CONCEPTUAL METAPHORS OF EMOTION IN SPOKEN LANGUAGE: GOOD IS UP IN SEMANTICS AND PROSODY

HAZEL K. GODFREY

A thesis
submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Masters of Science in Psychology

Victoria University of Wellington
2011

Abstract

Recent research on embodied cognition points to a role for the perceptuomotor system in conceptual representation. One way that the perceptuomotor system may be involved in conceptual representation is through metaphorical mappings, as described in Conceptual Metaphor Theory (Lakoff & Johnson, 1999). This theory accounts for the embodiment of abstract concepts with metaphoric mappings to perceptuomotor properties. Examples include INTELLIGENCE IS LIGHT (as in "that is a bright idea"), IMPORTANT IS BIG (as in "that is a big deal"), and INTIMACY IS CLOSENESS (as in "you are close to my heart"). The GOOD IS UP (as in "things are looking up") conceptual metaphor is the focus of this thesis. A prediction derived from Conceptual Metaphor Theory is that activation of the concept of "good" should automatically activate associated perceptuomotor processes, resulting in an attentional shift to upper visual space. Conversely activation of the concept "bad" should result in an attentional shift to lower visual space. There is experimental evidence for the existence of the GOOD IS UP conceptual metaphor. However, this past research has only assessed the validity of the GOOD IS UP conceptual metaphor with written emotion-related words. In order to paint an accurate picture of the nature of conceptual representation, both written and spoken language processing must be investigated.

The aim of this thesis was to determine whether the conceptual metaphor GOOD IS UP is activated by processing of spoken emotional words. Spoken language has two channels through which emotion can be conveyed; the semantic channel and the prosodic channel. This thesis assessed whether the GOOD IS UP conceptual metaphor was activated by emotional semantics and prosody separately. Semantically or prosodically valenced words were presented to participants. Positive and negative valence would be expected to elicit activation of the GOOD IS UP conceptual metaphor; thus GOOD IS UP congruent shifts in attention were expected. Following presentation of the spoken word, a visual target detection and identification task was completed to assess attention to upper and lower space. No metaphor congruent shifts in attention were observed, which suggests that the GOOD IS UP conceptual metaphor was not activated when words with semantic or prosodic emotion were processed. A thorough evaluation is provided of the differences between the previous studies, using written stimuli, and the current studies, using spoken stimuli. The discrepancies suggest that it is theoretically important to define the boundary

conditions under which evidence for conceptual metaphor congruent activation is (and is not) seen. Whether context is an important boundary condition especially needs to be considered. A multiple systems view of representation may need to be applied to Conceptual Metaphor Theory.

Acknowledgments

I would firstly like to thank my supervisor Dr. Gina Grimshaw. Thank you for your invaluable guidance. I really appreciate your feedback and all the time you put into this project.

Next I would like to thank my Mum and Dad for their endless encouragement and belief in me. Thank you also to my second family the Meade-Mumby-Walker's; especially to Kevin (for "tech support").

I wish to acknowledge Desiree Cheer for lending her voice to the recording of the stimuli. I also want to acknowledge the members of the Cognitive and Affective Neuroscience Lab at Victoria University for their support and proofreading.

Last (but definitely not least) thank you to Frances Bryson for your proofreading, friendship, and steadfast support.

Table of Contents

Abstract 2

Acknowledgments 4

Table of Contents 5

List of Tables 8

List of Figures 9

INTRODUCTION 10

Conceptual Metaphor Theory 11

Predictions derived from Conceptual Metaphor Theory 13

GOOD IS UP 14

Evidence for the GOOD IS UP conceptual metaphor 15

Pervasive metaphor 17

Spoken Language 18

Evolution 19

Complexity 20

The Current Studies 21

STUDY 1 25

Method 26

Participants 26

Stimuli and Apparatus 26

Procedure 27

Results and Discussion 30

Release Times 31

Press Times 33

Movement Times 33

STUDY 2 34

Method 35

Participants 35

Stimuli and Apparatus 35

Procedure 35

Results and Discussion 36

Prosodic Properties 36

Acoustic Properties 36

Semantic and Lexical Properties 39

STUDY 3 43

Method 43

Participants 43

Stimuli and Apparatus 43

Procedure 43

Results and Discussion 46

Release Times 49

Press Times 51

STUDY 4 53

Method 53

Participants 53

Stimuli and Apparatus 54

Procedure 54

Results and Discussion 55

Release Times 58

Press Times 62

GENERAL DISCUSSION 66

Methodology 67

Unlikely methodological explanations 67

Comparison to Meier and Robinson (2004) 69

Contextually dependent grounded cognition 71

Multiple systems 73

Grounded cognition evidence 73

Multiple systems in Conceptual Metaphor Theory 74

Explaining the current results 77

Spoken Language 79

Evolution 80

Complexity 80

Other Considerations 81

Time course 81

Dimensional versus categorical emotion 82

Conclusions 83

References 86

Appendix A 95

List of Tables

Table 1. *Mean (SD) release, press, and movement times (ms) for Study 1 by SOA, tone, and visual-field* 31

Table 2. *Prosodic and acoustic properties of the words used in Studies 3 and 4* 38

Table 3. *Emotional-semantic properties of the words used in Studies 3 and 4* 40

Table 4. *Lexical properties of the words used in Studies 3 and 4* 41

Table 5. *Mean (SD) subject release and press times (ms) for Study 3 by SOA, evaluation, and visual-field* 48

Table 6. *Mean (SD) item release and press times (ms) for Study 3 by SOA, meaning, and visual-field* 49

Table 7. *Mean (SD) subject release and press times (ms) for Study 4 by SOA, prosody-evaluation, and visual-field* 57

Table 8. *Mean (SD) item release and press times (ms) for Study 4 SOA, consensus-prosody, and visual-field* 58

List of Figures

Figure 1. Target-present trial procedure in Study 1 29

Figure 2. Release times for high- and low-tone trials for upper and lower visual-field targets in Study 1 32

Figure 3. Target-present trials in Study 3 44

Figure 4. Target-present trial procedure in Study 3 45

Figure 5. Subject release times for negative, neutral, and positively evaluated trials for upper and lower visual-field targets at the short and long SOA in Study 3 50

Figure 6. Item release times for negative, neutral, and positive semantics for upper and lower visual-field targets at the short and long SOA in Study 3 50

Figure 7. Subject press times for negative, neutral, and positively evaluated trials for upper and lower visual-field targets at the short and long SOA in Study 3 51

Figure 8. Item press times for negative, neutral, and positive semantics for upper and lower visual-field targets at the short and long SOA in Study 3 52

Figure 9. Target-present trial procedure in Study 4 55

Figure 10. Subject release times for trials evaluated as sad, neutral, and happy for upper and lower visual-field targets at the short and long SOA in Study 4 59

Figure 11. Item release times for sad, neutral, and happy prosody for upper and lower visual-field targets at the short and long SOA in Study 4 59

Figure 12. Subject release times by SOA and evaluation in Study 4 61

Figure 13. Subject release times by SOA and visual-field in Study 4 62

Figure 14. Subject press times for trials evaluated as sad, neutral, and happy prosody for upper and lower visual-field targets at the short and long SOA in Study 4 63

Figure 15. Item press times for sad, neutral, and happy prosody for upper and lower visual-field targets at the short and long SOA in Study 4 63

Figure 16. Subject press times by SOA and visual-field in Study 4 64

Conceptual metaphors of emotion in spoken language: GOOD IS UP in semantics and prosody

How concepts are represented in the mind has been the object of intense theorising and empirical investigation. As a result, many different theories have been developed to explain how representation is accomplished in the mind. The nature of conceptualisation, described by the different theories, is not a philosophy-free selection. The way we represent concepts and what is included in our representations is seen as key to what it means to be human (Johnson, 2007). It is not surprising, therefore, that some theories of representation are controversial.

Theories of conceptual representation can be divided into two broad dominant views. Proponents of the traditional view (e.g. Collins & Quillian, 1969; Katz & Fodor, 1963; Fodor, 1985), posit that conceptual representations are stored in their own independent system, which entails that there is no overlap with other systems (such as the perceptuomotor system) in the mind (Lakoff & Johnson, 1999; Winkielman, Niedenthal, & Oberman, 2008). The disconnection from the perceptuomotor system necessitates that the form of representation is symbolic and non-perceptual.

Proponents of the alternative view, grounded cognition (see Barsalou, 1999, 2008, 2010; Barsalou, Santos, Simmons, & Wilson, 2008; Gallese & Lakoff, 2005; Gibbs, 2006; Gibbs & Matlock, 2008; Johnson, 2007; Lakoff & Johnson, 1980, 1999; Niedenthal, Krauth-Gruber, & Ric, 2004; Niedenthal, 2007; Wilson, 2002; Winkielman et al., 2008), posit that conceptual representations are non-modular, that is, they are not instantiated in a separate representational system, but are distributed across the evolutionarily older perceptual-motor areas. Meaning is thus embodied. This is a controversial view (Barsalou, 2008; Grush, 2003; Haselarger, de Groot, & van Rappard, 2003). Grounded cognition theorists strongly propose that the mind uses the evolutionary older perceptual and motor systems to represent both concrete and abstract concepts. It is relatively easy to see how such an embodied representation system could work for concepts at the more concrete end of the concrete-abstract continuum, which have clear perceptuomotor components. For instance, in the grounded view the concept of an apple is not, as in a traditional semantic-network model, represented by a node connected to other nodes in an encapsulated representation system containing abstract information (such as an apple is a fruit, is red or green, and is juicy). Rather, areas of the perceptuomotor system, including visual, auditory, olfactory, tactile, taste, and motor movements, that were activated when experiencing an apple, are partially reactivated when the concept of an apple is retrieved (Barsalou, 1999).

Concrete concepts such as an apple have a solid real world counterpart which people can interact with and experience directly. It is harder to picture how such embodied grounding of concepts could work for more abstract concepts such as emotion-related concepts which do not have such intrinsic perceptuomotor properties as more concrete concepts do. In the traditional view there is debate about whether emotion and non-emotional cognition are independent, interact, or are integrated (Barnard, Duke, Byrne, & Davidson, 2007; Duncan & Feldman-Barrett, 2007; Gray, Braver, & Raichle, 2002; Lazarus, 1984; Leventhal & Scherer, 1987; Zajonc, 1984). Most traditional theories of representation assume some additional abstract aspect of meaning is attached to emotional concepts, as in the dimensional view (positive or negative; see Kövecses, 2000 and Niedenthal, 2008). Furthermore, in the case of semantic-network models, problems arise such as circularity in what defines meaning in a node; meaning is entirely achieved in the links from one node to another (Niedenthal, 2008). In the grounded view, parsimony is achieved by grounding representation of all concepts, not just concrete concepts, in the perceptuomotor systems. There are many instances in the literature that point to a role for embodied grounding in emotional concept representation (for example, Hauk, Johnsrude, & Pulvermüller, 2004; Havas, Glenberg, & Rinck, 2007; Havas, Glenberg, Gutowski, Lucarelli, & Davidson, 2010; Niedenthal, Winkielman, Mondillon, & Vermeulen, 2009; Willems, Labruna, D'Esposito, Ivry, & Casasanto, 2011). A larger number of theories have grounded cognition as the cornerstone of conceptual representation. The version of relevance to this thesis is Conceptual Metaphor Theory. According to conceptual metaphor theorists, the grounding problem for abstract emotion concepts is solved by grounding representations in the perceptuomotor system via metaphorical mappings.

Conceptual Metaphor Theory

Lakoff and Johnson (1980, 1999) developed Conceptual Metaphor Theory from the observation of three recurring effects: grounded cognition, unconscious processing, and the metaphoric nature of abstract thought. Like other grounded cognition theorists, proponents of Conceptual Metaphor Theory suggest that our conceptual system is not disembodied but is grounded; determined by the nature of our bodies, how we process the world through our perceptual system, and interact with it through our motor system (Lakoff & Johnson, 1999). Conceptual Metaphor Theory differs from other grounded cognition theories in that the primary focus is on the metaphoric mappings that are claimed to underlie abstract thought. These conceptual metaphors are deemed to be necessary to explain how abstract ideas, such as emotional concepts, are grounded.

The core principle of Conceptual Metaphor Theory is that metaphoric mappings, from source domains to target domains, underlie representation. The source domain is a perceptuomotor determined experience, for example brightness, verticality, or warmth. The target domain is a concept, for example happiness, dominance, or affection. The developmental origins of these conceptual metaphor mappings are a matter of debate, though most authors attribute the development of source-target mappings to repetitive co-activation of both domains (Grady, 1997 as cited in Lakoff & Johnson, 1999; Lakoff & Johnson, 1999; Tolaas, 1991). It is proposed that such repetitive co-activation is pervasive in development and results in the neural storage of many conceptual metaphors.

