $p = .683$; response latencies $H(8) = 15.67$, $p = .047$; *imageability-matched set:* accuracy $H(8) = 20.02$, $p = .044$; response latencies $H(8) = 6.51$, $p = .590$). There was no significant difference between monolinguals and bilinguals on reading accuracy (concreteness-matched set: $t(7) = 0.67$, $p = .521$; imageability-matched set: $t(7) = 2.02$, $p = .079$). However, a significant difference was found between the bilingual and monolingual participants for response latency on both word sets (concreteness-matched set: $t(7) = 3.53$, $p = .008$; imageability-matched set: $t(7) = 2.83$, $p =.022$). Observation of the monolingual participant outcomes shows that all participants had shorter response latencies for emotion words. In contrast, bilingual participants had either similar outcomes for emotion and non-emotion words or numerically slower (but non-significant) response times for emotion words.



	BBK	DTR	MMR	JDC	SPN	TLK	OLG	JOG	MLK
N (Emotion)	33	34	25	28	24	27	25	23	16
N (Non-emotion abstract)	36	36	23	25	19	33	22	21	14

	BBK	DTR	MMR	JDC	SPN	TLK	OLG	JOG	MLK
N (Emotion)	45	44	38	39	28	33	32	28	19
N (Non-emotion abstract)	44	44	33	27	20	42	28	28	17

Figure 3: Single-Word Reading results. Upper panels display accuracy data, lower panels display response time data. * Indicates significant result on Fisher Exact (two-tailed) (accuracy data) or Independent T-Test (response latency data). Group analysis: Stouffer's Test. Tables indicate number of analysed items for response latency tasks. Initials refer to participants, grouped into monolingual (BBK, DTR, MMR, JDC, SPN) and bilingual (TLK, OLG, JOG, MLK) and ordered by severity of reading difficulties. RT = response time. Error bars represent one standard error.

Table 5

Summary of Patterns of Performance Across Tasks.

		BBK	DTR	MMR	JDC	SPN	TLK	OLG	JOG	MLK	Group
Concreteness-Matched Set	Free Recall	ns	ns	↑ ns	↑ **.036**	↑ ns	ns	ns	ns	ns	↑ .088
	List Recall ID	ns	↑ ns	ns	↑ **.001**	↑ ns	↑ ns	ns	↑ ns	ns	↑ **.004**
	Lex Dec Accuracy	ns	ns	ns	ns	ns	ns	ns	ns	ns	↑ ns
	Lex Dec RT	↓ ns	↑ **.033**	↑ .076	↑ ns	↑ .065	↓ .062	↑ **.009**	↑ **.025**	↓ ns	↑ **.005**
	Reading Accuracy	ns	ns	ns	↑ ns	↑ ns	↓ ns	↑ ns	↓ ns	↑ ns	↑ ns
	Reading RT	↑ ns	↑ **.028**	↑ ns	↑ **.008**	↑ ns	ns	↓ ns	ns	↓ ns	ns
Imageability-Matched Set	Free Recall	ns	ns	↑ ns	ns	↑ **.015**	ns	↑ ns	ns	↑ ns	↑ **.018**
	List Recall ID	ns	↑ ns	ns	↑ **.006**	↑ ns	ns	ns	↑ ns	↑ ns	↑ **.025**
	Lex Dec Accuracy	ns	ns	ns	ns	ns	ns	ns	ns	↑ ns	↑ ns
	Lex Dec RT	ns	↑ .062	↑ **.017**	↑ ns	ns	↓ **.015**	↑ **.038**	↑ **.001**	↓ ns	↑ .084
	Reading Accuracy	↑ ns	ns	↑ ns	↑ **.015**	↑ **.047**	↓ **.034**	↓ ns	ns	↑ ns	↑ .088
	Reading RT	↑ ns	↑ **.014**	↑ **.024**	↑ ns	↑ .060	ns	ns	↑ ns	↓ ns	↑ **< .001**

Notes. Green up arrow indicates emotion words had higher accuracy (or shorter response latencies), blue down arrow indicates emotion words had lower accuracy (or longer response latencies), no arrow indicates equivalent performance (or difference within 5% of the total items for accuracy data or within 5% of one standard deviation for latency data). Statistics reported are Fisher Exact Tests (two-tailed) for accuracy data and two-sample t-tests (two-tailed) for response time data. Statistically significant results are reported in bold red font, marginal results ($.05 < p < .1$) are indicated in black font, non-significant results ($p < .1$) are indicated by "ns". RT = response time, ID = identification, Lex Dec = Lexical Decision.

5. Discussion

The purpose of this study was to investigate emotion word processing, and specifically whether, in people with aphasia, emotion words have a processing advantage compared with non-emotion abstract words. This question was inspired by previous research findings with unimpaired speakers which have demonstrated an advantage for emotion abstract words compared with non-emotion abstract words (e.g., Altarriba & Bauer, 2004; El-Dakhs & Altarriba, 2018). We also addressed a secondary issue: if there was an emotion word aprocessing dvantage, was it underpinned by the higher imageability of emotion words as predicted by Dual Coding Theory?

To address these questions, we tested nine people with aphasia using two tasks that have previously demonstrated an emotion word processing advantage in unimpaired speakers: list recall (Altarriba & Bauer, 2004; El Dahks & Altarriba, 2018) and lexical decision (Altarriba & Bauer, 2004; Kousta et al., 2009; Vinson et al., 2014); and additionally, included a single-word reading aloud task. Critically, we used stimuli that were matched for more psycholinguistic properties than had occurred in previous research, and, in particular had two subsets of items, one matched for concreteness and the other imageability.

Across both the concreteness- and imageability-matched sets, an emotion word processing advantage was found only for some tasks/analyses and only for some participants (see Table 5, above, for a summary of significant outcomes). If we first consider the two tasks on which an emotion word advantage has been shown in unimpaired speakers, list recall and lexical decision, we found mixed outcomes at both the group and individual level. For the list recall task, an emotion word advantage was found for the group as a whole for free recall (imageability- but not concreteness-matched set), and identification recall (both concreteness- and imageability matched set). However, at the individual level, only one participant showed significant effects on each measure (JDC, free recall in the concreteness-matched set and identification recall in both word sets; SPN, free recall in the imageability-matched set).

For lexical decision, at group level a significant emotion word advantage was found for response latencies on the concreteness-matched set, while the outcomes for the imageability-matched set were marginal ($p = .084$). However, homogeneity test outcomes indicated that there was significant variance across participants in the patterns shown for both word sets. This is consistent with individual level findings which showed that four participants produced significantly shorter response latencies for the emotion words (in both sets: OLG, JOG; imageability-matched set only: MMR; concreteness-matched set only: DTR), but one participant, TLK, displayed faster responses for the non-emotion abstract words (significant for the imageability-matched set). No significant differences were found at either the group or individual level for lexical decision accuracy with all participants (except MLK) performing at, or close to, ceiling.

On the reading task, a significant group level effect was found on only one measure: for the imageability-matched set response latencies were shorter for emotion words than non-emotion abstract words. The pattern of participants' performance was significantly heterogenous for response latencies on the concreteness-matched set and accuracy on the imageability-matched set. In addition, a significant effect of bilingualism was found for response latency data in both word sets. Monolingual participants displayed shorter response latencies for emotion words (significant for DTR in both word sets, JDC in the concreteness-matched set and MMR in the imageability-matched set). In contrast, bilingual participants showed similar response latencies for emotion words and non-emotion abstract words. Regarding accuracy data, two individual participants, JDC and SPN, showed higher accuracy for emotion words on the imageability-matched set, while one, TLK, showed the opposite effect on the same set.

Therefore, while evidence of an emotion word processing advantage was found, the results were far from consistent, or strong, across participants and tasks. It is possible, however, that the pattern of outcomes was influenced by participant characteristics, including bilingualism, severity and level of impairment. Each of these factors will be discussed in turn.

A potential source of variation in outcomes could be the language history of our participants. As described in the Introduction, previous studies with unimpaired bilingual speakers have shown mixed outcomes with regard to emotion word processing. While El-Dakhs & Altarriba (2018) found that bilinguals displayed an emotion word advantage in both their first and second languages on a recall task, this pattern was not evident in either language when the participants were asked to generate word associations. Zhang et al. (2019) found that their bilingual participants did not have an emotion word effect in their second language (however participants were not tested in their first language) and Altarriba and Basnight-Brown (2012) showed that English speakers had greater difficulty learning emotion words in Spanish compared with non-emotion words. Altarriba and Basnight-Brown suggest that emotion words may thus be more difficult to acquire in second language learning compared with first language acquisition. The possibility that emotion word processing may differ in second language speakers is also supported by evidence that bilinguals tend to associate emotion words in their second language with less intense emotions compared with their first language (Pavlenko, 2008). Interestingly, the four bilingual participants in this study (TLK, OLG, JOG and MLK) showed different patterns of outcomes. For example, TLK was the only (monolingual or bilingual) participant to display evidence of a reverse emotion word effect with significantly lower accuracy for emotion words compared with non-emotion abstract words in reading and slower response latencies for emotion words on the lexical decision task (both on the imageability-matched word set). In contrast, OLG and JOG both showed an advantage for emotion words with significantly faster response latencies for emotion words on the lexical decision task (both word sets). The final bilingual participant, MLK, showed no significant differences between emotion and non-emotion words. Moreover, as discussed above, comparison of bilingual and monolingual participants revealed a significant difference in the pattern of outcomes for response latencies in reading. On this measure, the bilingual participants had little difference between the word types, while a number of monolingual participants had significantly faster response times for the emotion words. However, there were no significant differences between bilingual and monolingual participants on any other task or outcome measure. Emotion word processing in bilingualism was not the focus of this study and a key limitation to drawing clear conclusions was the small number of bilingual participants. It may be interesting for future research to consider the different effects of emotion word processing in monolingual and bilingual populations with a larger sample size and also to examine emotion word processing by bilingual speakers with aphasia in both their first and second languages.

Among the monolingual participants, the degree to which an emotion word advantage was apparent in accuracy data appeared to be related to the participant's ability to do the task: the effect was more pronounced for participants with more severe impairments. To test this hypothesis, while being aware of limited power with this small sample, we performed correlations between accuracy and the difference between emotion and non-emotion word responses. Reading accuracy showed a large (but marginally significant) negative correlation between overall reading accuracy and the difference in accuracy between emotion and non-emotion reading (r(4) = -.855, p = .065). This corresponds with the pattern observed in individual participant outcomes: while mildly impaired participants (BBK and DTR) showed similar accuracy between emotion and non-emotion abstract words, moderately impaired participants (MMR, JDC and SPN) displayed a numerical advantage for emotion word reading with this reaching significance for JDC and SPN. Accuracy data for lexical decision and identification recall, as well as response latencies for reading and lexical decision also showed a (non-significant) negative correlation. In contrast, free recall showed a positive correlation, possibly in part due to floor effects in participants who had greater difficulty on this task (Table 6 provides full details of correlations). Therefore, from our small set of participants there appeared to be a relationship between performance and their overall accuracy on the experimental tasks, principally with regard to reading. This leads to a question of whether the nature and severity of an individual's underlying language impairments may predict their relative outcomes for emotion and non-emotion abstract words.

Table 6

Correlations between overall task accuracy and the emotion word advantage

Dependent Variable	**Monolingual Participants (r(4))**	**All Participants (r(8))**
Free Recall (items recalled)	0.859^{+}	0.523
Identification Recall (items correctly selected)	-0.614	-0.534
Lexical Decision Accuracy	-0.559	-0.428
Lexical Decision RT	-0.674	-0.411
Reading Accuracy	-0.855^{+}	-0.587^{+}
Reading RT	-0.065	-0.599^{+}

Notes. RT = response time. Pearson correlations are calculated between overall accuracy and the difference between emotion word and non-emotion abstract word accuracy or RT (emotion - non-emotion abstract) for each task.
$^{+}$ indicates marginal significance $0.05 < p < 0.10$.

With this in mind, we next considered whether the nature of a participant's impairments (and the underlying level of breakdown) may have influenced the degree to which an emotion word advantage was apparent. The participants who displayed the most clear and consistent emotion word effect (JDC and SPN) showed similarities in their background assessment outcomes: along with MLK, they were the only participants impaired on the Pyramids and Palm Trees Test, which is consistent with a possible breakdown at the semantic-level. Hence, it may be that emotion words are less vulnerable to impairment than non-emotion abstract words in people with impaired semantic processing. However, MLK did not show the same pattern of outcomes on the experimental tasks, which is possibly due to the influence of her bilingual language background.

The finding that the strongest effects of emotion on processing was in individuals with semantic impairment contrasts with previous studies with semantic-variant Primary Progressive Aphasia (sv-PPA). These investigations found no significant differences in outcomes for emotion and non-emotion abstract words (Hsieh et al., 2012; Joubert et al., 2017). One possibility is that the difference in findings can be explained by the difference in the tasks utilised in our study compared with the studies with people with sv-PPA. As noted, the sv-PPA studies focused solely on comprehension (e.g., synonym judgement task), while our study included reading and recall tasks. This may be significant as previous studies have found a dissociation between comprehension and production of semantic categories in people with aphasia (e.g., Hillis & Caramazza, 1991). For example, SPN, despite showing a difference for emotion and non-emotion words on the production tasks, was similarly impaired on both these word types in comprehension as measured by the Shallice and McGill Word-Picture Matching Test. A second possible explanation for the incongruency between our findings and those found for sv-PPA, is that it reflects the different nature of the semantic impairments in sv-PPA. For example, a number of cases have been reported in which people with sv-PPA have shown a reverse concreteness effect (better processing of abstract words compared with concrete words) (e.g., Breedin et al., 1994; Papagno, Capasso & Miceli, 2009), while we know of no studies that have reported such cases in stroke aphasia. Reilly, Peele and Grossman (2007) propose that this unusual pattern may be a result of the unique neural degradation in PPA that leads to poorer access to the sensory information used in processing of concrete concepts. In line with this hypothesis, people with sv-PPA may be unable to utilise the links to visual imagery that have been hypothesised to underpin the advantage for emotion word processing. As discussed in the Introduction, neuroimaging studies suggest that processing of emotion words involves activation of widespread neural networks outside the language system including activation of right-hemisphere regions and the limbic and motor cortices (e.g., Pulvermüller, 2013). It is thus possible that the pathways involved in processing of emotion words remain intact in our participants with semantic impairment (JDC, SPN) and hence their language deficits are less severe than for non-emotion abstract words for which processing is largely contained within the language system (e.g., Noppeney & Price, 2004). Unfortunately, no neuroimaging data was available for these participants in order to test this hypothesis.

The second question addressed in this study was whether imageability can explain differences in processing of emotion and non-emotion abstract words. If this were true, then we would expect to have seen an emotion word advantage only for the word set matched for concreteness but not the set matched for imageability. We found this was not the case. Indeed, contrary to expectations, on some measures significant findings were evident for the imageability-matched word set but not for the concreteness-matched set. This was most clear for accuracy on the reading task for which a significant emotion word advantage was found for two participants on the imageability-matched set, while there were no significant results for any participants on the concreteness-matched word set. However, differences in power may also have contributed to this pattern: the imageability-matched set was larger with 100 words compared to the 72 words in the concreteness-matched set. Nonetheless, that an effect was found for some participants for the imageability-matched set indicates that imageability alone cannot explain emotion words having a processing advantage.

The results of this research therefore present us with two key questions: (1) Why was the emotion word processing advantage less consistent and less strong than has been found in previous studies with unimpaired participants? And, (2) Why were differences between emotion and non-emotion abstract words found when imageability was matched?

In relation to the first of these issues, one consideration is whether our participants' aphasia impacted on the processing of emotion and non-emotion abstract words. While it is possible that particular individuals with aphasia may have a pattern of impairments that precluded a processing advantage for emotion words, the outcomes of this study are not in line with a general statement that aphasia reduces the emotion word processing advantage.

This is supported by the fact that if language impairments reduced the emotion word processing advantage then we would expect to see a positive correlation between overall accuracy and the presence of an emotion word advantage with individuals with a mild impairment showing a pattern similar to unimpaired speakers. Yet, as discussed previously, this is not consistent with our findings. Instead for three of the four accuracy measures (identification recall, lexical decision and reading) and for response latencies for both lexical decision and reading, we found a negative correlation between overall task accuracy and the size of the emotion word processing advantage with this pattern strengthening when only monolingual participants were included (see Table 6, earlier). This indicates that as the participants' severity increased (measured by the ability to do the task), the emotion word processing advantage usually became more pronounced (except for free recall, where a floor effect occurred). Therefore, a proposal that aphasia reduces the emotion word processing advantage appears unlikely. This is in line with previous studies that have shown that lexical processing in aphasia is influenced by many of the same semantic variables (such as imageability and concreteness) that impact on lexical processing in unimpaired speakers (e.g., Bird et al., 2003; Nickels & Howard, 1994, 1995).

Another possibility is that there were characteristics of the word sets that could explain the diminished effect for our participants. The word sets used in this study were tightly matched on a total of 13 lexical variables that are known to affect participant performance. This represents substantially more matching variables than used in previous studies that have compared emotion and non-emotion abstract words, the majority of which have matched stimuli on only length and frequency (see Table 1, earlier). Given that our findings were substantially less strong than those found for unimpaired participants, it is plausible that the improved matching on one or more of these additional criteria, such as familiarity or number of morphemes, may have played a role in lessening the significant outcomes found for our study.

With regard to the fact that an emotion word advantage was evident even when emotion and non-emotion words were matched for imageability, this suggests that other factors must influence emotion word processing. The first possibility is that the emotion word processing advantage is still linked to the concreteness effect, however it is driven by other factors that distinguish concrete and abstract words, such as context availability. Altarriba et al. (1999) found that emotion words have high ratings of context availability compared with non-emotion abstract words, which would be consistent with emotion words having better processing than non-emotion abstract words. As we did not specifically match our word lists on context availability, we cannot rule out the possibility that the sets may have differed on this variable.

A second possible explanation is that emotion words have a processing advantage due to their extreme ratings in valence as is suggested by Kousta et al. (2009; 2011) and Vinson et al. (2014). The emotion words in this study were strongly valanced, while the non-emotion words were neutral. Therefore, if valence underlies the emotion word processing advantage then we would expect an effect to be evident in both word sets. This is not inconsistent with our findings. We further explored the relative impact of imageability and valence using linear regression with all other matching variables included in the model as controls (see Method and Appendix B for a list of variables). The regression model was significant for all measures except free recall (see Appendix E for a full report of regression outcomes). However, imageability was a significant predictor only for the identification recall task (Overall model, adjusted $R^2 = 0.341$, $F(15,88) = 3.04$, $p < .001$, imageability $\beta = .282$, $t(102) = 2.43$, $p = .017$) while valence was not a significant predictor of accuracy or latency on any task.

This indicates that for the majority of measures, the variables we included in the model predicted outcomes (accuracy/response latencies), however neither imageability nor valence were significant predictors. In line with Reilly and Kean's (2007) findings that word-form characteristics may contribute to the concreteness effect, it may be that the lexical form features included in the model influenced outcomes. However, given many of our variables were highly intercorrelated (such as number of letters and number of phonemes), we could not determine which other variables predicted performance (due to multicollinearity indicated with high Variance Inflation Factors).

6. Limitations

In addition to the limitations already noted, there are a number of further areas where the current study's design limits the conclusions that can be drawn. Firstly, this study focused on nouns. However, research suggests that people with aphasia may show differences in processing of abstract and concrete verbs (Alyahya, Halai, Conroy. Lambon Ralph, 2018a) and furthermore, difficulties producing emotion verbs (e.g. "like", "hate") in conversation have been identified in aphasia (Armstrong et al., 2012). As such, future studies may consider similar question with regard to emotion verb processing. Furthermore, given that emotion words can commonly occur in adjective form (e.g., "happy", "angry"), future studies may additionally consider processing of emotion and non-emotion adjectives.

In addition, the use of written stimuli places a limitation on interpretation of the outcomes given the extent to which they rely on reading ability. It is not clear whether the same pattern would be observed when word production is prompted by other stimuli, such as pictures, or in natural connected speech.

This study was also limited by the small number of participants. For this reason, while some patterns of performance were identified, further investigation with a greater number of participants would help to clarify the factors that may have influenced participant performance. In particular, greater participant numbers may allow for analysis of the potential influences of bilingualism, aphasia type and/or lesion site on the processing of emotion words. In addition, future studies may consider perception of emotion more broadly and the possible effects of a deficit in this area on emotion word processing. Furthermore, the addition of data from unimpaired speakers, including bilinguals, on the same stimuli would further facilitate the interpretation of results from people with aphasia.