In the emotional domain, it is hypothesised that the source domain temperature is repetitively mapped onto the target domain affection (as in the body temperature observed during a hug between caregiver and child) to form the conceptual metaphor AFFECTION IS WARMTH; the source domain proximity is mapped onto the target domain intimacy (as in the proximity between the child and their caregivers) to form the conceptual metaphor INTIMACY IS CLOSENESS; the source domain smell is mapped onto the target domain evaluation (as in the negative evaluative response commonly paired with disgusting smells) to form the conceptual metaphor BAD IS STINKY; and the source domain verticality is mapped onto the target domain valence (as in the repeated pairing of the positive appearance of the caregiver from above the child), to form the conceptual metaphor GOOD IS UP/BAD IS DOWN[1] (Grady, 1997 as cited in Lakoff & Johnson, 1999).

Regardless of the origin of conceptual metaphors, theorists agree that they are used during linguistic processing; they are the representational system. Furthermore, as conceptual metaphors are formed early in development, through the strengthening

[1]From now on this will be referred to as GOOD IS UP. Lakoff and Johnson (1999) use the convention TARGET DOMAIN IS SOURCE DOMAIN to describe conceptual metaphors. I will also use this convention.

of neural connections between source and target domains, conceptual metaphors are activated and used for representation unconsciously and automatically. Conceptual metaphor theorists argue that these source-target domain mappings are not the result of shared linguistic conventional metaphors, such as “the sunny side is up”, but rather that the linguistic metaphors are the result of grounded source-target domain mappings (Gibbs, 2006; Lakoff & Johnson, 1999). Conceptual metaphors are thus embodied (Lakoff & Johnson, 1999).

In summary, according to Lakoff and Johnson (1980, 1999) conceptual metaphors are mappings between perceptuomotor source domains and conceptual target domains. These mappings develop through repetitive co-experience of the source and the target domains. More abstract thought, about domains such as emotion, would not be possible without conceptual metaphors. Notably, Conceptual Metaphor Theory is primarily a linguistic-philosophical theory. Conceptual metaphor theorists are concerned with how and why people categorise and process the world the way we do with the aim of answering philosophical questions about the nature of people and how to live (Johnson, 2007). However, Conceptual Metaphor Theory lends itself to empirical validation. Coming from an experimental psychological perspective, Meier and Robinson (2005) have derived three testable predictions from Conceptual Metaphor Theory to determine whether conceptual metaphors underlie representation for emotion-related concepts.

Predictions derived from Conceptual Metaphor Theory.

Meier and Robinson’s (2005) first prediction (consistency) is that, if emotion concepts are represented using grounded conceptual metaphors, like GOOD IS UP, then a processing advantage should be observed for stimuli that have properties consistent with the conceptual metaphor. For example, positive stimuli in the upper visual-field should be processed faster than positive stimuli in the lower visual-field. This prediction has been supported for several conceptual metaphors of emotion including GOOD IS UP (Meier & Robinson, 2004), POSITIVE IS BRIGHT (with manipulations and judgements of brightness; Meier et al., 2004), and DOMINANCE IS UP (measuring trait dominance and with manipulations of verticality; Robinson, Zabelina, Ode, & Moeller, 2008).

Meier and Robinson’s (2005) second prediction (congruency) is that, if emotion concepts are represented using conceptual metaphors, then activating target domain concepts (like emotion concepts) should activate the perceptuomotor source domain in a metaphor consistent manner. For example consistent with the conceptual metaphor GOOD IS UP, processing positive words should activate the perceptuomotor source domain of upward-verticality and direct attention to the upper space and processing negative words should activate downward-verticality and direct attention to the lower space. A metaphor congruent shift in visual attention would be observed in an advantage for upper visual-field targets over lower visual-field targets after evaluating words as positive, and in an advantage for lower visual-field targets over upper visual-field targets after evaluating words as negative. This congruency prediction has been supported for the conceptual metaphors GOOD IS UP (Meier and Robinson, 2004), and POSITIVE IS BRIGHT (with manipulations of and judgements of brightness; Meier, Robinson, Crawford, & Ahlvers, 2007).

Meier and Robinson's (2005) third prediction (automaticity) is that, if conceptual metaphor mapping is necessary for representation, then conceptual metaphor consistent source-target mappings should be present at automatic processing stages. For example, the shifting of attention to the upper visual-field after processing a positive word should occur after only a very short delay. This prediction has been supported for the conceptual metaphor POSITIVE IS BRIGHT (Meier et al., 2007).

GOOD IS UP

The GOOD IS UP conceptual metaphor is the focus of this thesis. The conceptual metaphor GOOD IS UP describes the mapping between the perceptuomotor source domain, verticality, and the conceptual target domain, valence. Speculation as to the development of the metaphor focuses on the repeated experiential co-occurrence between upper space, from the child's perspective, and appearance of parents and caregivers who provide nutrition and care; on the co-occurrence of being prone with being helpless; on the co-occurrence of erect posture with confidence and happiness and slumped posture with depression (Tolaas, 1991); and on the co-occurrence of death with being buried in the ground below (Crawford, 2009). Cross linguistic studies suggest that the GOOD IS UP conceptual metaphor is universal. For example, Luodonpää-Manni and Viimaranta (2010) examined the validity of metaphors that use the source domain, verticality, in Russian and French. They used dictionary sources to see if the conceptual metaphors listed by Lakoff and Johnson (1980) as being present for English speakers were descriptive of Russian and French speakers' source-target domain mappings. The analysis conducted by

Luodonpää-Manni and Viimaranta demonstrates that the verticality metaphor mapping the source domain verticality to GOOD and BAD[2] is a basic important metaphor across cultures.

Evidence for the GOOD IS UP conceptual metaphor.

Researchers have developed paradigms in which emotional valence and verticality of the stimuli are manipulated with the direct aim of testing the cognitive reality of the GOOD IS UP conceptual metaphor. Meier and Robinson (2004) presented positive and negative words in the upper or lower visual-field on a computer screen. Participants were required to evaluate the word as positive or negative by saying “positive” or “negative” out loud after the word was presented. Response times were faster when the emotional valence of the words matched the vertical position as predicted by the conceptual metaphor GOOD IS UP. That is, participants were faster to evaluate words as positive in the upper visual-field and as negative in the lower visual-field. This pattern of results is in line with Meier and Robinson’s (2005) first prediction of consistency, that if emotional concepts are represented using conceptual metaphors, then a processing advantage should be observed for stimuli that have properties consistent with the conceptual metaphor (in this case vertical position).

In Meier and Robinson’s (2004) second study, a similar result was found when participants evaluated an emotional word before completing a visual-attention task. As in their first study, the evaluation response was given orally using the valence labels “positive” or “negative”. After evaluating a centrally presented positive word, participants were faster to indicate whether a visual target was the letter *p* or *q* in the upper visual-field than in the lower visual-field. Conversely, after evaluating a centrally presented negative word participants were faster to discriminate between a *p* and *q* in the lower visual-field than in the upper visual-field. Thus, activating the conceptual metaphor GOOD IS UP shifted visual attention to the conceptual metaphor appropriate position. This is consistent with Meier and Robinson’s (2005) second prediction of metaphor congruent perceptual processing, that if emotion concepts are represented using conceptual metaphors, then activating target domain

---

[2] I will follow Luodonpää-Mannii and Viimaranta and call the mapping between verticality and emotion GOOD IS UP, rather than HAPPY IS UP, or POSITIVE IS UP. The name of the conceptual metaphor is not as important as the relevant source and target domains, verticality (upper and lower space) and dimensional valence (positive and negative).

concepts should activate the perceptuomotor source domain (and associated processing) in a metaphor congruent manner.

While Meier and Robinson (2005) were confident that the paradigms used in their 2004 study were appropriate for testing the cognitive reality of conceptual metaphors, patterns of responding consistent with the GOOD IS UP conceptual metaphor are also observed with paradigms that used more subtle manipulations of verticality. Casasanto (2008, as cited in Brookshire, Ivry, & Casasanto, 2010) and Brookshire et al. (2010) used tasks in which the shift between upper and lower target position was not so noticeable. In a spatial-interference antonym-judgement task, Casasanto presented participants with words positioned above fixation and below fixation. Participants were faster to say that the word pairs were antonyms (they had the opposite meaning) when the word pair positioning was consistent with the GOOD IS UP conceptual metaphor, that is, when the positive word was above fixation, and the negative word below, than when it was inconsistent. In a spatial-interference lexical decision task, Casasanto again presented participants with word pairs, one word of the pair was positioned above and one below fixation. One word of the pair was a real word, either positive or negative, and one was a non-word. Participants were faster to make a lexical decision when the real word of the pair was in the position consistent with the GOOD IS UP conceptual metaphor, that is, when the real positive word was presented above the non-word, and the real negative word below, than vice versa.

In the Casasanto (2008) studies there were stimuli in both the upper and lower visual-field on each trial. It was the positioning of the valenced word of the pair that was critical. Because both positions were filled on each trial, the vertical positioning of the valenced word was less salient. Yet speed of responding was consistent with the GOOD IS UP conceptual metaphor; which fits with Meier and Robinson's (2005) first prediction of consistency, that if emotional concepts are represented using conceptual metaphors, then a processing advantage should be observed for stimuli that have properties consistent (in this case in terms of their vertical position) with the conceptual metaphor.

In Brookshire et al. (2010) a single word was presented on each trial. That word was coloured purple or green, and the participants' task was to decide on the font colour. Participants pressed and held a centre key to start the trial. To identify the font colour they released the centre key and moved to the purple or green key, which were positioned above and below the centre key. Even though the emotional valence of the stimuli was irrelevant to the task, participants were faster to release the centre key and press the key in the upper-position when identifying the font colour of positive words than of negative words, and faster to release the centre key, and press the key in the lower-position when identifying the font colour of negative words than of positive words. Brookshire et al's results are consistent with Meier and Robinson's (2005) second prediction of metaphor congruent perceptual processing, that if emotion concepts are represented using conceptual metaphors, then activating target domain concepts should activate the perceptuomotor source domain in a metaphor congruent manner.

Pervasive metaphor.

The GOOD IS UP conceptual metaphor is pervasive. First, it is not limited to verticality in the visual domain. There is evidence that verticality effects extend to the auditory and bodily domains. Weger, Meier, Robinson, and Inhoff (2007) reported that evaluating positive words biased participants to identify tones as high pitched and evaluating negative words biased participants to identify tones as low pitched. This mapping is consistent with the conceptual metaphor GOOD IS UP, as high tone and low tones are also mapped to upper and lower space (see Bernstein & Edelstein, 1971; Chiou & Rich, 2011; Evans & Treisman, 2010 for evidence of the HIGH PITCH IS UP metaphor). Meier and Hauser (2008; as cited in Crawford, 2009) reported consistency between the valence of the word participants were evaluating and the part of the body with which they responded. Participants were faster to evaluate positive words with their finger (part of the upper body) than with their foot (part of the lower body), and were faster to evaluate negative words with their foot than with their finger.

Second, the GOOD IS UP conceptual metaphor is not only activated by evaluation of single word stimuli. General mood experience also shifts visual attention in a pattern consistent with the conceptual metaphor (Meier & Robinson, 2006). Degree of neuroticism was correlated with vertical attention bias. The higher participants were on neuroticism scores, the faster they were to respond to targets in the lower visual-field (regardless of the stimulus valence). A stronger correlation was found with depression. The higher participants scored on a measure of depression, the faster they were to respond to targets in the lower visual-field. As an aside, it is interesting to consider what role body specific effects may have played in these correlations in addition to valence specific effects. For example, depressed people generally have a more slumped posture compared to non-depressed controls (Michalak, Troje, Fischer, Vollmar, Heidenreich, & Shulte, 2009), and focusing on achieving a more erect posture is a part of some therapies for depression (Steckler & Young, 2009).

Third, there is also non-linguistic evidence for the GOOD IS UP conceptual metaphor, which reaffirms that conceptual metaphor mapping is a general cognitive process, and not a representation specific to language. Meier and Hauser (2008; as cited in Crawford, 2009) reported that participants' intuitions of valenced tattoo positions were biased in the direction of the GOOD IS UP conceptual metaphor. Participants preferred positive tattoos to be on the upper body, and negative tattoos to be on the lower body. Crawford, Margolies, Drake, and Murphy (2006) explored whether valence biased participants' memory for the position of pictorial stimuli. The vertical position in which participants remembered a positive picture being presented was higher than its original position, and the position in which participants remembered a negative picture was lower than its original presentation. This GOOD IS UP congruent memory bias was evident both with pictures drawn from the International Affective Picture System (IAPS) and with yearbook pictures paired with valenced descriptions; and was evident immediately and after a long delay between viewing the picture and position recall. The Crawford et al. study is additional evidence for the processing of valenced stimuli activating metaphor congruent perceptuomotor processing (Meier & Robinson's, 2005, second prediction). Viewing a valenced picture with the aim to remember its position activated GOOD IS UP consistent perceptuomotor processes and biased the remembered location.