7. Conclusions

Previous research has shown that lexical processing in aphasia is often influenced by concreteness. In this research, the focus moved away from the concrete/abstract dichotomy and considered whether emotion abstract words are differentially processed compared with non-emotion abstract words. Consistent with previous research with unimpaired speakers, we showed that abstract words that describe emotions *can have* a processing advantage over other non-emotion abstract words. However, our findings revealed inconsistencies across participants and tasks, with bilingualism, severity of impairments and semantic-level breakdown each exerting an influence over participant results. Investigation of the factors underlying emotion word processing suggested that the heightened imageability ratings of emotion words could not explain the effect. In sum, while this study confirms that emotion may be an important factor in the processing of abstract words by people with aphasia, there is need for continued research to better understand when, why and how a processing advantage for emotion words is observed in clinical populations.

References

Altarriba, J., & Basnight-Brown, D. (2011). The representation of emotion vs. emotion-laden words in English and Spanish in the Affective Simon Task. *International Journal of Bilingualism, 15*(3), 310-328. doi: 10.1177/1367006910379261

Altarriba, J., & Basnight-Brown, D. (2012). The acquisition of concrete, abstract, and emotion words in a second language. *International Journal of Bilingualism, 16*(4), 446-452. doi: 10.1177/1367006911429511

Altarriba, J., & Bauer, L. M. (2004). The distinctiveness of emotion concepts: A comparison between emotion, abstract, and concrete words. *American Journal of Psychology, 117*(3), 389-410. doi: 10.2307/4149007

Altarriba, J., Bauer, L., & Benvenuto, M. (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. doi: 10.3758/BF03200738

Alyahya, R., Halai, A., Conroy, P., & Lambon Ralph, M. (2018a). The behavioural patterns and neural correlates of concrete and abstract verb processing in aphasia: A novel verb semantic battery. *NeuroImage: Clinical, 17*, 811-825. doi: 10.1016/j.nicl.2017.12.009

Alyahya, R., Halai, A., Conroy, P., & Lambon Ralph, M. (2018b). The relationship between the multidimensionality of aphasia and psycholinguistic features. *Aphasiology, 32*(Sup1), 5-6. doi: 10.1080/02687038.2018.1486374

Armstrong, E., Mortensen, L., Ciccone, N., & Godecke, E. (2012). Expressing opinions and feelings in a conversational setting. *Seminars in Speech and Language, 33*(1), 16-26. doi: 10.1055/s-0031-1301160

Armstrong, E., & Ulatowska, H. (2007). Making stories: Evaluative language and the aphasia experience. *Aphasiology, 21*(6-8), 763-774. doi: 10.1080/02687030701192364

Audacity Team (2019). Audacity(R): Free Audio Editor and Recorder (Version 2.3.1) [computer application]. Retrieved from https://audacityteam.org

Ayçiçegi-Dinn, A., & Caldwell-Harris, C. (2009). Emotion-memory effects in bilingual speakers: A levels-of-processing approach. *Bilingualism: Language and Cognition, 12*(3), 291-303. doi: 10.1017/S1366728909990125

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

Bauer, L., Olheiser, E., Altarriba, J., & Landi, N. (2009). Word type effects in false recall: Concrete, abstract, and emotion word critical lures. *The American Journal of Psychology, 122*(4), 469-481. Retrieved from http://www.jstor.org.simsrad.net.ocs.mq.edu.au/stable/27784422

Binder, J. R., Westbury, C. F., McKiernan, K. A., Possing, E. T., & Medler, D. A. (2005). Distinct Brain Systems for Processing Concrete and Abstract Concepts. *Journal of Cognitive Neuroscience, 17*(6), 905-917. doi: 10.1162/0898929054021102

Bird, H., Howard, D., & Franklin, S. (2003). Verbs and nouns: The importance of being imageable. *Journal of Neurolinguistics, 16*(2-3), 113-149. doi: 10.1016/S0911-6044(02)00016-7

Bird, A., & Reese, E. (2006). Emotional Reminiscing and the Development of an Autobiographical Self. *Developmental Psychology, 42*(4), 613-626. doi: 10.1037/0012-1649.42.4.613

Blonder, L., Heilman, K., Ketterson, T., Rosenbek, J., Raymer, A., Crosson, B., . . . Rothi, L. (2005). Affective facial and lexical expression in aprosodic versus aphasic stroke patients. *Journal of the International Neuropsychological Society, 11*(6), 677-685. doi: 10.1017/S1355617705050794

Borod, J., Andelman, F., Obler, L., Tweedy, J., & Wilkowitz, J. (1992). Right hemisphere specialization for the identification of emotional words and sentences: Evidence from stroke patients. *Neuropsychologia, 30*(9), 827-844. doi: 10.1016/0028-3932(92)90086-2

Breedin, S., Saffran, E., & Coslett, H. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology, 11*(6), 617-660. doi: 10.1080/02643299408251987

Brumfitt, S. (1993). Losing your sense of self: What aphasia can do. *Aphasiology, 7*(6), 569-575. doi: 10.1080/02687039308248631

Butler, S. (Ed.) (2019). *Macquarie Dictionary (Online Version)*. Retrieved February 21, 2019 from https://www.macquariedictionary.com.au

Christian, J., Bickley, W., Tarka, M., & Clayton, K. (1978). Measures of free recall of 900 English nouns: Correlations with imagery, concreteness, meaningfulness, and frequency. *Memory & Cognition, 6*(4), 379-390. doi: 10.3758/BF03197470

Code, C., & Petheram, B. (2011). Delivering for aphasia. *International Journal of Speech-Language Pathology, 13*(1), 3-10, doi: 10.3109/17549507.2010.520090

Cowie, R., & Cornelius, R. (2003). Describing the emotional states that are expressed in speech. *Speech Communication, 40*(1-2), 5-32. doi: 10.1016/S0167-6393(02)00071-7

Cowie, R., Douglas-Cowie, E., Apolloni, B., Taylor, J., Romano, A., & Fellenz, W. (1999). What a neural net needs to know about emotion words. *Computational Intelligence and Applications, 404*, 5311-5316. Retrieved from http://www.lfsag.unito.it/antonio_romano/1999_CSCC_Cowie_et_al.pdf

Crystal, D. (2003). *The Cambridge encyclopedia of the English language (2nd ed.).* Cambridge, United Kingdom: Press Syndicate of the University of Cambridge.

Davidson, B., Howe, T., Worrall, L., Hickson, L., & Togher, L. (2008). Social Participation for Older People with Aphasia: The Impact of Communication Disability on Friendships. *Topics in Stroke Rehabilitation, 15*(4), 325-340. doi: 10.1310/tsr1504-325

DeKosky, S., Heilman, K., Bowers, D., & Valenstein, E. (1980). Recognition and discrimination of emotional faces and pictures. *Brain and Language, 9*(2), 206-214. doi: 10.1016/0093-934X(80)90141-8

Dellantonio, S., Mulatti, C., Pastore L., & Job R. (2014). Measuring inconsistencies can lead you forward. The case of imageability and concreteness ratings. *Frontiers in Psychology, 5,* 708. doi: 10.3389/fpsyg.2014.00708

Dreyer, F., Frey, D., Arana, S., von Saldern, S., Picht, T., Vajkoczy, P., & Pulvermuller, F. (2015). Is the Motor System Necessary for Processing Action and Abstract Emotion Words? Evidence from Focal Brain Lesions. *Frontiers in Psychology, 6*, 1661. doi: 10.3389/fpsyg.2015.01661

El-Dakhs, D., & Altarriba, A. (2018). The Distinctiveness of Emotion Words: Does It Hold for Foreign Language Learners? The Case of Arab EFL Learners. *Journal of Psycholinguistic Research, 47*(5), 1133-1149. doi: 10.1007/s10936-018-9583-6

Ferreira, D. L. (2012). *Aphasia incidence and intervention in the acute hospital setting* (Honours thesis, Edith Cowan University, Perth, Australia). Retrieved from https://ro.ecu.edu.au/theses_hons/64

Fivush, R. (2007). Maternal Reminiscing Style and Children's Developing Understanding of Self and Emotion. *Clinical Social Work Journal, 35*(1), 37-46. doi: 10.1007/s10615-006-0065-1

Franklin, S., Howard, D., & Patterson, K. (1994). Abstract word meaning deafness. *Cognitive Neuropsychology, 11*(1), 1-34. doi: 10.1080/02643299408251964

Franklin, S., Howard, D., & Patterson, K. (1995). Abstract word anomia. *Cognitive Neuropsychology, 12*(5), 549-566. doi: 10.1080/02643299508252007

Frattaroli, J. (2006). Experimental Disclosure and Its Moderators: A Meta-Analysis. *Psychological Bulletin, 132*(6), 823-865. doi: 10.1037/0033-2909.132.6.823

Ghio, M., Vaghi, M., Perani, D., & Tettamanti, M. (2016). Decoding the neural representation of fine-grained conceptual categories. *NeuroImage, 132*, 93-103. doi: 10.1016/j.neuroimage.2016.02.009

Head, H. (1921). Discussion on Aphasia. *Brain, 43*(4), 412-450. doi: 10.1093/brain/43.4.412

Hilari, K., Northcott, S., Roy, P., Marshall, J., Wiggins, R., Chataway, J., & Ames, D. (2010). Psychological distress after stroke and aphasia: The first six months. *Clinical Rehabilitation, 24*(2), 181-190. doi: 10.1177/0269215509346090

Hillis, A., & Caramazza, A. (1991). Category-specific naming and comprehension impairment: A double dissociation. *Brain, 114*(5), 2081-94. doi: 10.1093/brain/114.5.2081.

Howard, D. & Patterson, K. (1992). *The Pyramids and Palm Trees est. A test of semantic access from words and pictures.* Bury St. Edmunds, UK: Thames Valley Company.

Hsieh, S., Foxe, D., Leslie, F., Savage, S., Piguet, O., & Hodges, J. (2012). Grief and Joy: Emotion Word Comprehension in the Dementias. *Neuropsychology, 26*(5), 624-630. doi: 10.1037/a0029326

James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance, 1*(2), 130-136. doi: 10.1037/0096-1523.1.2.130

Jessen, F., Heun, R., Erb, M., Granath, D., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The Concreteness Effect: Evidence for Dual Coding and Context Availability. *Brain and Language, 74*(1), 103-112. doi: 10.1006/brln.2000.2340

Joubert, S., Vallet, G., Montembeault, M., Boukadi, M., Wilson, M., Laforce, R., . . . Brambati, S. (2017). Comprehension of concrete and abstract words in semantic variant primary progressive aphasia and Alzheimer's disease: A behavioral and neuroimaging study. *Brain and Language, 170*, 93-102. doi: 10.1016/j.bandl.2017.04.004

Kay, J., Lesser, R. & Coltheart, M. (1992) *PALPA: Psycholinguistic Assessments of Language Processing in Aphasia.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Kazanas, S., & Altarriba, A. (2016). Emotion Word Processing: Effects of Word Type and Valence in Spanish–English Bilinguals. *Journal of Psycholinguistic Research, 45*(2), 395-406. doi: 10.1007/s10936-015-9357-3

Knickerbocker, H., & Altarriba, J. (2013). Differential repetition blindness with emotion and emotion-laden word types. *Visual Cognition, 21*(5), 599-627. doi: 10.1080/13506285.2013.815297

Knickerbocker, H., Johnson, R., & Altarriba, J. (2015). Emotion effects during reading: Influence of an emotion target word on eye movements and processing. *Cognition and Emotion, 29*(5), 784-806. doi: 10.1080/02699931.2014.938023

Kousta, S., Vigliocco, G., Vinson, D., Andrews, M., & Del Campo, E. (2011). The Representation of Abstract Words: Why Emotion Matters. *Journal of Experimental Psychology: General, 140*(1), 14-34. doi: 10.1037/a0021446

Kousta, S., Vinson, D., & Vigliocco, G. (2009). Emotion words, regardless of polarity, have a processing advantage over neutral words. *Cognition, 112*(3), 473-481. doi: 10.1016/j.cognition.2009.06.007

Laakso, M. (2014). Aphasia Sufferers' Displays of Affect in Conversation. *Research on Language and Social Interaction, 47*(4), 404-425. doi: 10.1080/08351813.2014.958280

Lambon Ralph, M., Sage, K., & Roberts, J. (2000). Classical anomia: A neuropsychological perspective on speech production. *Neuropsychologia, 38*(2), 186-202. doi: 10.1016/S0028-3932(99)00056-1

Landis, T. (2006). Emotional Words: What's so Different from Just Words? *Cortex, 42*(6), 823-830. doi: 10.1016/S0010-9452(08)70424-6

Lorch, M., Borod, J., & Koff, E. (1998). The role of emotion in the linguistic and pragmatic aspects of aphasic performance. *Journal of Neurolinguistics, 11*(1-2), 103-118. doi: 10.1016/S0911-6044(98)00008-6

Moseley, R., Shtyrov, Y., Mohr, B., Lombardo, M., Baron-Cohen, S., & Pulvermüller, F. (2015). Lost for emotion words: What motor and limbic brain activity reveals about autism and semantic theory. *NeuroImage, 104*, 413-422. doi: 10.1016/j.neuroimage.2014.09.046

Nickels, L., & Cole-Virtue, J. (2004). Reading tasks from PALPA: How do controls perform on visual lexical decision, homophony, rhyme, and synonym judgements? *Aphasiology, 18*(2), 103-126. doi: 10.1080/02687030344000517-633

Nickels, L., & Howard, D. (1994). A frequent occurrence? factors affecting the production of semantic errors in aphasic naming. *Cognitive Neuropsychology, 11*(3), 289-320. doi: 10.1080/02643299408251977

Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia, 33*(10), 1281-1303. doi: 10.1016/0028-3932(95)00102-9

Niedenthal, P., Halberstadt, J. B., & Setterlund, J. (1997). Being Happy and Seeing "Happy' ': Emotional State Mediates Visual Word Recognition. *Cognition and Emotion, 11*(4), 403-432. doi: 10.1080/026999397379863

Noppeney, U., & Price, C. (2004). Retrieval of abstract semantics. *NeuroImage, 22*(1), 164-170. doi: 10.1016/j.neuroimage.2003.12.010

Nyström, M. (2006). Aphasia - an existential loneliness: A study on the loss of the world of symbols. *International Journal of Qualitative Studies on Health and Well-being, 1*(1), 38-49. doi: 10.1080/17482620500501883

Ofek, E., Purdy, S., Ali, G., Webster, T., Gharahdaghi, N., & McCann, C. (2013). Processing of emotional words after stroke: An electrophysiological study. *Clinical Neurophysiology, 124*(9), 1771-1778. doi: 10.1016/j.clinph.2013.03.005

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review 76*(3), 241-263. doi: 10.1037/h0027272

Paivio, A. (1991). Dual Coding Theory: Retrospect and Current Status. *Canadian Journal of Psychology/Revue Canadienne De Psychologie, 45(*3), 255-287. doi: 10.1037/h0084295

Paivio, A., Rogers, T., & Smythe, B. (1968). Why are pictures easier to recall than words? *Psychonomic Science, 11*(4), 137-138. doi:10.3758/BF03331011

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. doi: 10.1037/0278-7393.20.5.1196

Paivio, A., Yuille, J., & Madigan, S. (1968). Concreteness, Imagery, and Meaningfulness Values for 925 Nouns. *Journal of Experimental Psychology, 76*(1 Pt.2), 1-25. doi: 10.1037/h0025327

Papagno, C., Capasso, R., & Miceli, G. (2009). Reversed concreteness effect for nouns in a subject with semantic dementia. *Neuropsychologia, 47*(4), 1138-1148. doi: 10.1016/j.neuropsychologia.2009.01.019

Parr, S. (2007). Living with severe aphasia: Tracking social exclusion. *Aphasiology, 21*(1), 98-123. doi: 10.1080/02687030600798337

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and cognition*, *11*(2), 147-164. doi: 10.1017/S1366728908003283

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195-203. doi: 10.3758/s13428-018-01193-y

Pennebaker, J., & Graybeal, A. (2001). Patterns of Natural Language Use: Disclosure, Personality, and Social Integration. *Current Directions in Psychological Science, 10*(3), 90-93. doi: 10.1111/1467-8721.00123

Pulvermüller, F. (2013). How neurons make meaning: Brain mechanisms for embodied and abstract-symbolic semantics. *Trends in Cognitive Sciences, 17*(9), 458-470. doi: 10.1016/j.tics.2013.06.004

Reilly, J., & Kean, J. (2007). Formal Distinctiveness of High- and Low-Imageability Nouns: Analyses and Theoretical Implications. *Cognitive Science, 31*(1), 157-168. doi: 10.1080/03640210709336988

Reilly, J., Peelle, J., & Grossman, M. (2007). A unitary semantics account of reverse concreteness effects in semantic dementia. *Brain and Language, 103*(1-2), 86-87. doi: 10.1016/j.bandl.2007.07.057

Reilly, J., Westbury, C., Kean, J., & Peelle, J. E. (2012). Arbitrary symbolism in natural language revisited: when word forms carry meaning. *PloS one, 7*(8), e42286. doi:10.1371/journal.pone.0042286

Russell, J. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178. doi: 10.1037/h0077714

Schrauf, R., & Sanchez, J. (2004). The Preponderance of Negative Emotion Words in the Emotion Lexicon: A Cross-generational and Cross-linguistic Study. *Journal of Multilingual and Multicultural Development, 25*(2-3), 266-284. doi: 10.1080/01434630408666532

Schwanenflugel, P., & Shoben, E. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(1), 82-102. doi: 10.1037/0278-7393.9.1.82

Scott, G., Keitel, G., Becirspahic, A., Yao, M., & Sereno, B. (2019). The Glasgow Norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods, 51*(3), 1258-1270. doi: 10.3758/s13428-018-1099-3

Scott, G., O'Donnell, P., Leuthold, H., & Sereno, S. (2009). Early emotion word processing: Evidence from event-related potentials. *Biological Psychology, 80*(1), 95-104. doi: 10.1016/j.biopsycho.2008.03.010

Setti, A., & Caramelli, N. (2005) Different domains in abstract concepts. In B. Bara, L. Barsalou, and M. Bucciarelli (Eds.), *Proceedings of the XXVII Annual Conference of Cognitive Science Society* (pp.1997-2002). New Jersey: Cognitive Science Society. Retrieved from https://cora.ucc.ie/handle/10468/2709

Sheikh, J., & Yesavage, J. (1986). Geriatric Depression Scale (GDS): Recent Evidence and Development of a Shorter Version. *The Journal of Aging and Mental Health, 5*(1-2), 165-173. doi: 10.1300/J018v05n01_09

Slatcher, R., & Pennebaker, J. (2006). How Do I Love Thee? Let Me Count the Words: The Social Effects of Expressive Writing. *Psychological Science, 17*(8), 660-664. doi: 10.1111/j.1467-9280.2006.01762.x

Stroke Foundation. (2017). *No Postcode Untouched.* Retrieved from https://strokefoundation.org.au/en/What-we-do/Research/No-postcode-untouched

Swinburn, K., Porter, G., & Howard, D. (2004). *The Comprehensive Aphasia Test*, Hove, UK: Psychology Press.

Thomas, S. A., & Lincoln, N. B. (2008). Predictors of Emotional Distress After Stroke. *Stroke, 39*(4), 1240-1245. doi: 10.1161/STROKEAHA.107.498279

Vigliocco, G., Kousta, S., Vinson, D., Andrews, M., & Del Campo, E. (2013). The Representation of Abstract Words: What Matters? Reply to Paivio's (2013) Comment on Kousta Et Al. (2011), *Journal of Experimental Psychology: General, 142*(1), 288-291. doi: 10.1037/a0028749

Vinson, D., Ponari, M., & Vigliocco, G. (2014). How does emotional content affect lexical processing? *Cognition & Emotion, 28*(4), 737-746. doi: 10.1080/02699931.2013.851068

Warrington, E., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107*(3), 829-54. doi: 10.1093/brain/107.3.829

Zaykin, D. (2011). Optimally weighted Z-test is a powerful method for combining probabilities in meta-analysis. *Journal of Evolutionary Biology, 24*(8), 1836-1841. doi: 10.1111/j.1420-9101.2011.02297.x

Zhang, J., Teo, T., & Wu, C. (2019). Emotion Words Modulate Early Conflict Processing in a Flanker Task: Differentiating Emotion-Label Words and Emotion-Laden Words in Second Language. *Language and Speech, 62*(4), 641–651. doi: 10.1177/0023830918807509

Appendix A

Word Stimuli

Words are organised alphabetically in word set: Emotion (E), Non-emotion Abstract (A) and Concrete (C). Ratings for imageability (Im), concreteness (CN), arousal (Ar), valence (VAL), familiarity (FAM) are from the Glasgow Norms (Scott et al., 2019). Written frequency (Log Freq HAL), spoken frequency (Log Freq SUBLECT), number of letters (let), number of phonemes (phon), number of syllables (syll), orthographic neighbourhood size (Orth N), phonological neighbourhood size (phon N), Orthographic Levenshtein Distance (OLD), Phonological Levenshtein Distance (PLD), mean bigram frequency (bigram freq) and number of morphemes (Mor) are from English Lexicon Project (Balota et al., 2007). Concreteness-matched (CN-match) set and imageability-matched (Im-match) set indicate "Y" for "yes", "N" for "no".