Spoken Language

The studies given as evidence for the cognitive reality of the conceptual metaphor GOOD IS UP can be mostly divided into two types: those that used manipulations of mood, or measures of personality, to assess the presence of the verticality-emotion mapping; and those that used manipulation of linguistic stimuli. Those studies which used non-linguistic manipulations contribute to our understanding of the nature of conceptual metaphoric representation because they demonstrate that conceptual metaphoric processing is not specific to linguistic processing. Those which use linguistic stimuli are useful too, as exploration of the nature of conceptual metaphoric representation during linguistic processing is one way to assess the cognitive reality of conceptual metaphors. However, all of the experimental-linguistic studies use written language. This generalisation is also true of empirical studies exploring the validity of conceptual metaphors other than GOOD IS UP.

While the studies using written stimuli all report GOOD IS UP consistent responding, this does not mean that we should stop questioning the nature of conceptual metaphors. This thesis will assess the cognitive validity of the conceptual metaphor GOOD IS UP in spoken language processing. Assessing the cognitive validity of conceptual metaphors in spoken language will add to the theoretical understanding of Conceptual Metaphor Theory. If the same source (verticality) and target (positive/negative) mappings are observed with spoken linguistic stimuli as with written linguistic stimuli, this would strengthen arguments for conceptual metaphoric based representation. If no verticality-emotion mappings are observed when processing spoken linguistic stimuli, I would question how broad ranging conceptual metaphoric representation is. Investigation of the GOOD IS UP conceptual metaphor in *spoken* language is useful theoretically for several reasons.

Evolution.

First, Conceptual Metaphor Theory and other grounded cognition theories emphasise repeatedly that there is no separate representation system for concepts. The mind uses the *evolutionary older* perceptual and motor systems (Barsalou, 1999; Lakoff & Johnson, 1999). The earliest evidence of written language is approximately 5000 years old (Harley, 2001) therefore written language developed very recently in our cognitive history and presumably makes use of many processes beyond the perceptuomotor system. Furthermore, developmentally, people learn to speak before they learn to write, and a cognitively normal adult may not be able to read but have normal speech (Wurm, Vakoch, Strasser, Calin-Jageman, & Ross, 2001). A more stringent test of Conceptual Metaphor Theory, and grounded-cognition theories in general, is to examine whether conceptual metaphor congruent processing is present when assessed with spoken linguistic stimuli. This theme is emerging in other avenues of research. Wurm et al. (2001), and Wurm, Vakoch and Seaman (2004) have argued that as spoken language is evolutionarily older than written language; if emotional and linguistic processing interact, evidence is more likely to be seen in studies of spoken, than written, language. Cook (2002) states that, in our evolutionary history, pitch in animal calls conveyed information regarding dominance, danger, and mate selection. As an emotion system in the brain developed, pitch, as a component of prosody (see the complexity section below), also came to be used to convey emotional information. Evidence for this claim is the similarity of animal vocalisations and pitch in the human voice (Cook, 2002). If this evolution argument is valid, and there are stronger links between the grounded representation system and spoken language than with written language, then the role of conceptual metaphors in emotional language processing may be more pervasive than is indicated by studies using written language.

Complexity.

Second, spoken language is more complex than written language. The use of speech allows the controlled manipulation of two channels: semantics (what we say), and prosody (how we say it). Prosody is one of the ways that emotion is expressed in language and is a feature of spoken language that expresses information at a level above segmental features like phonemes. Prosody changes the quality of the segments in terms of their pitch, intensity, and length, but not their phonemic nature (Ladd, 1996). Happy speech has high mean pitch and sad speech has low mean pitch (Banse & Scherer, 1996; Scherer, 2003).

Studies examining the conceptual metaphor GOOD IS UP with written words are purely semantic in focus. Although prosody is an extralinguistic property of language, there is no reason to think that the conceptual metaphor GOOD IS UP is not recruited during processing of emotional prosody. In other areas of research interactions between linguistic and prosodic processing have been demonstrated. For example, emotional prosody seems to play a role in lexical access. Using a homophone spelling task, in which participants listened to a homophone spoken in happy, neutral, or sad prosody, then transcribed it, Nygaard and Lunders (2002) demonstrated that participants transcribed the emotional spelling of a homophone more often when the homophone was spoken in emotional prosody than in neutral prosody. Nygaard and Queen (2008) extended the observation of prosodic modulation of linguistic processing to non-ambiguous words. Participants were faster to name words when the semantics and prosody of the word were congruent. That is, they were faster to repeat a spoken semantically-positive word when it was spoken in happy prosody (than in sad or neutral prosody) and were faster to repeat a spoken semantically-negative word when it was spoken in sad prosody (than in happy or neutral prosody).

Such studies remind researchers that language is not uni-dimensional. Any theory of representation of emotion-related concepts, including Conceptual Metaphor Theory, must be able to account for effects of emotion across the range of linguistic complexity; in both written and spoken language, and in both the semantic and prosodic channels of spoken language. The research conducted up to the current date has only evaluated the cognitive validity of the GOOD IS UP conceptual metaphor with written words. An investigation of the relevance of the GOOD IS UP conceptual metaphor to representations accessed during spoken word processing is long overdue.

The Current Studies

No studies have yet been conducted that were specifically designed to assess the cognitive reality of conceptual metaphors in spoken emotional language. For evolutionarily and complexity reasons, a better test of the cognitive reality of the GOOD IS UP conceptual metaphor is to use spoken words rather than written words. The aim of studies in this thesis was to examine the role of conceptual metaphors in spoken language processing. This thesis explores whether shifts in attention congruent with the GOOD IS UP conceptual metaphor are induced by emotional semantics and emotional prosody separately. That is, this thesis tests Meier and Robinson's (2005) second prediction of metaphor congruent perceptual processing for the GOOD IS UP conceptual metaphor, that activation of emotion-related concepts activates GOOD IS UP congruent shifts in attention.

Using a spatial attention paradigm, analogous to that used by Meier and Robinson (2004) with visual words, four studies were conducted. Study 1 was conducted to ensure that the visual attention paradigm was sensitive to attentional manipulation. Study 2 resulted in the creation of well balanced sets of words for use in Studies 3 and 4. Study 3 was conducted to determine if spoken words that were semantically emotional resulted in GOOD IS UP congruent shifts in attention. Study 4 was conducted to determine if spoken words that were prosodically emotional resulted in GOOD IS UP congruent shifts in attention.

Studies 1, 3, and 4 were similar in procedure. All used the same dual-task procedure involving evaluation of an auditory stimulus followed by a visual attention task. The only major difference between the studies was the auditory stimuli used. The paradigm was dual-task. In the auditory task component, participants identified the auditory cue on a categorical dimension. In Study 1, the auditory cue was a high or low pitched tone and participants' task was to decide if it was Tone X or Tone Y.

In Study 3, the auditory cues were semantically negative, neutral, or positive words spoken in neutral prosody and participants evaluated the words semantically. In Study 4, the auditory cues were semantically-neutral words spoken in sad, neutral, and happy prosodies, and participants evaluated the words prosodically.

In the visual attention task component, participants made a speeded target detection and identification response to a visual target. In Studies 1, 3, and 4 the visual targets were black shapes; a square and a circle. On each experimental trial, the auditory cue was presented first, then, after a short or long SOA, the visual target could appear. As soon as a shape appeared participants indicated with a key release that they had detected the shape and then with a key press identified the shape as a square or circle. As the visual attention task was go-no-go there were catch trials on which no shape was presented. After responding (or not, on catch trials) to the visual target, participants identified (in Study 1) or evaluated (in Studies 3 and 4) the auditory cue in terms of its pitch (Study 1), semantic emotion (Study 3), or prosodic emotion (Study 4).

The visual attention task used was inspired by that in Meier and Robinson's (2004) Study 2. In their design participants were presented with a positive or negative visual word cue, which they evaluated with a spoken response as positive or negative, and subsequently saw a *p* or a *q*. The letter target could appear in the upper or lower visual-field; however the position of the letter was irrelevant to the task. Participants were required to identify the letter by pressing the p key on the keyboard with their right index finger or the q key with their left index finger. This paradigm induced GOOD IS UP congruent shifts in attention; responses were faster to targets in the upper visual-field after presentation of a positive word and faster to targets in the lower visual-field after presentation of a negative word. However, in addition to using spoken words, the present study included several major methodological modifications to the paradigm used by Meier and Robinson (2004). These changes were made in order to conduct a more stringent test of the predicted metaphor congruent perceptual processing.

First, the visual targets used in Study 2 of Meier and Robinson (2004) were letters, which are linguistic stimuli. A more powerful test of the induction of perceptual processing consistent with emotion-verticality mappings is to use non-linguistic targets. A black square and a black circle were used. The participants' task required a multiple step response. To start each trial participants pressed and held the

5 key on the number pad. When a shape was detected participants were instructed to release the key as quickly as possible and then to press the key to the left or the right of the 5 key to indicate if the shape was a square or circle.

Second, in everyday life, as well as in experimental settings, there are many potential spatial mappings to be considered. In addition to the mappings of interest there are also stimulus-response compatibility (SRC) mappings. Participants in Meier and Robinson (2004) viewed stimuli that appeared in the upper or lower visual-field and responded on keys that were positioned to the left (*q*) and right (*p*). People are generally faster to respond to lower visual-field targets with a left key and to upper visual-field targets with a right key (Weeks & Proctor, 1990). Furthermore, right handed participants generally map positive to the right position and negative to the left position (Casasanto, 2009). Such SRC and handedness mappings could confound any shifts in attention due to the GOOD IS UP conceptual metaphor and were not considered by Meier and Robinson (2004). The verticality paradigm used in this thesis was designed to minimise the contribution of these potential mappings to response time. First, as in Brookshire et al. (2010), three reaction times were recorded. Release time, the time to release a key on detection of a target, should not be affected by left-down/ up-right SRC mappings. Press time, the time to press a key to the left or the right of a central key, and movement times, the time to move after releasing the central key to the left or right key, could be affected by left-lower/ right-upper SRC mappings, and so the assignment of shape to key was counterbalanced across participants. Finally, to minimise any effect of valence-handedness mappings, all participants were right handed.

Third, in order to be able to make a more powerful claim regarding the automaticity of any verticality mappings, the order of the task components was changed. In Meier and Robinson (2004), the evaluation of the emotional words occurred before the presentation of the visual target. That is, participants saw a word, evaluated it, and then saw a visual target to which they responded. A powerful way of elucidating the time course of processing is to manipulate Stimulus Onset Asynchrony (SOA), the time between the onset of stimulus one, the word, and the onset of stimulus two, the shape. With the task component order used by Meier and Robinson (2004) their 2005 prediction of automaticity (that metaphor congruent perceptual processing, including shifts in attention, will be observed at automatic processing stages) is hard to assess. However, by reversing the order of the visual stimulus and the evaluation in conjunction with the use of two SOAs the automaticity prediction can be tested. Two SOAs between the spoken word and the visual target are used; a short SOA, at which attentional orienting is thought to be automatic, and a long SOA, at which attentional orienting is thought to be controlled (Posner, 1980; Posner & Snyder, 2004). The auditory cue was presented first, then the visual target to which participants made a speeded response, and finally participants made their evaluation response to the word. The SOA manipulation also adds unpredictability to the timing of the onset of the shape target. With a randomly varying SOA, participants cannot get into a regular rhythm of responding.

Even though the evaluation response does not occur until the end of the trial, after the presentation of and response to the visual target, it is still possible to be fairly sure that participants were evaluating the valence of the word by comparing response times at the short and long SOA. The psychological refractory period (PRP) effect describes the phenomenon in dual-task situations where, as the SOA between two stimuli decreases, the time to respond to the second stimulus increases (Pashler, 1992; 1993). Pashler reports that this interference is not due to a delay at stimulus perception or response production, but rather to a cognitive-bottleneck at response-selection. Participants cannot begin the response-selection process for the second stimulus (in this case the shape) until a response has been selected, but not necessarily produced for the first stimulus (in this case the auditory cue). Thus if in the current paradigm participants are selecting their evaluation response before selecting their shape response, response times will be faster at the long SOA than at the short SOA.