Word	Set	Im	CN	Ar	Val	Log Freq HAL	Log Freq SUBT	Fam	Lett	Phon	Syll	Orth N	Phon N	OLD	PLD	Bigram Freq	Mor	CN-Match	Im-Match
affection	E	4.60	3.12	6.82	8.18	7.55	2.64	5.88	9	6	3	0	0	2.65	1.95	1942	2	Y	Y
amusement	E	4.57	3.77	6.49	7.63	7.68	2.28	5.79	9	9	3	0	0	3.00	3.60	1798	2	Y	Y
anger	E	4.76	3.65	5.64	2.53	9.06	3.00	6.24	5	4	2	2	2	1.65	1.80	3694	1	Y	Y
anguish	E	3.63	2.56	4.77	2.46	6.88	2.05	3.74	7	6	2	0	0	2.75	2.55	2045	1	N	Y
animosity	E	2.68	2.22	6.48	2.61	6.49	1.51	3.61	9	9	5	0	0	4.00	4.50	1534	1	N	Y
annoyance	E	3.40	2.97	4.46	2.47	7.25	1.42	5.79	9	6	3	0	0	3.20	2.85	1427	2	Y	Y
anxiety	E	4.06	2.62	5.27	2.26	7.89	2.54	5.59	7	8	4	0	0	3.00	4.25	1158	2	Y	Y
awe	E	3.11	2.53	6.66	7.34	7.91	2.13	5.38	3	1	1	12	9	1.30	1.00	355	1	N	Y
bliss	E	3.94	2.67	6.46	8.09	7.70	2.21	5.13	5	4	1	2	3	1.90	1.80	2065	1	N	Y
confidence	E	3.77	3.06	6.85	7.77	9.60	3.00	6.39	10	8	3	0	1	2.60	3.60	2056	2	Y	Y
depression	E	3.77	3.06	3.79	1.73	8.95	2.61	6.18	10	7	3	1	1	2.40	2.35	2670	3	Y	Y
despair	E	3.71	3.13	4.59	2.03	8.26	2.48	5.09	7	6	2	0	1	2.90	2.35	1791	1	Y	Y
distress	E	3.94	3.29	5.46	2.29	7.62	2.57	5.38	8	7	2	1	0	2.55	2.40	2927	1	Y	Y
ecstasy	E	4.35	4.03	6.97	7.15	7.70	2.21	4.85	7	7	3	0	0	3.50	2.90	1421	1	Y	Y
embarrassment	E	4.09	3.07	4.85	2.26	7.66	2.39	6.03	13	11	4	0	0	5.25	3.90	1750	2	Y	N
empathy	E	3.21	2.30	5.38	7.27	7.68	1.77	5.83	7	6	3	0	0	2.85	2.80	1457	3	N	Y
enjoyment	E	4.39	2.97	7.47	8.18	8.31	1.85	6.27	9	8	3	0	0	3.55	3.50	1609	3	Y	Y
enthusiasm	E	4.18	2.82	6.91	8.00	7.91	2.36	6.32	10	9	5	1	0	4.30	5.20	1635	2	N	Y
euphoria	E	3.79	1.97	7.82	8.31	6.28	1.42	4.09	8	7	4	1	0	3.80	3.60	1293	2	N	Y
excitement	E	4.94	3.03	7.70	8.29	8.58	2.80	6.47	10	9	3	0	1	3.65	3.30	1981	3	Y	Y
fear	E	3.94	2.47	5.17	2.56	10.45	3.55	5.97	4	3	1	12	40	1.25	1.00	1783	1	Y	Y

frustration	E	3.65	2.44	5.11	2.40	8.37	2.18	6.22	11	9	3	0	0	3.25	3.40	2694	2	N	Y
fury	E	4.71	3.27	5.79	2.43	8.26	2.29	4.94	4	5	2	4	2	1.65	1.90	799	1	Y	Y
glee	E	4.53	2.80	6.24	7.62	6.45	1.77	4.24	4	3	1	3	9	1.80	1.20	1574	1	Y	Y
gratitude	E	3.77	2.55	5.43	7.85	7.27	2.60	5.24	9	8	3	0	0	3.20	3.05	2336	2	Y	Y
grief	E	3.89	3.14	4.88	1.94	8.27	2.74	5.47	5	4	1	1	10	1.95	1.45	1331	1	Y	Y
guilt	E	3.00	2.53	4.63	2.29	8.52	2.88	6.00	5	4	1	4	12	1.70	1.35	654	1	N	Y
happiness	E	5.27	2.97	6.97	8.57	9.07	3.10	6.65	9	7	3	0	0	2.90	2.55	2322	2	Y	Y
hatred	E	3.03	3.55	5.73	1.78	8.78	2.44	5.84	6	6	2	0	0	1.95	2.65	3208	2	Y	Y
hope	E	3.50	2.46	6.85	8.00	11.92	4.21	6.36	4	3	1	12	18	1.20	1.00	1203	1	N	Y
humiliation	E	3.85	2.79	4.97	2.09	7.38	2.38	5.47	11	10	5	0	0	3.45	5.25	2366	2	N	Y
irritation	E	3.44	3.07	5.19	2.58	7.04	1.52	5.27	10	7	4	1	2	2.65	2.80	2711	2	Y	Y
jealousy	E	4.03	2.68	5.53	2.20	7.31	2.47	5.71	8	6	3	0	0	3.20	2.55	1409	2	Y	Y
joy	E	4.94	3.24	6.94	8.06	9.37	3.16	6.09	3	3	1	10	10	1.35	1.75	142	1	Y	N
love	E	4.47	2.85	8.15	8.65	12.02	4.76	6.91	4	3	1	12	9	1.15	1.35	1304	1	Y	Y
malice	E	2.63	2.79	5.39	2.44	7.13	1.79	4.03	6	5	2	0	6	1.95	1.65	2283	1	Y	Y
misery	E	4.18	2.74	3.27	2.15	7.97	2.72	5.16	6	6	3	0	1	1.95	2.25	2592	2	Y	Y
optimism	E	2.97	2.06	6.77	8.47	6.78	1.90	6.07	8	8	4	1	0	3.40	3.65	1633	2	N	Y
panic	E	3.74	3.87	5.69	2.27	8.79	3.05	5.97	5	5	2	2	1	1.90	1.80	2098	1	Y	Y
passion	E	3.70	3.19	7.58	8.13	8.76	3.00	6.03	7	4	2	0	3	2.15	1.75	2156	2	Y	Y
pity	E	3.34	2.32	3.52	3.23	8.50	3.08	5.65	4	4	2	3	16	1.70	1.15	1233	1	N	Y
rage	E	5.24	3.44	6.49	2.06	8.75	2.76	5.27	4	3	1	13	22	1.10	1.35	1702	1	Y	Y
remorse	E	2.88	2.49	3.55	3.79	6.60	2.19	4.50	7	6	2	0	0	2.60	2.30	2129	1	Y	Y
resentment	E	2.94	2.29	4.80	1.94	7.20	1.89	5.39	10	10	3	0	0	3.45	3.80	2951	1	N	Y
sadness	E	4.13	2.76	3.77	1.85	7.11	2.39	6.23	7	6	2	2	2	2.35	1.90	1698	2	N	Y
satisfaction	E	3.14	2.40	6.69	7.79	8.87	2.55	6.00	12	10	4	0	0	4.15	4.40	2494	3	Y	Y
shame	E	3.41	2.80	3.97	2.29	9.33	3.33	5.65	5	3	1	7	17	1.45	1.00	1282	1	N	Y
shock	E	3.75	3.45	5.81	3.25	9.31	3.17	5.61	5	3	1	6	25	1.55	1.00	1020	1	Y	Y
sorrow	E	3.60	2.77	3.66	2.06	7.32	2.55	4.79	6	4	2	2	2	2.10	1.85	1389	1	Y	Y
suspicion	E	3.23	3.00	5.20	2.77	7.88	2.57	5.24	9	7	3	0	0	3.60	2.75	1796	2	Y	Y
terror	E	4.35	2.85	5.79	1.68	9.06	2.66	5.21	6	4	2	0	9	2.50	1.50	3206	2	Y	Y
woe	E	2.71	2.50	4.28	2.42	6.66	2.08	3.83	3	2	1	11	32	1.25	1.00	242	1	Y	Y
accent	A	3.62	3.60	5.53	5.63	8.56	2.79	5.88	6	6	2	2	0	1.85	2.20	2031	1	Y	Y
advice	A	2.88	2.35	4.09	6.23	10.70	3.39	6.09	6	5	2	1	1	2.30	2.40	1163	1	Y	Y
agility	A	3.66	3.52	5.94	6.78	7.11	1.38	4.78	7	7	4	1	1	2.60	2.90	1404	2	N	Y
algebra	A	5.18	4.39	2.32	3.84	7.83	1.96	5.40	7	7	3	0	0	3.10	4.05	1375	1	N	Y
apology	A	3.00	3.06	4.50	6.09	8.33	2.91	5.94	7	7	4	0	0	2.70	2.60	893	2	Y	Y
aroma	A	3.59	3.51	5.53	6.58	6.69	1.85	4.71	5	5	3	0	0	1.95	2.70	2001	1	N	Y

conformity	A	3.14	2.80	3.69	3.62	6.99	1.30	4.46	10	10	4	0	0	3.30	3.60	1733	3	Y	Y
connection	A	4.03	3.75	5.67	6.56	10.88	3.14	6.00	10	7	3	2	1	2.15	1.95	2672	2	N	Y
courtship	A	3.88	4.12	5.26	6.69	5.89	1.78	3.39	9	7	2	0	0	3.80	2.95	1302	2	N	Y
debt	A	4.20	4.34	3.34	1.97	9.56	2.86	5.27	4	3	1	2	22	1.75	1.00	962	1	N	Y
deed	A	3.29	3.94	4.00	5.36	7.59	2.68	4.70	4	3	1	13	28	1.30	1.00	2652	1	N	Y
destiny	A	2.61	2.00	6.14	6.57	8.44	3.07	5.27	7	7	3	1	1	2.35	2.65	3747	2	Y	Y
drama	A	3.74	4.03	5.65	4.47	8.46	3.01	6.13	5	5	2	0	0	1.90	1.95	1533	1	Y	Y
duty	A	3.06	3.06	3.82	5.03	9.62	3.42	5.38	4	5	2	1	11	1.90	1.45	668	1	Y	N
ego	A	3.26	2.46	4.31	3.03	8.73	2.58	5.07	3	3	2	2	3	1.85	1.90	398	1	N	Y
errand	A	3.94	4.44	3.29	5.10	5.72	2.31	5.41	5	5	2	1	1	2.55	1.90	3019	1	N	Y
fame	A	4.06	2.68	5.84	6.00	8.71	2.65	5.97	4	3	1	14	23	1.15	1.00	1038	1	Y	Y
fashion	A	5.52	4.09	6.32	6.68	9.50	2.98	6.31	6	4	2	0	5	2.50	1.55	1896	1	N	Y
fate	A	2.44	1.79	5.50	5.32	9.05	3.14	5.41	4	3	1	17	31	1.00	1.00	3001	1	Y	Y
gesture	A	4.91	3.97	5.09	5.88	7.58	2.58	5.70	7	5	2	0	1	2.40	2.60	2650	1	Y	Y
gist	A	2.06	2.40	3.56	5.67	6.95	1.76	5.03	4	4	1	6	13	1.65	1.80	2301	1	Y	Y
gossip	A	4.00	3.00	5.57	3.59	7.51	2.60	6.14	6	5	2	0	0	2.90	2.35	1000	1	Y	Y
habit	A	3.47	3.76	3.44	4.71	8.78	2.87	6.02	5	5	2	0	2	1.95	1.85	1236	1	Y	Y
hallucination	A	4.00	3.38	5.06	3.71	6.42	1.96	5.03	13	10	5	0	0	4.05	4.20	2686	2	Y	Y
identity	A	3.09	2.22	5.53	6.18	9.44	2.82	5.50	8	8	4	1	0	2.80	3.40	2565	2	N	Y
illusion	A	3.41	2.94	6.00	4.74	8.52	2.63	5.06	8	5	3	1	1	2.45	1.95	1989	2	Y	Y
immunity	A	2.67	3.47	4.57	6.49	7.62	2.34	4.91	8	8	4	1	0	3.00	2.95	1127	2	N	Y
impulse	A	2.67	3.03	6.64	5.55	8.08	2.43	5.39	7	6	2	0	0	2.50	2.45	943	2	Y	N
inspection	A	3.65	3.47	4.91	3.97	8.26	2.48	5.24	10	8	3	0	0	2.60	2.70	2843	3	Y	Y
irony	A	2.30	2.50	4.24	4.65	8.03	2.37	5.39	5	5	3	2	0	1.85	1.90	1932	2	Y	Y
magic	A	4.94	2.17	6.58	6.69	10.78	3.43	6.06	5	5	2	1	1	1.90	2.60	1334	1	Y	Y
necessity	A	2.69	2.27	5.81	5.16	8.47	2.17	5.30	9	8	4	0	0	3.60	2.90	1981	2	Y	Y
nutrition	A	4.12	4.50	5.24	6.94	8.48	1.69	5.71	10	7	3	0	0	2.90	3.10	2480	2	N	Y
obedience	A	4.06	3.27	4.62	5.12	8.49	1.95	6.00	9	8	4	0	1	3.65	3.50	1960	2	Y	Y
occasion	A	4.00	3.71	6.17	6.74	8.80	2.93	5.71	8	5	3	0	0	2.70	2.40	1849	1	Y	Y
penalty	A	4.00	4.13	4.09	3.26	9.15	2.67	5.10	7	6	3	0	0	2.85	2.80	1841	2	N	Y
personality	A	2.19	2.30	6.03	6.88	9.42	2.91	6.41	11	9	5	0	0	3.60	4.85	2480	3	Y	Y
poetry	A	5.00	4.15	4.46	6.77	9.09	2.83	5.25	6	6	3	0	1	2.30	2.65	979	2	N	Y
pursuit	A	3.77	3.24	5.85	5.79	8.42	2.56	4.79	7	5	2	0	3	2.60	2.00	1115	2	Y	Y
ration	A	3.51	3.97	3.24	4.00	6.31	1.83	4.06	6	5	2	2	2	1.75	1.70	3836	1	Y	Y
reflex	A	3.97	4.09	4.94	5.74	7.16	2.07	5.44	6	7	2	1	1	2.40	2.70	1896	2	Y	Y
reminder	A	3.34	3.88	4.03	5.53	8.17	2.29	5.85	8	7	3	1	1	2.25	2.50	3478	3	Y	Y
reunion	A	3.97	3.48	5.38	6.67	7.98	2.65	5.03	7	7	3	0	0	2.70	3.45	2530	3	Y	Y

revolution	A	4.73	3.33	6.47	5.71	9.51	2.80	4.74	10	8	4	2	2	2.35	2.45	2221	2	Y	Y
riddle	A	3.79	3.50	5.29	4.97	7.18	2.37	4.91	6	4	2	3	9	1.70	1.20	1531	1	Y	Y
rumour	A	3.03	2.66	3.00	3.26	8.83	2.73	6.00	6	4	2	2	7	1.95	1.70	1253	1	Y	Y
semester	A	2.59	4.03	3.94	5.10	8.07	2.55	5.52	8	7	3	0	0	2.95	3.00	3342	1	N	Y
style	A	4.26	3.03	5.91	6.44	10.67	3.19	6.00	5	4	1	2	9	1.80	1.60	1944	1	Y	Y
tact	A	2.37	2.60	3.68	5.76	6.63	1.76	4.42	4	4	1	6	16	1.55	1.00	1493	1	Y	Y
therapy	A	4.40	3.85	4.74	6.24	9.19	2.98	5.76	7	6	3	0	0	2.95	2.50	2217	2	Y	Y
vision	A	3.97	3.52	4.85	5.91	9.93	3.09	6.55	6	4	2	0	1	2.00	1.95	2470	2	Y	Y
voyage	A	5.42	4.35	6.63	6.69	7.73	2.49	4.24	6	4	2	0	0	2.45	2.50	503	1	Y	Y
alligator	C	6.68	6.55	5.66	4.36	6.58	2.25	5.27	8	7	4	0	0	3.55	2.90	2145	1	Y	Y
ambulance	C	6.89	6.77	5.41	4.72	7.24	3.06	6.52	9	9	3	0	0	3.20	3.15	1433	2	Y	Y
amplifier	C	5.23	5.56	5.90	5.87	7.97	1.23	4.21	9	8	4	1	1	3.45	3.70	1758	3	Y	Y
aspirin	C	5.82	5.90	3.85	5.03	7.12	2.60	6.00	7	6	2	0	0	2.70	2.60	2280	1	N	Y
asteroid	C	5.68	6.03	6.00	4.56	7.47	1.94	4.62	6	7	3	0	0	3.10	3.85	2702	2	Y	Y
auditorium	C	6.06	5.06	4.38	5.31	7.37	1.91	5.24	10	9	5	0	0	4.35	3.90	1395	2	N	Y
bacteria	C	4.82	5.62	4.06	3.47	8.44	2.19	5.41	8	8	4	0	0	2.80	3.30	2562	2	N	Y
banjo	C	6.32	6.49	4.64	5.61	6.93	1.92	4.59	5	5	2	0	0	2.30	2.70	1114	1	N	Y
beer	C	6.71	6.63	5.18	5.54	10.12	3.59	6.62	4	3	1	12	40	1.30	1.00	2630	1	Y	Y
blender	C	6.12	6.00	4.09	5.53	6.88	1.94	5.28	7	6	2	3	3	1.80	1.75	3156	2	Y	Y
bracelet	C	6.47	6.74	5.15	6.03	6.09	2.60	5.71	8	7	2	0	0	2.95	2.80	1898	2	Y	Y
camel	C	6.63	6.81	4.47	5.52	8.12	2.41	4.79	5	4	2	1	6	1.85	1.75	1556	1	N	Y
casino	C	6.72	6.60	5.44	4.43	8.44	3.02	4.55	6	6	3	1	0	2.00	2.90	2529	1	N	Y
castle	C	6.76	6.50	5.88	6.07	9.36	3.04	5.93	6	4	2	1	10	1.85	1.50	2131	1	Y	Y
clock	C	6.53	6.50	4.18	5.36	10.11	3.48	5.16	5	4	1	8	13	1.45	1.15	910	1	Y	Y
commander	C	5.62	5.31	5.09	5.17	9.54	3.28	4.39	9	7	3	1	1	2.35	2.40	2470	3	Y	Y
compass	C	6.42	6.62	5.97	6.65	7.37	2.52	5.08	7	6	2	0	1	2.45	1.90	2639	1	Y	Y
corn	C	6.29	6.35	3.41	5.40	8.51	2.86	5.29	4	4	1	11	20	1.20	1.00	1900	1	Y	Y
cotton	C	6.03	6.44	3.56	5.94	8.65	2.86	5.64	6	4	2	0	4	2.00	1.80	2004	1	Y	Y
curry	C	6.53	6.53	5.09	5.85	7.66	2.10	6.25	5	4	2	4	5	1.75	1.75	806	1	Y	Y
dictionary	C	6.21	6.39	3.12	6.00	9.51	2.28	5.59	10	9	4	0	0	3.55	3.75	2469	2	N	Y
dinosaur	C	6.82	6.37	5.77	5.40	7.78	2.31	5.58	8	7	3	0	0	3.75	3.15	1853	2	Y	Y
emerald	C	6.16	5.90	5.77	6.27	8.74	2.12	4.59	7	7	3	0	0	2.75	2.90	2570	1	Y	Y
employee	C	5.57	5.60	3.65	5.97	9.31	2.77	6.21	8	6	3	2	3	2.55	2.20	761	2	Y	Y
envelope	C	6.49	6.43	2.94	5.15	9.29	2.71	5.88	8	7	3	0	0	3.10	3.65	1652	1	Y	N
fortress	C	6.09	6.18	5.28	5.39	8.03	2.24	4.19	8	7	2	0	0	2.85	2.60	2377	2	N	Y
harmonica	C	6.65	6.77	4.09	6.06	6.51	1.95	4.79	9	9	2	0	0	3.35	3.75	1985	2	Y	Y
iceberg	C	5.97	6.18	3.46	5.53	7.73	2.41	5.52	7	5	3	0	0	2.95	2.60	1857	2	Y	Y