Fourth, the modality of the evaluation response was changed. Participants were required to click on a box labelled with tone types, semantic valences (positive, neutral, negative) or prosodic valences (happy, neutral, sad). In Meier and Robinson's (2004) paradigm participants spoke the words "positive" or "negative" to evaluate the words. Mouse clicks were thought to be less likely, compared to explicit spoken production of valenced labels, to result in conceptual metaphor activation due directly to the labels used.

Fifth, neutral valenced words and prosody were included. In everyday language there is not a clear contrast between positive and negative themes. They are intermixed with neutral words and voices. The inclusion of neutral semantics and prosody allows the examination of the contribution of grounded representation in a more ecologically valid setting. Furthermore, in order to look at the independent contribution of emotional semantics and prosody separately one channel must be neutral.

These five changes were not expected to reduce the contribution of the GOOD IS UP metaphor in conceptual processing. Rather, these changes allowed a more stringent test of the cognitive reality of the GOOD IS UP metaphor to be conducted. As many confounds as possible have been removed or controlled for and the paradigm has been adapted to be more suitable for assessing Meier and Robinson's (2004) predictions of congruency and automaticity. If the GOOD IS UP metaphor underlies representation for emotional words then metaphor congruent shifts in attention should be observed. After evaluating words that are positive in terms of their semantics or prosody, participants should be faster to respond to visual targets in the upper visual-field than in the lower-visual field. After evaluating words that are negative in terms of their semantics or prosody, participants should be faster to respond to visual targets in the lower-visual field than in the upper visual-field.

STUDY 1

Study 1 was conducted to ensure that the revised paradigm was sensitive to verticality mappings. The conceptual metaphor HIGH PITCH IS UP was chosen to be the test of whether metaphor congruent shifts in attention can be observed with this paradigm. The conceptual metaphor HIGH PITCH IS UP describes the mapping between the perceptuomotor source domain, verticality, and the conceptual target domain, pitch. The HIGH PITCH IS UP conceptual metaphor is especially relevant to this thesis where prosody is considered, as pitch is a key component of prosody. In experiments investigating pitch-verticality mappings participants are generally presented with an auditory and a visual stimulus. The auditory stimulus can be high or low in pitch. The visual stimulus can be presented in the upper or lower visual-field. Facilitation is observed for high pitch upper visual-field and low pitch lower visual-field pairings, compared to the opposite pairings. It is thought that the HIGH PITCH IS UP metaphor originates from repeated experience of the spatial position in which high and low pitches resonate in the body. When a speaker produces low pitched sounds the vocalisation resonates in the speaker's chest, whereas when a speaker produces high pitch sounds the vocalisation resonates higher than the chest and feels like it is resonating in the head area (Zbikowski, 1998). As would be expected from such a frequently occurring collocation between pitch and verticality, the pitch-verticality mapping is very robust (Ben-Artzi & Marks, 1999; Chiou & Rich, 2011;

Bernstein & Edelstein, 1971; Evans & Treisman, 2010; Maeda, Kanai & Shimojo, 2004; Melara & O'Brien, 1987; Occelli, Spence & Zampini, 2009; Sadaghiani, Maier & Noppeney, 2009; Shintel, Nusbaum & Okrent, 2006). Infants as young as nine months old map ascending tones to upward pointing arrows and descending tones to downward pointing arrows (Wagner, Winner, Cicchetti, & Gardner, 1981). In Study 1, a high- (2000Hz) and a low- (500Hz) tone were used as the auditory stimuli. The purpose of Study 1 was primarily to determine if the modified paradigm is suitable for investigating auditory-visual verticality mappings. Participants should be faster to respond to visual targets in the upper visual-field after listening to the high-tone, and faster to respond to visual targets in the lower visual-field after listening to the low-tone.

Method

Participants

Participants were 22 (15 female, 7 male; mean age 18.23 years) undergraduate students. All had normal or corrected-to-normal vision, had no hearing deficits, were right handed (as assessed by the Waterloo Handedness Questionnaire–Revised; Elias, Bryden, & Bulman-Fleming, 1998), and were in the sub-clinical range (not greater than 57 out of 80) on anxiety or depression (as assessed by the Zung Anxiety, 1965, and Depression Questionnaires, 1971).

Stimuli and Apparatus

Both the auditory and visual stimuli were presented using a Dell PC running Psychology Software Tools' E-Prime Suite version 1.1 (Schneider, Eschman, & Zuccolotto, 2002). Visual stimuli appeared on a 31cm x 23cm Dell CRT monitor with a vertical refresh rate of 60 Hz. Auditory stimuli were presented with Manhattan noise-cancelling stereo headphones with circumaural cushions. The output was verified using a Phillips sound meter as being approximately 75dB.

The auditory cues were low- (500Hz) and high- (2000Hz) pure tones of 500ms duration. The tones were created in Audacity (version 1.2.6) using the Sine Tone Generator. The visual targets were a black square and circle 47 x 47 pixels in size (1.4 cm x 1.4 cm). The computer monitor was positioned approximately 50cm from the participant, on a stand so that centre fixation was 37cm from the table top, and approximately at eye level for the participant.

Procedure

This experiment used a dual-task paradigm. Participants performed a visual-attention task and a tone-identification task. Specifically, on experimental trials participants heard a tone, and then saw a shape. Participants were required to make a two-step speeded response to detect and then identify the shape, and subsequently to make a non-speeded response to identify the tone.

The tone-identification task required participants to identify the tone they heard as Tone X or Tone Y. This letter labelling allowed the experimenter to avoid use of the terms “high” and “low” when describing the tones. Such explicit labelling using location terms may activate HIGH PITCH IS UP congruent shifts in attention alone (Evans & Treisman, 2010). The visual-attention task was a go no-go target discrimination task. Participants were required to make a speeded two-step response, detection and shape identification, to a present shape. To add an element of unpredictability to the target discrimination component of the task, the target could appear in one of four of locations; 2cm from the top of the screen (high upper visual-field; 9.5cm from the centre), 2 cm from the bottom of the screen (low lower visual-field; 9.5cm from the centre), 3 cm from the top of the screen (medium upper visual-field; 8.5cm from the centre), or 3 cm from the bottom of the screen (medium lower visual-field; 8.5cm from the centre). At an approximate viewing distance of 50cm, the high upper visual-field and low lower visual-field positions corresponded to approximately 11 degrees to the centre of the shape, and the medium visual-field positions corresponded to approximately 10 degrees to the centre of the shape. The target appeared 24 times in each of these locations.

The experiment consisted of 120 trials; 96 target-present trials and 24 catch trials in which a tone was heard, but no shape was presented. For the target-present trials, 48 were preceded by the low-tone and 48 by the high-tone. For each of the 48 low and high trials, 24 had a short SOA (500ms) and 24 a long SOA (1200ms). For each of the SOA durations, on 12 trials a square shape followed the beep, and on 12 a circle; six were presented in the upper visual-field and six in the lower visual-field. Of these, three were presented in the high/low region of the visual-field and three in the medium region of the visual-field. In the data analysis stage, the trials were collapsed across the high/low and medium positions, and across the shape-type, to give 24 upper and lower visual-field targets following a low-tone, and 24 upper and lower visual-field targets following a high-tone for each SOA.

See Figure 1 for the target-present trial procedure. A trial started when participants pressed and held the 5 key on the number pad with their right index finger. Then a fixation cross appeared for a random duration between 1000ms and 1500ms; participants were instructed to stare at the cross. During the last 500ms of this interval the tone was played. After a 0ms or 700ms ISI (to make a 500ms or 1200mm SOA) a shape appeared on target-present trials. The fixation was displayed during the 700ms ISI. However, participants could not use the fixation offset as a cue to the onset of a shape target because the catch trials also had a fixation offset that corresponded half of the time to the timing used on the target-present trials with a short (500ms) SOA and half of the time to the timing used on the target-present trials with a long (1200ms) SOA. The shape was displayed for up to 4000ms but terminated after a key release and press. On target present trials, participants were instructed to release the 5 key as quickly as possible once they had detected the target. Once they had detected a target and released the 5 key, participants were instructed to press 4 or 6 (left or right movement) to indicate if they saw a square or circle. On the catch trials participants were instructed to keep pressing the 5 key. 500ms after a response (or the full 4000ms in the event of no response as in catch trials), a question screen was displayed. Participants used the mouse with their left hand to click on the box (labelled Tone X and Y) that corresponded to the tone that they had heard at the beginning of a trial. The tone-identification component was included to ensure that participants would evaluate the tones as they processed them. After a 1000ms ISI the next trial began.

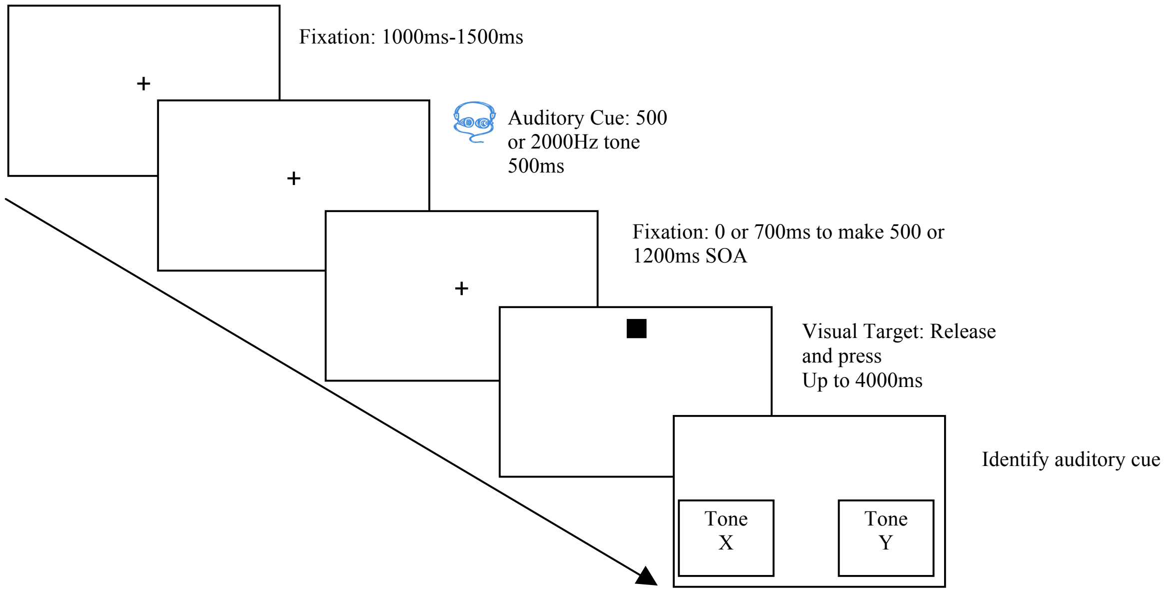

Figure 1. Target-present trial procedure in Study 1.

Reaction times for releases and presses from the onset of the shape were recorded using E-Prime and the computer's internal timer. In order to control for possible response mapping influences, the assignment of tone to label (500Hz, 2000Hz; Tone X, Tone Y) and of shape to key (square, circle; 4, 6) were counterbalanced across participants.

Before the 120 experimental trials, participants were given practice trials for each task component. As each trial is made up of several components, it was not desirable for task complexity to mask any underlying effects. As I did not control for typing or game-playing experience, it was especially important that participants be trained to use the same finger to release and press keys in order to control for different comfort levels with pressing multiple keys. Thus participants completed four sets of practice trials. First, they completed ten tone-practice trials on which they only heard a tone (five low- and five high-tones) and identified the tone they heard. They were given feedback on their tone identification. Second, they completed six release-practice trials on which they might see a shape (two square, two circle, two catch) and released the 5 key upon seeing it. Third, they completed ten press-practice trials on which they might see a shape (four square, four circle, and two catch trials) and released the 5 key upon seeing it and then identified it by pressing the 4 or 6 keys. For the second and third set of practice trials it was made clear to the participants that they should only release and press keys with their right index finger (i.e. not the middle or ring fingers). Finally, they completed ten practice trials with all components.

Results and Discussion

Two participants were removed from the analysis (see below), resulting in a sample of 20 participants (14 female, 6 male; mean age 18.15 years). There were three response time variables; the time from the onset of the shape to release the 5 key on detection of a target (release time), the time from the onset of the shape to press the 4 or 6 key to identify the target (press time), and the time between the release and the press (movement time). As the key release component of the shape task was a go-no-go target detection task, the number of catch trials on which participants responded was inspected. One participant responded on more than two (out of 24) catch trials and was removed from the analysis. Release times below 200ms were deemed anticipatory and times above 1500ms as prolonged detection. Therefore, the release times for the remaining participants were filtered so that only trials on which the release time was greater than 200ms and less than 1500ms were used to calculate a median release time for each SOA, tone, and visual-field combination. All participants had at least 93 trials (out of 96; maximum 3% data excluded) with which to calculate a median release time. Thus, almost no data points were excluded from the analysis of release times.