jungle	C	6.36	6.49	5.79	5.56	8.61	3.06	5.32	6	5	2	4	3	1.75	2.25	1961	1	Y	Y
lace	C	4.54	5.59	4.88	5.83	8.10	2.28	6.35	4	3	1	12	27	1.10	1.00	2401	1	N	Y
leopard	C	6.82	6.73	5.75	5.64	6.55	2.44	5.28	7	5	2	2	6	2.65	1.60	1517	1	Y	Y
magazine	C	6.48	6.74	5.12	5.38	11.01	3.23	5.89	8	7	3	0	0	2.95	3.70	1782	1	Y	Y
medallion	C	6.27	6.11	4.42	5.94	6.72	1.84	3.40	9	8	4	0	0	3.45	3.40	2663	2	N	Y
microwave	C	6.62	6.71	3.59	5.57	8.38	2.30	6.44	9	8	3	0	0	3.95	4.40	1280	2	Y	Y
mountain	C	6.83	6.86	5.77	6.16	10.46	3.26	6.30	8	5	2	1	2	2.55	2.55	2458	1	Y	Y
ocean	C	6.50	6.57	5.64	6.47	9.31	3.19	6.00	5	3	2	0	1	1.95	1.70	1895	1	Y	Y
olive	C	6.62	6.18	4.79	5.70	7.99	2.58	6.27	5	4	2	1	0	1.80	1.90	1705	1	N	Y
parcel	C	6.31	5.79	5.43	6.00	7.09	1.86	6.03	6	5	2	1	2	1.90	1.75	1572	1	Y	Y
pearl	C	6.57	6.63	5.56	6.58	9.28	2.90	5.25	5	3	1	1	23	1.70	1.35	1681	1	Y	Y
pianist	C	6.29	6.23	5.11	6.19	6.81	1.94	5.91	7	7	3	0	0	2.75	2.60	2211	2	Y	Y
projector	C	6.03	6.29	3.63	5.31	7.23	1.80	5.57	9	8	3	1	2	2.55	2.70	1276	3	N	Y
radiator	C	6.36	6.40	3.47	5.74	6.96	2.02	5.78	8	7	4	0	0	2.55	3.05	2208	3	N	Y
raven	C	6.64	6.58	4.55	4.69	8.13	2.20	4.36	5	4	2	2	3	1.75	1.75	2303	1	N	Y
razor	C	6.48	6.65	5.64	4.41	8.09	2.55	5.68	5	4	2	0	7	2.00	1.75	1502	2	Y	Y
scooter	C	6.36	6.59	4.44	5.28	6.61	2.15	4.71	7	5	2	2	4	2.00	1.70	2623	2	Y	Y
shark	C	6.81	6.76	6.05	3.82	8.09	2.88	5.16	4	4	1	8	8	1.50	1.40	1419	1	Y	Y
spine	C	6.77	6.24	5.71	6.14	8.55	2.47	6.21	5	4	1	8	15	1.40	1.20	2705	1	N	Y
stallion	C	6.21	6.44	5.88	5.97	6.89	2.22	4.15	8	7	3	1	2	2.70	2.60	2855	1	Y	Y
television	C	6.09	5.79	4.53	5.71	10.02	3.24	5.28	10	8	4	0	0	3.70	4.55	2485	3	Y	Y
triangle	C	6.50	5.35	3.54	5.03	8.60	2.34	6.09	8	7	3	0	0	2.70	3.45	2419	2	Y	N
turtle	C	6.03	6.06	4.15	5.57	8.41	2.94	5.70	6	4	2	1	11	1.90	1.70	1386	1	Y	Y
walnut	C	6.33	6.21	3.09	5.36	8.14	2.00	4.53	6	6	2	0	0	2.65	3.55	998	2	N	Y

Appendix B

Matching information for emotion, non-emotion abstract and concrete word sets.

Two-sample t-tests were conducted between the emotion set and other word sets to establish a significant difference on valence and arousal, and between the concrete set and other sets on imageability and concreteness. All other analyses were intended to determine that sets were not significantly different and used Bayesian t-test comparisons for which the number indicates Bayes Factor (BF). A Bayes Factor of less than 1 supports the hypothesis that the sets do not differ. 0.33-1.00 is anecdotal evidence, 0.10-0.33 is moderate evidence, 0.03-0.10 is strong evidence, 0.01-0.03 is very strong evidence, < 0.01 is extreme evidence. For the concreteness matched-set, emotion and non-emotion abstract sets have a Bayes Factor of 3.419, indicating moderate evidence that the sets differ. For the imageability-matched set, emotion and non-emotion abstract sets have a Bayes Factor of >100 indicating extreme evidence that the sets differ.

Log Frequency (HAL) = Written Frequency, Log Frequency (SUBT) = Spoken Frequency, Orthographic N = Orthographic neighbourhood size, Phonological N = phonological neighbourhood size, OLD = orthographic Levenshtein distance, PLD = phonological Levenshtein distance, Mean Bigram Freq = mean bigram frequency.

CONCRETENESS MATCHED SET (n=36)						
	MEAN (STANDARD DEVIATION)			STATISTICAL COMPARISONS		
	Emotion	Non-Emotion Abstract	Concrete	Emotion vs. Non-Emotion Abstract	Emotion vs. Concrete	Non-Emotion Abstract vs. Concrete
Imageability	3.99 (0.68)	3.54 (0.83)	6.36 (0.39)	t(70) = -2.51, p = .100	t(70) = 18.20, p < .001	t(70) = -18.50, p<.001
Concreteness	3.05 (0.40)	3.14 (0.66)	6.34 (0.43)	t(70) = 0.70, p = .489; BF = 0.300	t(70) = 33.68, p < .001	t(70) = -24.44, p<.001

Arousal	5.72 (1.22)	5.07 (1.05)	4.95 (0.94)	t(70) = 2.01, p = .041	t(70) = -2.97, p = .004	BF = 0.271
Valence	4.38 (2.79)	5.41 (1.01)	5.58 (0.64)	t(70) = -2.41, p = .019	t(70) = 2.51, p = .015	BF = 0.331
				BF	**BF**	**BF**
Log Frequency (HAL)	8.26 (1.13)	8.41 (1.13)	8.25 (1.28)	0.275	0.243	0.276
Log Frequency (SUBT)	2.59 (0.60)	2.60 (0.51)	2.58 (0.54)	0.243	0.243	0.243
Familiarity	5.55 (0.72)	5.47 (0.62)	5.49 (0.65)	0.277	0.258	0.246
No. Letters	7.08 (2.58)	6.78 (2.15)	6.92 (1.61)	0.277	0.255	0.253
No. Phonemes	5.83 (2.22)	5.78 (1.84)	5.78 (1.73)	0.244	0.245	0.243
No. Syllables	2.36 (0.96)	2.53(1.08)	2.39 (0.84)	0.298	0.245	0.285
Orthographic N	2.25 (4.05)	1.81 (3.70)	1.78 (3.12)	0.269	0.277	0.243
Phonological N	5.25 (9.64)	3.81 (7.08)	4.72 (8.22)	0.305	0.250	0.271
OLD	2.49 (0.93)	2.35 (0.69)	2.49 (0.75)	0.305	0.243	0.317
PLD	2.35 (0.92)	2.34 (0.85)	2.44 (0.98)	0.243	0.262	0.270
Mean Bigram Freq	1895 (770)	1929 (830)	1952 (578)	0.246	0.257	0.245
No. Morphemes	1.64 (0.68)	1.64 (0.72)	1.53 (0.65)	0.243	0.301	0.297

IMAGEABILITY-MATCHED SET (n=50)						
	MEAN (STANDARD DEVIATION)			STATISTICAL COMPARISONS		
	Emotion	Non-Emotion Abstract	Concrete	Emotion vs. Non-Emotion Abstract	Emotion vs. Concrete	Non-Emotion Abstract vs. Concrete
Imageability	3.80 (0.65)	3.67 (0.84)	6.28 (0.49)	t(98) = - 0.81, p = .419; BF = 0.280	t(98) = 21.55, p < .001	t(98) = -19.02, p < .001
Concreteness	2.86 (0.46)	3.36 (0.74)	6.30 (0.42)	t(98) = 4.08, p < .001	t(98) = 39.10, p < .001	t(98) = -42.46, p < .001
Arousal	5.66 (1.24)	4.92 (1.06)	4.83 (0.89)	t(98) = -3.17, p=.002	t(98) = -3.81, p<.001	BF = 0.230
Valence	4.48 (2.79)	5.43 (1.23)	5.52 (0.66)	t(98)= 2.20, p=.030	t(98) = 2.56, p=.012	BF = 0.232
				BF	**BF**	**BF**
Log Frequency (HAL)	8.13 (1.21)	8.33 (1.23)	8.10 (1.16)	0.283	0.212	0.314
Log Frequency (SUBT)	2.51 (0.65)	2.52 (0.51)	2.48 (0.52)	0.212	0.215	0.224
Familiarity	5.50 (0.78)	5.38 (0.65)	5.37 (0.73)	0.284	0.288	0.211
No. Letters	7.06 (2.37)	6.80 (2.13)	6.86 (1.74)	0.245	0.234	0.213
No. Phonemes	5.90 (2.30)	5.8 (1.83)	5.84 (1.81)	0.216	0.213	0.212
No. Syllables	2.46 (1.15)	2.6 (1.05)	2.46 (0.97)	0.253	0.211	0.261
Orthographic N	2.32 (3.94)	1.74 (3.58)	1.80 (3.22)	0.274	0.266	0.212

Phonological N	5.12 (8.97)	3.76 (7.50)	4.46 (8.05)	0.285	0.225	0.231
OLD	2.52 (0.88)	2.41 (0.68)	2.47 (0.78)	0.259	0.221	0.224
PLD	2.46 (1.13)	2.40 (0.85)	2.46 (0.94)	0.219	0.211	0.221
Mean Bigram Freq	1849 (724)	1957 (811)	1975 (583)	0.263	0.318	0.219
No. Morphemes	1.60 (0.67)	1.6 (0.67)	1.56 (0.67)	0.211	0.219	0.220

Appendix C

Raw Scores

Numbers indicate total score for accuracy data and mean response times. Data is presented for each word set: emotion, non-emotion abstract and concrete, organised by task within the concreteness-matched and imageability-matched sets. Raw scores are presented for non-target words: nonwords from the lexical decision task and distractor words from the identification recall task. RT = response time, ID Recall = Identification recall, NE Abstract = Non-Emotion Abstract.