The key press component of the shape task was a target discrimination task; participants were required to report whether they saw a square or a circle. The press times excluded trials on which participants identified the shape incorrectly and the reaction times were filtered so that only trials on which the press time was greater than 200ms were used to calculate a median press time for each SOA, tone, and visual-field combination. All participants had at least 90 trials (out of 96; maximum 6% data excluded) with which to calculate the median press times. One participant was removed from the analysis for not meeting this criterion.

The difference between the auditory stimuli was non-subjective. Tone X and Tone Y always differed in fundamental frequency by 1500Hz. All participants scored at least 82 percent accuracy on the tone identification task (M = 91%, SD = 8%). Nevertheless the analyses below were conducted twice; once with all trials regardless of tone-identification accuracy and repeated with only trials on which participants had identified the tone correctly. The same main effects and interactions were found. The analyses reported below for release, press, and movement times are therefore based on all trials regardless of tone accuracy, contingent on the data filters described above. See Table 1 for a summary of the release, press, and movement times.

Table 1.

Mean (SD) release, press, and movement times (ms) for Study 1 by SOA, tone, and visual-field.

	Release Times			
	Short SOA		Long SOA	
Tone	Lower VF *M (SD)*	Upper VF *M (SD)*	Lower VF *M (SD)*	Upper VF *M (SD)*
High-Tone	475	463	463	439
	(111)	(108)	(89)	(87)
Low-Tone	472	474	443	446
	(100)	(101)	(79)	(91)
	Press Times			
	Lower VF	Upper VF	Lower VF	Upper VF
High-Tone	856	870	815	775
	(208)	(284)	(263)	(205)
Low-Tone	916	846	782	776
	(325)	(248)	(224)	(189)
	Movement Times			
	Lower VF	Upper VF	Lower VF	Upper VF
High-Tone	359	390	348	311
	(199)	(241)	(219)	(174)
Low-Tone	400	369	328	305
	(301)	(230)	(215)	(172)

Release Times

The median release times were analysed in a 2 (SOA: 500ms, 1200ms) x 2 (Tone: low, high) x 2 (Visual-field: upper, lower) repeated-measures ANOVA. Most importantly, there was a significant tone x visual-field interaction, $F(1, 19) = 7.387$, $MSE = 598$, $p = .014$, $\eta_p^2 =.280$; that did not interact with SOA, $F(1, 19) = .387$, $MSE = 899$, $p = .541$, $\eta_p^2 = .020$. Follow up paired-samples t-tests indicated that across

SOAs for the high-tone trials participants were significantly faster to release the key on detecting an upper visual-field target (*M* = 453ms, *SD* = 98ms) than a lower visual-field target (*M* = 472ms, *SD* = 99ms), $t(19) = -2.777$, $p = .012$. However, although in the predicted direction, for the low-tone trials participants were non-significantly faster to release the key following a lower visual-field target (*M* = 457ms, *SD* = 88ms) than an upper visual-field target (*M* = 461ms, *SD* = 95ms), $t(19) = .662$, $p = .516$. See Figure 2. This interaction demonstrates that the paradigm is sensitive to a metaphoric shift in attention. Processing of high pitch seemed to shift attention to the upper visual-field and the processing of low pitch to the lower visual-field. This pattern of mapping is congruent with the HIGH PITCH IS UP conceptual metaphor. As the short SOA was 500ms, which is relatively long in terms of processing time, no claims can be made about the automaticity of the HIGH PITCH IS UP conceptual metaphor. However, in Studies 3 and 4 the short SOA will be 400ms, allowing investigation of automaticity.

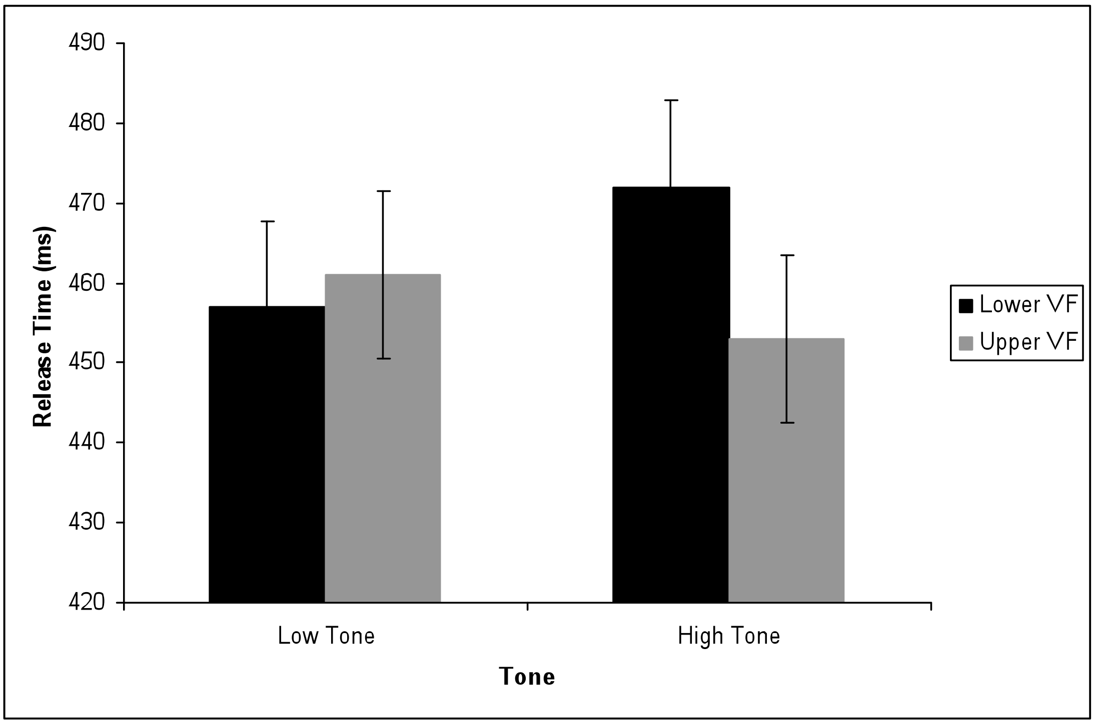

Figure 2. Release times for high- and low-tone trials for upper and lower visual-field targets in Study 1.

There was also a main effect of SOA such that participants were faster to release on detection of a target in the 1200ms SOA trials (*M* = 448ms *SD* = 83ms) than in the 500ms SOA trials (*M* = 471ms, *SD* = 101ms), $F(1, 19) = 11.380$, *MSE* = 1921, $p = .003$, $\eta_p^2 = .375$.

Press Times

The median press times were analysed in a 2 (SOA: 500ms, 1200ms) x 2 (Tone: low, high) x 2 (Visual-field: upper, lower) repeated-measures ANOVA. In contrast to the release times there was no significant tone x visual-field interaction, $F(1, 19) = 1.009$, *MSE* = 6288, $p = .328$, $\eta_p^2 = .050$. The only significant effect was a main effect of SOA, such that participants were faster to identify a target on the 1200ms SOA trials (*M* = 787ms, *SD* = 211ms) than on the 500ms SOA trials (*M* = 872ms, *SD* = 255ms), $F(1, 19) = 25.482$, *MSE* = 11346, $p < .001$, $\eta_p^2 = .573$.

Movement Times

The predicted tone x visual-field interaction was only observed for release responses. Participants were instructed to release the 5 key as soon as they detected a shape even if they had not identified it yet. The release and press times were recorded from the onset of the shape. To rule out any opposing effects present in the release and press times, movement time was calculated for the lag between the release response and the press response and the SOA x tone x visual-field ANOVA was repeated for the movement times. As in the press times there were no interactions and the only main effect was of SOA, $F(1, 19) = 12.933$, *MSE* = 9795, $p = .002$, $\eta_p^2 = .405$; with faster movements at the long SOA (*M* = 323, *SD* = 186) than at the short SOA (*M* = 379, *SD* = 223).

That the same main effect and no interactions were obtained in both the press and movement times indicates that there were not any effects in the press times that were cancelled out by the release times. The movement time analysis also demonstrated that the lag between releasing and pressing keys was short, often less than 100ms. As mentioned in the introduction, SRC mappings could mask conceptual metaphoric consistent shifts in attention in the press and movement times. Thus, it is not surprising that, if conceptual metaphoric shifts in attention are observed in only one reaction time variable, they are observed in the release times. The release times should be free of any confounding SRC mapping influence. Furthermore, Brookshire et al. (2010) state that movement times tap action execution rather than action planning. Thus it is primarily the release times, and secondly the press times, that are of theoretical interest and in Studies 3 and 4 the movement times will not be analysed.

Participants were faster to make both detection (indexed by the release time) and identification (indexed by the press and movement times) responses to shapes for trials with a long SOA than a short SOA. The psychological refractory period (PRP) effect could account for this long SOA advantage (Pashler, 1992; 1993). At the short SOA participants may have been delayed at selecting their response to the shape until they had selected their response to the word. The long SOA advantage observed in Study 1 thus suggests that even though participants were not required to respond to the tones until after the shape task, they were most probably evaluating and selecting their tone-identification response before they responded to the shapes. A strength of the current paradigm is that the long SOA advantage can be used as a marker of the evaluation process.

The presence of a significant tone x visual-field interaction in the predicted direction in the release times combined with evidence of a PRP effect suggests that the paradigm is sensitive to attentional manipulation and is suitable for investigating auditorally induced verticality mappings.

STUDY 2

The purpose of Study 2 was to investigate and control the psycholinguistic properties of the stimuli used in Studies 3 and 4. Studies 3 and 4 are similar in structure to Study 1. The experiments differ in the nature of the auditory stimuli. In Study 3, the auditory cues were semantically positive, negative, and neutral words, spoken in neutral prosody; which allows investigation of attentional shifts when processing emotional semantics. In Study 4, the auditory cues were semantically neutral words spoken in happy, sad, and neutral prosody; which allows investigation of attentional shifts when processing emotional prosody.

Studies 3 and 4 combined required 160 words for the target present trials: 96 semantically-neutral words, 32 positive words, and 32 negative words. Thirty-two of the neutral words needed to be spoken in neutral prosody, 32 in happy prosody and 32 in sad prosody. The positive and negative words needed to be spoken in only neutral prosody. Ratings for semantic emotion were taken from the Affective Norms of English words (ANEW; Bradley & Lang, 1999); however, it was necessary to conduct a ratings experiment to assess evaluation of prosodic emotion. An initial pool of 250 words (50 positive words spoken in neutral prosody, 50 negative words spoken in neutral prosody, 50 neutral words spoken in neutral prosody, 50 neutral words spoken in happy prosody, and 50 neutral words spoken in sad prosody[3]) were included in a prosodic ratings experiment to select the final 160 words.

Method

Participants

Twenty participants (4 male, 16 female; mean age 27.00 years), who did not participate in other studies, completed the prosodic identification component of this study.

Stimuli and Apparatus

The words were drawn from the ANEW (Bradley & Lang, 1999) and were spoken by an adult female voice actress. The prosodies were recorded in blocks to enhance consistency. The digital stimuli were recorded in one channel (mono) with a Neumann U87 microphone at 24bits and 44100Hz using the software Protools version 7, controlled by a Macintosh G5 computer. The editing software Audacity version 1.2.6 was used to duplicate the mono recordings to make stereo tokens, convert the files to 16bits, add 40ms of silence to the beginning of each word, and to equate the tokens for peak amplitude.

Procedure

Each participant identified the prosody of 150 tokens: 25 positive words spoken in neutral prosody, 25 negative words spoken in neutral prosody, 25 neutral words spoken in happy prosody, 25 neutral words spoken in sad prosody, 25 neutral words spoken in fearful prosody[4], and 25 neutral words spoken in neutral prosody. Each of the tokens was a unique word for each participant, that is, they heard each individual word only once, in one of the four prosodies. As there were 20 participants, overall each token was identified by ten individuals. Participants were instructed to listen to the word, ignore the meaning, and click on the box that best described the tone of voice. The emotions participants could select from were: angry, fearful, sad, happy, neutral, and other. In order to exclude poor tokens, participants were also instructed to click a box labelled "could not hear" if they could not understand the word. For each word an identification score was calculated: the percentage of times participants identified the prosody as the valence intended by the voice actress.

---

[3] The neutral words spoken in sad prosody were also spoken in fearful prosody for use in another study.
[4]For use in another study, not reported here.

Results and Discussion

Prosodic Properties

The 250 words identified in the prosodic identification experiment were narrowed down to 160 words and assigned to lists. The final 160 words were identified as their intended prosody by at least 70 percent of participants. Seventy percent is in the high range for identification of emotional prosody. Accuracy of prosodic identification is roughly 50 percent in most studies, which is well above chance, given the number of prosodies from which to choose (see Banse & Scherer, 1996).

See Table 2 for the prosodic identification scores for each list type. The properties of the words reported here are for the lists used for the target-present trials in Studies 3 and 4. There were two lists from which words would be drawn from for Study 3: the positive-semantics list (32 semantically positive words spoken in neutral prosody) and the negative-semantics list (32 semantically negative words spoken in neutral prosody). There were two lists from which words would be drawn from for Study 4: the happy-prosody list (32 semantically neutral words spoken in happy prosody) and the sad-prosody list (32 semantically neutral words spoken in sad prosody). Finally one list was for use in Studies 3 and 4: the neutral list (32 semantically neutral words, spoken in neutral prosody). None of the final 160 words had any reports of not being able to hear what the word said. See Appendix A for the word lists.

Acoustic Properties

Acoustic analysis of the stimuli was also conducted. The characteristics of duration and pitch were selected. Duration was measured in ms (excluding the 40ms of silence at the beginning of each of the files). The Fundamental Frequency (F0) parameters mean, median, and standard deviation were extracted using PRAAT (Boersma & Weenink, 2007). Values between 75 and 600Hz were submitted to the auto-correlation method and used for the extraction of the F0. See Table 2 for a summary of the acoustic parameters for each of the five stimulus sets.

Multivariate ANOVAs for the three prosodically-neutral lists to be used in Study 3 showed that the lists only differed on F0 standard deviation $F(2, 93) = 8.059$, $MSE = 150$, $p = .001$, $\eta_p^2 = .148$, and duration $F(2, 93) = 18.660$, $MSE = 12613$, $p < .001$, $\eta_p^2 = .286$. The lists did not differ on F0 mean or median. Post-hoc Tukey Tests showed that for F0 standard deviation, semantically-negative words had less within word variation than positive (*p* = .006) and neutral words (*p* = .001), which did not differ from each other (*p* = .831). For duration, the post-hoc Tukey tests showed that semantically-positive words were longer than negative (*p* <.001) and neutral words (*p* = .001), which did not differ from each other (*p* = .082). Semantically-negative words were thus shorter in duration and were spoken in a neutral prosody that had less variance than semantically-positive and neutral words, but did not differ from semantically-positive or neutral words in mean or median pitch.

Multivariate ANOVAs for the three semantically-neutral lists to be used in Study 4 showed that the lists differed on F0 mean, $F(2, 93) = 50.948$, $MSE = 1729$, $p < .001$, $\eta_p^2 = .523$, F0 median, $F(2, 93) = 33.081$, $MSE = 2456$, $p < .001$, $\eta_p^2 = .416$, F0 standard deviation $F(2, 93) = 38.167$, $MSE = 752$, $p < .001$, $\eta_p^2 = .451$, and duration $F(2, 93) = 14.241$, $MSE = 12285$, $p < .001$, $\eta_p^2 = .234$. Post-hoc Tukey Tests showed that for F0 mean and median, happy prosody was the highest in pitch and neutral prosody was the lowest in pitch with sad at an intermediary level (mean F0: happy-neutral *p* < .001, happy-sad *p* < .001, neutral-sad *p* = 005; median F0: happy-neutral *p* < .001, happy-sad *p* < .001, neutral-sad *p* = .037) Post-hoc Tukey tests showed that for F0 standard deviation, happy prosody had the greatest within-word variation (happy-neutral *p* < .001, happy-sad *p* < .001), and that neutral and sad did not differ significantly from each other (*p* = .551). Post-hoc Tukey tests showed that for duration (ms), sad prosodic words were the slowest, and happy prosodic words the fastest, with neutral at an intermediary level (happy-neutral *p* = .012, happy-sad *p* < .001, neutral-sad *p* = .049).

Table 2.

Prosodic and acoustic properties of the words used in Studies 3 and 4.

	Semantically Emotional			Prosodically Emotional	
Property	Positive-semantics List *M (SD)*	Negative-semantics List *M (SD)*	Neutral List *M (SD)*	Happy-prosody List *M (SD)*	Sad-prosody List *M (SD)*
			Prosodic Properties		
Prosodic identification percentage	94.37 (7.59)	94.69 (9.15)	95.31 (7.61)	97.19 (6.83)	93.44 (10.04)
			Acoustic Properties		
F0 Mean (Hz)	180 (10)	197 (82)	185^{eg} (17)	288^{ef} (58)	219^{fg} (39)
F0 Median (Hz)	178 (10)	182 (12)	184^{hj} (24)	283^{hi} (74)	215^{ij} (37)
F0 Standard Deviation (Hz)	25^{a} (17)	15^{ab} (8)	27^{bj} (10)	82^{jk} (27)	34^{k} (38)
Duration (ms)	762^{cd} (129)	593^{c} (114)	654^{dln} (91)	573^{lm} (125)	720^{mn} (113)

Note. a-n indicate statistically significant differences within the property.

Semantic and Lexical Properties

In contrast to previous conceptual metaphor experiments (e.g. Brookshire et al., 2010; Meier & Robinson, 2004), the words in this experiment were highly controlled for lexical variables. The final word lists for the target-present trials were balanced for the semantic properties valence and arousal (both retrieved from the ANEW; Bradley & Lang, 1999); the lexical properties of length, frequency (Kucera-Francis, 1967), imageability, familiarity, and concreteness (all three drawn from the MRC Psycholinguistics Database; Coltheart, 1981; Wilson, 1988), and orthographic neighbourhood size, phonological neighbourhood size, and bigram frequency (Ortho_N, OG_N, BG_Mean; drawn from measures obtained from the English Lexicon project; Balota et al., 2007). It is important to control for such lexical variables because if GOOD IS UP congruent shifts in attention are observed with a non-controlled word set it would be impossible to know if certain items were driving the shifts in attention due to their lexical properties. Lexical properties have been demonstrated to affect lexical processing. For example concreteness (Levy-Drori & Henik, 2006), frequency (Navarrete, Basagni, Alario, & Costa, 2006; Whaley, 1978), and orthographic neighbourhood size (Samson & Pillon, 2004). See Table 3 for a summary of the semantic properties for each of the five stimulus sets, and see Table 4 for a summary of the lexical properties for each of the five stimulus sets.

Table 3.

Emotional-semantic properties of the words used in Studies 3 and 4.

Property	Semantically Emotional			Prosodically Emotional	
	Positive-semantics List *M (SD)*	Negative-semantics List *M (SD)*	Neutral List *M (SD)*	Happy-prosody List *M (SD)*	Sad-prosody List *M (SD)*
Valence	7.71[a] (.44) Range 7.05-8.72	1.99[c] (.35) Range 1.25-2.74	5.50[ac] (.48) Range 4.51-6.45	5.40 (.30) Range 5.05-6.02	5.54 (.51) Range 4.02-6.68
Arousal	5.49[b] (1.41)	5.77[d] (.90)	4.06[bd] (.64)	3.89 (.55)	4.06 (.60)

Note 1. a- d indicate statistically significant differences within the property, $p < .005$.

Note 2. The valence ratings on the ANEW range from 1 (negative) to 9 (positive) and the arousal ratings on the ANEW range from 1 (low arousal) to 9 (high arousal).

Table 4.

Lexical properties of the words used in Studies 3 and 4.

Property	Semantically Emotional			Prosodically Emotional	
	Positive-semantics List *M (SD)*	Negative-semantics List *M (SD)*	Neutral List *M (SD)*	Happy-prosody List *M (SD)*	Sad-prosody List *M (SD)*
Length (letter)	5.72 (1.59)	5.72 (1.40)	5.53 (1.50)	5.63 (1.29)	5.53 (1.32)
Frequency	65 (69)	43 (84)	65 (79)	53 (60)	99 (127)
Familiarity	557[a] (44)	510[a] (56)	531 (53)	535 (58)	557 (46)
Concreteness	403[b] (122)	418[c] (98)	527[bc] (108)	540 (85)	530 (105)
Imageability	500 (89)	500 (61)	534 (97)	541 (88)	552 (86)
Orthographic neighbourhood size	4.34 (5.78)	3.72 (6.03)	4.50 (6.32)	4.03 (4.88)	5.22 (5.03)
Phonological neighbourhood size	7.50 (8.22)	7.50 (10.80)	8.13 (10.07)	8.81 (8.89)	8.53 (8.86)
Bigram Frequency	3180 (1460)	3537 (1288)	3615 (1249)	3863 (1562)	3743 (1641)

Note. a- c indicate statistically significant differences within the property, $p < .005$.

Paired samples t-tests with a Bonferroni corrected alpha level of $p = .005$ were conducted to compare the emotional and lexical properties of the semantically positive, negative, and neutral words used in Study 3. Emotionally, the positive-semantics and negative-semantics lists differed significantly on valence, $t(62) = 57.647$, $p < .001$. The positive-semantics and neutral lists differed significantly on valence, $t(62) = 19.192$, $p < .001$, and arousal, $t(62) = 5.249$, $p < .001$. The negative-semantics and neutral lists differed significantly on valence $t(62) = -33.569$, $p < .001$, and arousal, $t(62) = 8.744$, $p < .001$. Positive words were more positive than negative and neutral words, negative words were more negative than positive and neutral words, and positive and negative words were higher in arousal than neutral words. That is, the word types used in Study 3 differed as expected in terms of emotional semantics.

The positive-semantics and negative semantics lists differed lexically. Positive words were more familiar than negative words, $t(62) = 3.713$, $p < .001$, and positive words were more concrete than negative words, $t(62) = -4.291$, $p < .001$. The negative-semantics and neutral lists differed significantly on concreteness $t(62) = -4.223$, $p < .001$. Negative words were less concrete. It is not desirable that the positive words were more familiar than the negative words, or that neutral words were more concrete than the positive and negative words. While every possible effort was made to the balance the lists on these properties, it seems to be the nature of neutral words to be concrete, and positive words to be more familiar. However, these two variables should not influence any emotion-verticality mappings. In terms of concreteness, the neutral list is the baseline. The positive-semantics and negative-semantics lists do not differ on concreteness; and the positive-negative comparison is where any shifts to upper or lower space should be seen. The positive and negative emotional lists do however differ on familiarity. If familiarity results in greater activation of emotion verticality mappings then greater shifts in attention should be observed for positive words. The role of concreteness and familiarity in the results of Study 3 will be addressed in the discussion of Study 3.

Paired samples t-tests with a Bonferroni corrected alpha level of $p = .005$ were conducted to compare the emotional and lexical properties of the semantically neutral words used in Study 4. The three semantically neutral lists to be used in Study 4 (happy-prosody, sad-prosody, neutral) did not differ significantly from each other on any of the semantic or lexical variables. Aside from the concreteness and familiarity differences in the emotional semantics lists to be used in Study 3, the words selected are well controlled and balanced. Therefore, they are suitable for a stringent test of the GOOD IS UP metaphor.

STUDY 3: EMOTIONAL SEMANTICS

The aim of Study 3 was to determine if GOOD IS UP consistent shifts in attention are induced by spoken words that are semantically-emotional but not prosodically-emotional. The auditory cues were semantically negative, neutral, and positive words, spoken in neutral prosody. If processing of emotional semantics alone recruits emotion-verticality mappings, as seemingly demonstrated by studies that use written emotion words (Brookshire et al., 2010; Meier & Robinson, 2004), then participants' attention should be directed to GOOD IS UP metaphorically congruent space. Participants should be faster to respond to targets in the upper visual-field than the lower visual-field after evaluating positive words, and faster to respond to targets in the lower visual-field than the upper visual-field after evaluating negative words.

Method

Participants

Participants were 32 (29 female, 3 male; mean age 20.41 years) undergraduate students. All had normal or corrected-to-normal vision, had no hearing deficits, were right handed (as assessed by the Waterloo Handedness Questionnaire–Revised; Elias et al., 1998), and were in the sub-clinical range (participants scored no greater than 52 out of 80) on anxiety and depression (as assessed by the Zung Anxiety, 1965, and Depression Questionnaires, 1971).

Stimuli and Apparatus

See Studies 1 and 2 for details of the computer set up and stimuli used.

Procedure

As in Study 1, the participants completed a dual-task experiment. Participants performed a visual-attention task and a meaning-evaluation task. Specifically, participants heard a word, and then saw a shape. Participants were required first to make a speeded detection and identification response to the shape and subsequently a non-speeded evaluation of the word. Catch trials were included in which no shape was presented to ensure that participants did not anticipate their response to the target.