		BBK	DTR	MMR	JDC	SPN	TLK	OLG	JOG	MLK	Group Average
CONCRETENESS-MATCHED SET (n=36)											
Free Recall	Emotion	5	5	8	11	12	5	3	6	5	6.7
	NE Abstract	5	6	4	3	7	6	1	6	5	4.8
	Concrete	2	8	2	4	6	7	1	5	1	4.0
ID Recall	Emotion	33	23	33	34	30	34	32	30	26	30.6
	NE Abstract	32	18	33	22	27	29	31	27	26	27.2
	Concrete	28	28	30	30	30	33	31	28	27	29.4
Lex Dec Accuracy	Emotion	36	36	36	36	36	36	35	36	28	35.0
	NE Abstract	36	34	34	36	35	36	35	35	28	34.3
	Concrete	36	36	36	35	36	36	36	35	30	35.1
Lex Dec RT	Emotion	1530 (319)	1051 (191)	1229 (319)	1155 (382)	1599 (345)	1140 (313)	1357 (291)	1520 (316)	2069 (527)	1406
	NE Abstract	1462 (264)	1201 (351)	1374 (351)	1187 (305)	1800 (478)	1016 (232)	1581 (382)	1716 (383)	2109 (451)	1494
	Concrete	1598 (331)	1191 (236)	1444 (365)	1175 (313)	1646 (415)	1085 (271)	1542 (335)	1667 (405)	2045 (525)	1488
Reading Accuracy	Emotion	34	33	31	30	25	27	28	24	20	28.0
	NE Abstract	34	34	30	26	18	33	24	29	13	26.8
	Concrete	34	34	34	25	27	35	29	35	24	30.8
Reading RT	Emotion	921 (131)	864 (110)	1002 (391)	963 (135)	1368 (506)	1323 (369)	1026 (283)	1003 (293)	1285 (380)	1084
	NE Abstract	932 (101)	939 (151)	1078 (252)	1113 (225)	1429 (476)	1322 (339)	1019 (146)	895 (166)	1090 (241)	1091
	Concrete	961 (90)	888 (161)	1013 (135)	1085 (158)	1610 (652)	1348 (458)	1017 (233)	933 (193)	1092 (193)	1105

IMAGEABILITY-MATCHED SET (n=50)											
Free Recall	Emotion	8	8	9	11	16	6	4	9	7	8.7
	NE Abstract	6	7	5	7	6	7	1	8	5	5.8
	Concrete	5	9	4	5	8	10	2	6	3	5.8
ID Recall	Emotion	45	34	44	46	42	45	43	42	37	42.0
	NE Abstract	43	27	46	34	37	44	44	38	36	38.8
	Concrete	40	37	43	38	39	45	40	35	36	39.2
Lex Dec Accuracy	Emotion	50	48	50	50	49	50	47	50	39	48.1
	NE Abstract	49	47	50	49	48	48	47	50	36	47.1
	Concrete	49	49	50	48	50	50	50	48	42	48.4
Lex Dec RT	Emotion	1496 (323)	1103 (295)	1271 (327)	1116 (333)	1778 (471)	1152 (294)	1394 (283)	1501 (350)	2098 (531)	1434
	NE Abstract	1506 (299)	1224 (321)	1448 (386)	1201 (354)	1797 (464)	1029 (178)	1539 (373)	1767 (403)	2025 (409)	1504
	Concrete	1600 (353)	1197 (245)	1526 (477)	1223 (351)	1597 (413)	1102 (295)	1511 (346)	1676 (461)	2014 (474)	1494
Reading Accuracy	Emotion	47	46	43	42	31	37	32	33	24	37.2
	NE Abstract	44	47	36	30	20	46	35	34	18	34.4
	Concrete	42	48	46	35	37	49	42	46	35	42.2
Reading RT	Emotion	918 (122)	884 (134)	929 (259)	1020 (184)	1329 (369)	1325 (361)	962 (194)	969 (277)	1196 (340)	1059
	NE Abstract	943 (94)	981 (218)	1113 (295)	1100 (220)	1635 (631)	1335 (382)	1024 (170)	970 (241)	1108 (228)	1134
	Concrete	957 (126)	914 (148)	1000 (132)	1083 (153)	1675 (593)	1485 (579)	1049 (273)	940 (183)	1105 (206)	1134
NON-TARGET WORD SETS											
Lex Dec Nonwords (n=156)		153	153	154	156	145	140	142	145	148	148.4
Lex Dec Nonwords RT		1987	1403	1552	1282	1843	1526	1837	1779	2218	1714
Recall ID Emotion Distractors (n=26)		24	23	21	24	25	13	18	11	21	20.0
Recall ID NE Abstract Distractors (n=26)		25	25	22	21	24	14	18	17	18	20.4
Recall ID Concrete Distractors (n=26)		26	25	26	21	26	15	22	18	24	22.6

Appendix D

Statistical Results

Free recall, identification recall (ID Recall), lexical decision (Lex Dec) and reading accuracy are Fisher Exact Tests. Lexical decision and reading response times (Lex Dec RT; Reading RT) are independent t-tests. Group are Stouffer's tests. Unless otherwise stated numbers denote two-tailed p-values.

		BBK	**DTR**	**MMR**	**JDC**	**SPN**	**TLK**	**OLG**	**JOG**	**MLK**	**Group**
Concreteness-Matched Set	Free Recall	1.000	1.000	.343	.036	.285	1.000	.614	1.000	1.000	z = -1.70, p = .088
	ID Recall	1.000	.341	1.000	.001	.563	.151	1.000	.563	1.000	z = -2.87, p = .004
	Lex Dec Accuracy	1.000	.493	.493	1.000	1.000	1.000	1.000	1.000	1.000	z = -0.39, p = .697
	Lex Dec RT	t(70) = 0.99, p = .327	t(66) = 2.19, p = .033	t(68) = 1.81, p = .076	t(65) = .039, p = .702	t(58) = 1.84, p = .065	t(69) = 1.89, p = .062	t(64) = 2.66, p = .009	t(65) = 2.28, p = .026	t(31) = 0.23, p = .817	z = -2.78, p =.005
	Reading Accuracy	1.000	1.000	1.000	.396	.149	.111	.430	.285	.155	z = -0.71, p = .477
	Reading RT	t(65) = 0.52, p = .602	t(65) = 2.25, p=.028	t(39) = 0.87, p = .389	t(50) = 2.82, p = .008	t(39) = 0.54, p = .591	t(55) = .11, p = .913	t(43) = .04, p = .966	t(40) = 1.27, p = .210	t(27) = 1.50, p = .145	z = -1.27, p = .205
Imageability-Matched Set	Free Recall	.774	1.000	.390	.437	.031	1.000	.361	1.000	.760	z = -2.37, p = .018
	ID Recall	.760	.221	.740	.006	.329	1.000	1.000	.456	1.000	z = -2.25, p = .025
	Lex Dec Accuracy	1.000	1.000	1.000	1.000	1.000	.477	1.000	1.000	.646	z = -0.52, p = .604
	Lex Dec RT	t(97) = 0.16, p = .871	t(92) = 1.89, p = .062	t(95) = 2.44, p = .017	t(93) = 1.20, p = .232	t(88) = 0.10, p =.918	t(94) = 2.49, p=.015	t(90) = 2.09, p = .038	t(89) = 3.35, p = .001	t(43) = 0.51, p = .604	z = -1.73, p = .084
	Reading Accuracy	.487	1.000	.142	.015	.047	.034	.672	1.000	.314	z = -1.70, p = .088
	Reading RT	t(87) = 1.10, p = .275	t(86) = 2.51, p = .014	t(49) = 2.37, p = .024	t(64) = 1.61, p = .126	t(46) = 2.11, p = .062	t(73) = 0.11, p=.908	t(58) = 1.3, p = .195	t(54) = 0.01, p = .991	t(34) = 0.89, p = .367	z = -3.76, p < .001

Appendix E

Regression Results

Dependent variables are group level outcome data. "RT" refers to response time. Parameters included in the base model are: imageability, valence, concreteness, arousal, familiarity, spoken and written frequency, length (letters, phonemes, syllables), orthographic neighbourhood size, phonological neighbourhood size, orthographic Levenshtein distance, phonological Levenshtein distance, mean bigram frequency, number of morphemes. Variance inflation factors (VIF) were less than 2.5 for both Imageability and Valence for every model.

Dependent Variable	**ANOVA**	**Imageability**	**Valence**
Free Recall (items recalled)	R^2 = 0.167, F(15,88) = 1.180, p = .302	β = .188, t(102) = 1.44, p = .153	β = .054, t(102) = 0.42, p = .679
Identification Recall (items correctly selected)	R^2 = 0.341, F(15,88) = 3.039, p < .001	β = .282, t(102) = 2.43, p = .017	β = .155, t(102) = 1.36, p = .178
Lexical Decision Accuracy	R^2 = 0.300, F(15,88) = 2.518, p = .004	β = .226, t(102) = 1.89, p =.062	β = .015, t(102) = 0.13, p = .897
Lexical Decision RT	R^2 = 0.288, F(15,88) = 2.376, p = .006	β = -.170, t(102) = -1.41, p = .162	β = .024, t(102) = 0.20, p = .842
Reading Accuracy	R^2 = 0.462, F(15,88) = 5.036, p < .001	β = -.035, t(102) = -0.34, p = .738	β = -.035, t(102) = -0.34, p = .735
Reading RT	R^2 = 0.270, F(15,88) = 2.149, p = .014	β = -.156, t(102) = -1.28, p = .204	β = -.015, t(102) = -0.12, p = .901

Appendix E of this thesis has been removed as it may contain sensitive/confidential content