The specific details of the procedure are mostly the same as in Study 1. There are four exceptions. First, instead of tones participants heard semantically-emotional words spoken in a neutral prosody. Thus, instead of tone identification, there was a meaning-evaluation component which required participants to evaluate the meaning of the word they heard as negative, neutral, or positive by clicking on the corresponding box. The words ranged from 393 - 1013ms in length. Second, instead of two tone types (high, low) there were three word valences (negative, neutral, positive). As in Study 1, there were 96 critical trials and 24 catch trials. Therefore for the critical trials there were 32 trials that presented a negative word, 32 that presented a neutral word, and 32 that presented a positive word. Of the 32 critical trials, for each valence half (16) were presented with a short SOA (400ms) between the word and the visual target and half with a long SOA (1200ms). Of these half (8) had an upper visual-field target, half a lower visual-field target. Of these half (4 trials) presented a shape in the high-upper/lower position and half in the medium-upper/lower position. See the method section of Study 1 for visual angles. Half of the time the target was a circle, and half of the time a square. At the analysis stage the data was collapsed across high/low and medium location and shape type to give a score for the upper and lower visual-field with eight trials per condition. See Figure 3 for a visual illustration of the target-present trial makeup.

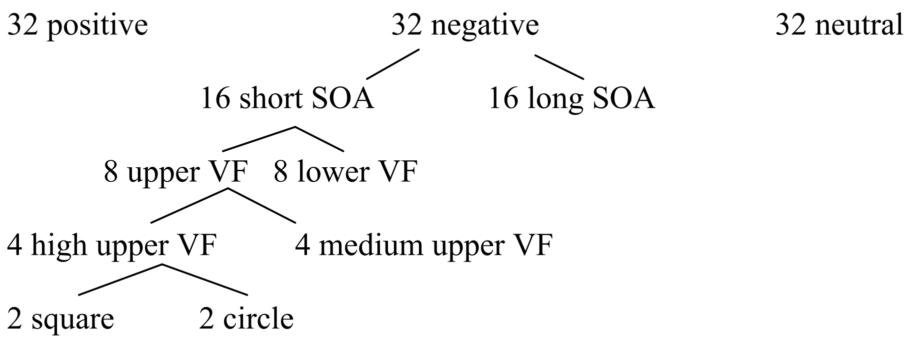

Figure 3. Target-present trials in Study 3.

SOA was manipulated across items; each item was allocated one SOA. Third, to allow assessment of activation of the GOOD IS UP conceptual metaphor at automatic processing stages, the short SOA was reduced to 400ms from the 500ms used in Study 1. The word lists were ordered alphabetically and every second word was assigned the 400ms SOA, and alternating words the 1200ms SOA. As the words ranged from 313 - 1013ms in duration, on some trials the shape could appear while the word was still being presented. See Figure 4 for an illustration of the target-present trial procedure.

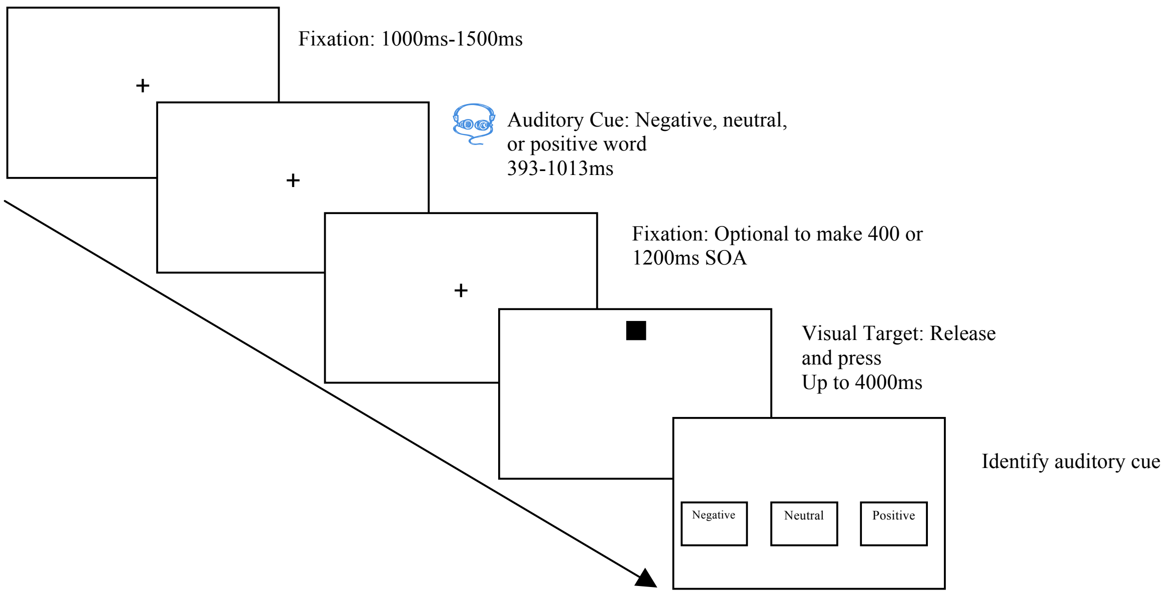

Figure 4. Target-present trial procedure in Study 3.

As in Study 1, reaction times for releases and presses from the onset of the shape were recorded using E-Prime and the computer's internal timer. In order to control for possible response mapping influences, the assignment of shape to key (square, circle; 4, 6) was counterbalanced across participants.

Before the 120 experimental trials, participants were given similar training as in Study 1. The fourth difference between Study 1 and 3 is that the number of practice trials was slightly increased to allow for even numbers of practice trials for each valenced word type. First, the participants completed twelve semantic-evaluation practice trials on which they only heard a word (four negative, neutral, and positive words) and evaluated the meaning they heard. They were given feedback on their meaning evaluation to help them understand the task demands. However, they were also instructed that there is individual variation in what people judge as positive and negative, and to respond with their own evaluation. Second, they completed twelve release-practice trials on which they might see a shape (four square, four circle, four catch) and released the 5 key upon seeing it. Third, they completed twelve press-practice trials on which they might see a shape (four square, four circle, and four catch trials) and released the 5 key upon seeing it and identified it by pressing the 4 or 6 key. For the second and third set of practice trials it was made clear to the participants that they should only release and press keys with their right index finger (i.e. not the middle or ring fingers). Fourth, they completed twelve practice trials with all trial components.

Results and Discussion

In contrast to Study 1, the evaluation component of the task did not have an objectively correct answer. In Study 3, participants were required to evaluate the meaning of the words. The "correct" answer was defined by using the ANEW ratings (Bradley & Lang, 1999). An examination of the answers given by participants for the target-present trials showed that the participants generally agreed with these ratings (M = 91%, SD = 5%). However, a closer look showed that participants agreed with the ANEW ratings more for negative words (M = 98%, SD = 2%) than for neutral words (M = 85%, SD = 13%); $t(27) = 5.306$, $p < .001$, or positive words (M = 90%, SD = 9%); $t(27) = 5.116$, $p < .001$. There was no significant difference between agreement for neutral and positive semantics; $t(27) = -1.475$, $p = .152$. Meaning is much more subjective than tones that differ consistently by 1500Hz (Study 1). In fact the conceptual-metaphor literature stresses that evaluation, or at least salience of meaning, is necessary to induce conceptual metaphoric mappings (Brookshire et al., 2010; Lakoff & Johnson, 1999; Meier & Robinson, 2004). Thus, as in Experiment 3 in Crawford et al. (2006), all subject analyses were conducted using the answer participants provided for the meaning evaluation rather than the predetermined ANEW meaning.

Four participants were removed from the analysis (see below), resulting in a sample of 28 participants[5] (27 female, 1 male; mean age 18.79 years). There were two reaction time variables: the time to release the 5 key from the onset of the shape on detection of a target (release time) and the time to press the 4 or 6 key from the onset of the shape on identification of the shape (press time). As the key release component of the shape task was a go-no-go target detection task, the number of catch trials on which participants responded was inspected. Two participants responded on more than two (out of 24) catch trials and were removed from the analysis. Times below 200ms were deemed anticipatory and times above 1500ms as prolonged detection. Therefore, the release times for the remaining participants were filtered so that only trials on which the release time was greater than 200ms and less than 1500ms were used to calculate a median release time for each SOA, evaluation, and visual-field

[5]Meier and Robinson (2004) included 28 participants in their Study 2, which closely parallels the design of the current experiment.

combination. All participants had at least 93 trials (out of 96; maximum 3% data excluded) with which to calculate a median release time.

The key press component of the shape task was a target discrimination task; participants were required to report whether they saw a square or a circle. The press times excluded trials on which participants identified the shape incorrectly and the press times were filtered so that only trials on which the press time was greater than 200ms were used to calculate a median press time for each SOA, evaluation, and visual-field combination. All participants had at least 84 trials (out of 96; maximum 12% data excluded) with which to calculate the median response times. Two participants were removed from the analysis for not meeting this criterion. See Table 5 for a summary of the subject release and press times.

Item analyses were also conducted. Unlike for the subject analysis, for the item analysis I had to use the averages for each item based on the ANEW determined semantics in order to classify the valence, not the actual evaluations given by participants (which varied for some items). Thus, there will be eight trials in each prosody x SOA x visual-field cell for the item analysis, but the number of trials in each evaluation x SOA x visual-field cell will vary in the subject analysis. The subject and item analyses will be reported together. F_1 denotes the subject analysis with data by evaluation, F_2 denotes the item analysis with data by ANEW determined semantics. See Table 6 for a summary of the item release and press times.

Table 5.

Mean (SD) subject release and press times (ms) for Study 3 by SOA, evaluation, and visual-field.

	Release Times			
	Short SOA		Long SOA	
Evaluated Emotion	Lower VF *M (SD)*	Upper VF *M (SD)*	Lower VF *M (SD)*	Upper VF *M (SD)*
Positive	522	516	461	475
	(109)	(127)	(95)	(99)
Neutral	521	512	470	459
	(119)	(130)	(102)	(97)
Negative	510	510	463	477
	(109)	(143)	(109)	(118)
	Press Times			
	Lower VF	Upper VF	Lower VF	Upper VF
Positive	781	801	693	685
	(170)	(191)	(141)	(148)
Neutral	765	756	687	703
	(155)	(170)	(131)	(170)
Negative	802	772	689	698
	(200)	(198)	(149)	(124)

Table 6.

Mean (SD) item release and press times (ms) for Study 3 by SOA, meaning, and visual-field.

	Release Times			
	Short SOA		Long SOA	
Semantic Emotion	Lower VF *M (SD)*	Upper VF *M (SD)*	Lower VF *M (SD)*	Upper VF *M (SD)*
Positive	513	502	451	462
	(20)	(28)	(14)	(21)
Neutral	504	491	458	443
	(24)	(14)	(25)	(28)
Negative	508	486	451	459
	(18)	(33)	(26)	(17)
	Press Times			
	Lower VF	Upper VF	Lower VF	Upper VF
Positive	765	746	679	658
	(44)	(24)	(44)	(34)
Neutral	744	753	682	686
	(25)	(47)	(8)	(47)
Negative	761	739	687	681
	(74)	(32)	(34)	(45)

Release Times

The median release times were analysed in 2 (SOA: 400ms, 1200ms) x 3 (Evaluation F_1/Meaning F_2: negative, neutral, positive) x 2 (Visual-field: upper, lower) repeated-measures ANOVA (F_1) and univariate ANOVA (F_2). Importantly, there was no significant evaluation x visual-field interaction, $F_1(2, 54) = 1.693$, $MSE = 1365$, $p = .194$, $\eta_p^2 = .059$, or meaning x visual-field interaction $F_2(2, 84) = .692$, $MSE = 532$, $p = .503$, $\eta_p^2 = .016$; nor was there a SOA x evaluation x visual-field interaction, $F_1(2, 54) = .507$, $MSE = 1753$, $p = .605$, $\eta_p^2 = .018$, or a SOA x meaning x visual-field interaction $F_2(2, 84) = 1.056$, $MSE = 532$, $p = .352$, $\eta_p^2 = .025$. See Figure 5 for the subject data and Figure 6 for the item data.

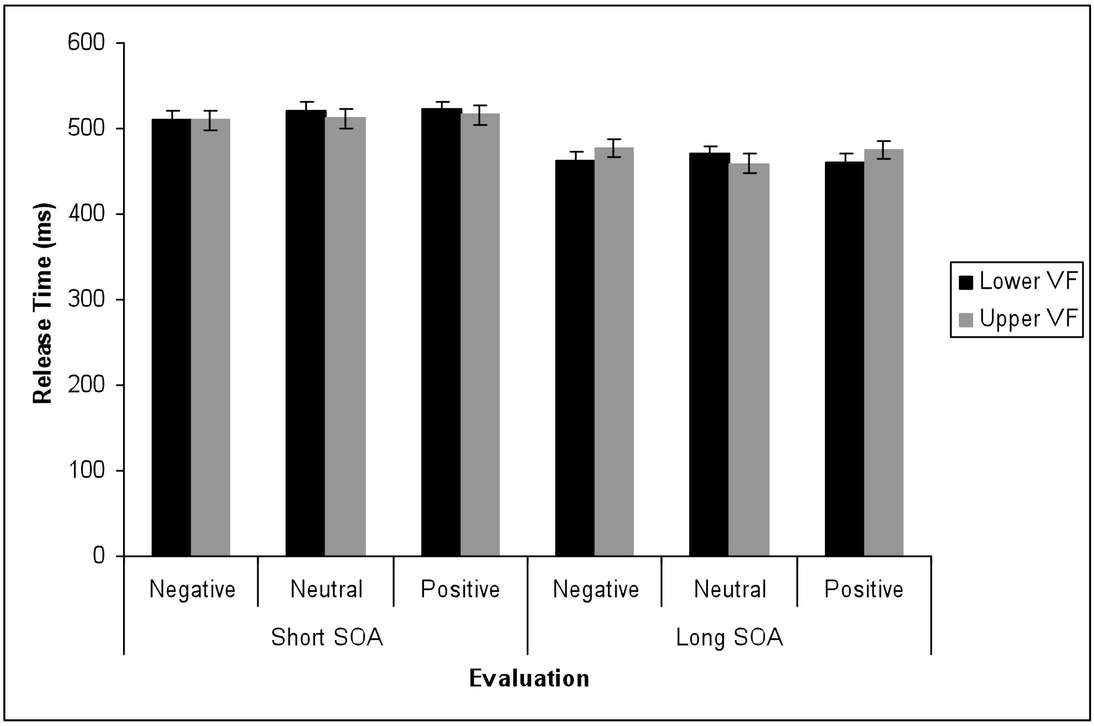

Figure 5. Subject release times for negative, neutral, and positively evaluated trials for upper and lower visual-field targets at the short and long SOA in Study 3.

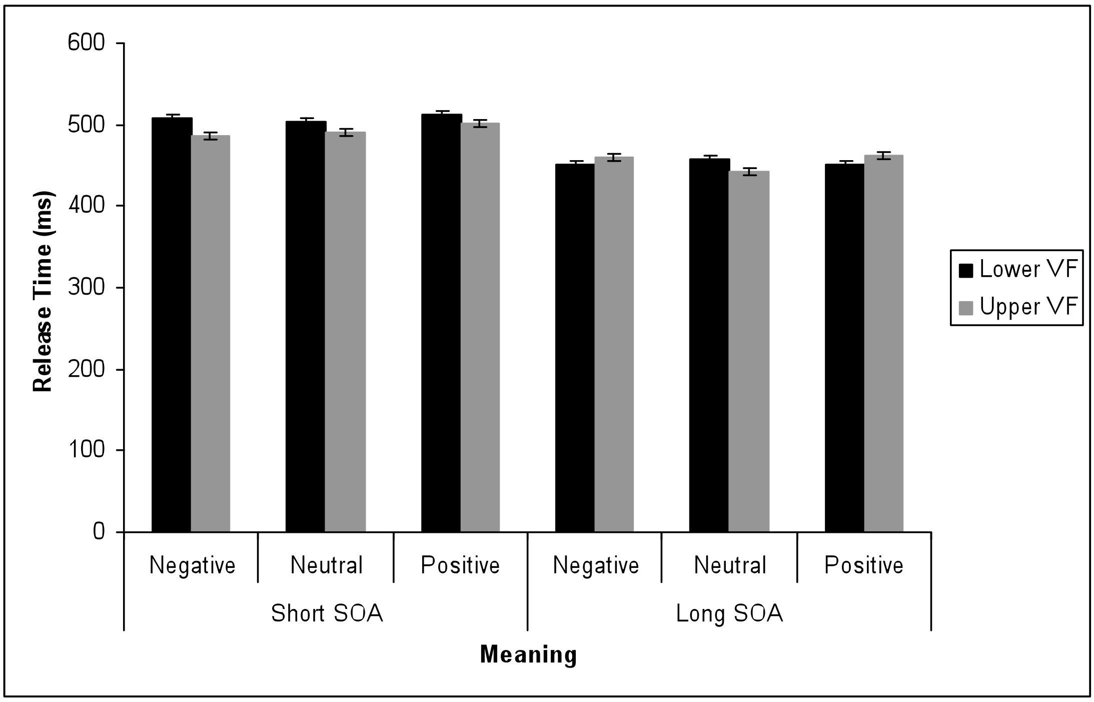

Figure 6. Item release times for negative, neutral, and positive semantics for upper and lower visual-field targets at the short and long SOA in Study 3.

The only significant effect in the release times was of SOA. Participants were significantly faster to release a key on detecting a shape on the long SOA trials (F_1 *M* = 467ms, *SD* = 98 ms; F_2 *M* = 454 ms, *SD* = 22 ms) than on the short SOA trials (F_1 *M* = 515 ms, *SD* = 118 ms; F_2 *M* = 501 ms, *SD* = 24 ms), $F_1(1, 27) = 35.856$, *MSE* = 5294, $p < .001$, $\eta_p^2 = .570$, $F_2(1, 84) = 97.873$, *MSE* = 532, $p < .001$, $\eta_p^2 = .538$.

Press Times

The median press times were analysed in 2 (SOA: 400ms, 1200ms) x 3 (Evaluation F_1/Meaning F_2: negative, neutral, positive) x 2 (Visual-field: upper, lower) repeated-measures ANOVA (F_1) and univariate ANOVA (F_2). Mirroring the release time data there was no significant evaluation x visual-field interaction, $F_1(2, 54) = .332$, *MSE* = 6632, $p = .719$, $\eta_p^2 = .012$, or meaning x visual-field interaction F_2 $(2, 84) = .919$, *MSE* = 1701, $p = .403$, $\eta_p^2 = .021$; nor was there a SOA x evaluation x visual-field interaction, $F_1(2, 54) = 1.981$, *MSE* = 4286, $p = .148$, $\eta_p^2 = .068$, or a SOA x meaning x visual-field interaction F_2 $(2, 84) = .147$, *MSE* = 1701, $p = .863$, $\eta_p^2 = .003$. See Figure 7 for the subject data and Figure 8 for the item data.

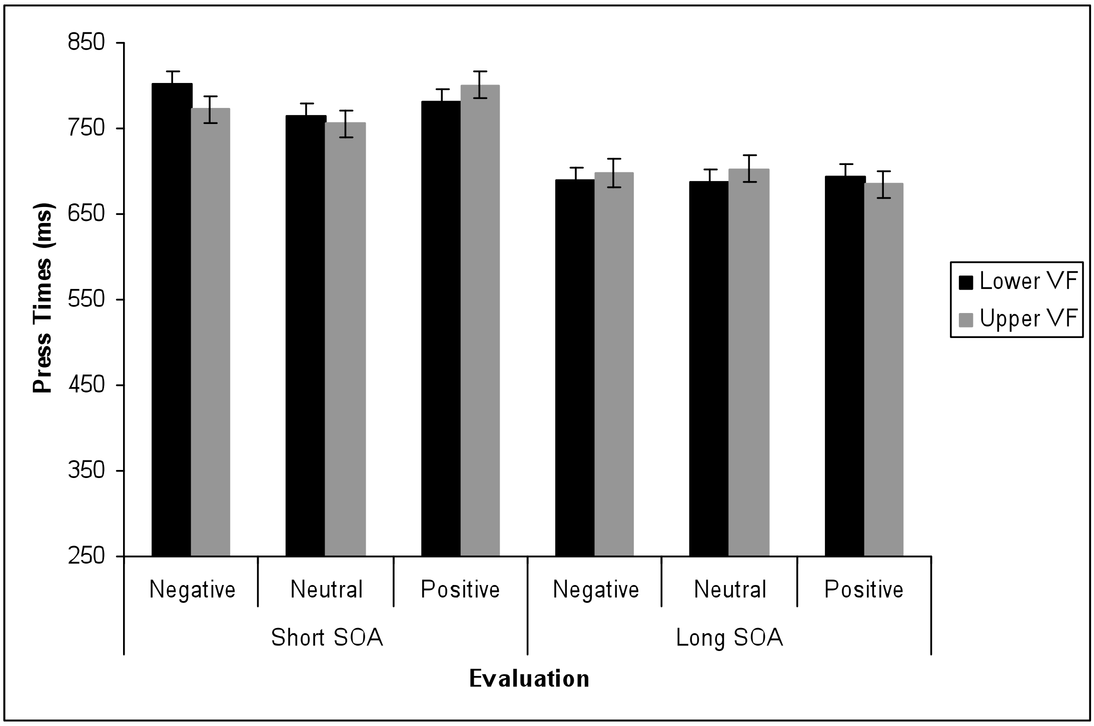

Figure 7. Subject press times for negative, neutral, and positively evaluated trials for upper and lower visual-field targets at the short and long SOA in Study 3.

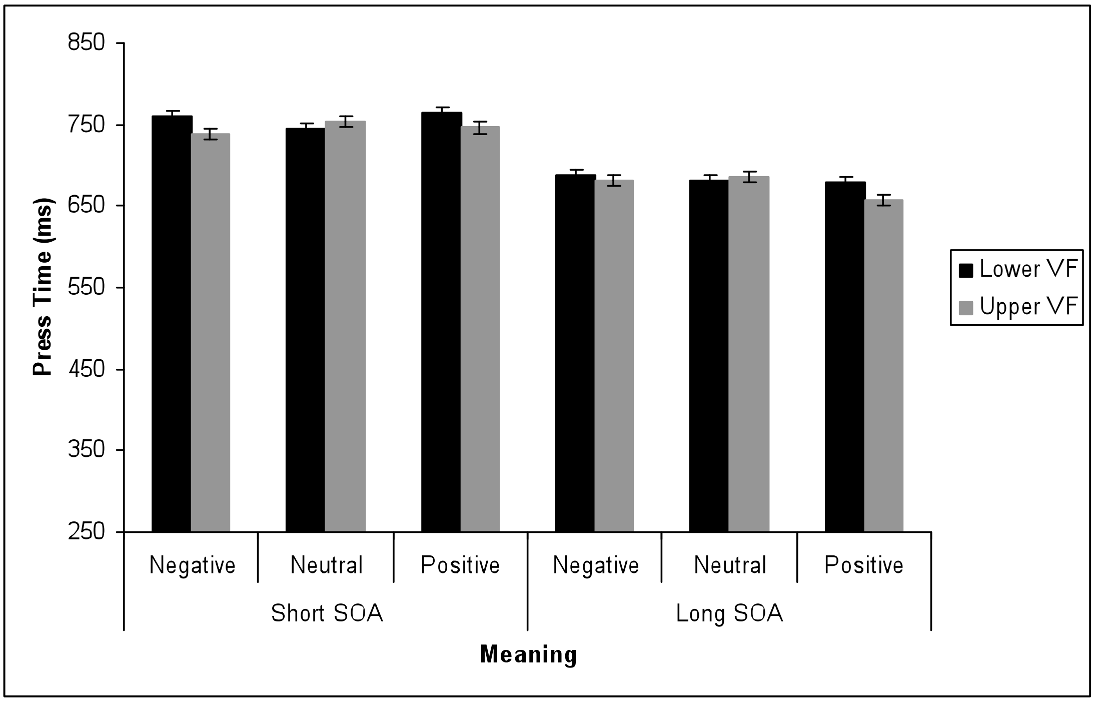

Figure 8. Item press times for negative, neutral, and positive semantics for upper and lower visual-field targets at the short and long SOA in Study 3.

Again mirroring the release time data the only significant effect was of SOA such that participants were faster to identify a shape on the long SOA trials (F_1 *M* = 693 ms, *SD* = 134 ms; F_2 *M* = 679 ms, *SD* = 37 ms) than on the short SOA trials (F_1 *M* = 779 ms, *SD* = 166 ms; F_2 *M* = 751 ms, *SD* = 43 ms), $F_1(1, 27) = 54.129$, *MSE* = 11696, $p < .001$, $\eta_p^2 = .667$, $F_2(1, 84) = 74.207$, *MSE* = 1701, $p < .001$, $\eta_p^2 = .469$.

Contrary to the hypothesis, evaluation of meaning did not induce GOOD IS UP congruent shifts in visual attention There was no significant evaluation/meaning x visual-field or SOA x evaluation/meaning x visual-field interactions in the release or press times. Study 3 failed to replicate the results of Meier and Robinson (2004). Possible reasons for the lack of replication of the GOOD IS UP congruent shifts in attention with spoken language will be addressed in the General Discussion.

As in Study 1, with high- and low-pitch tones, participants were faster to make both detection and identification responses to shapes for trials with a long SOA (1200ms) than a short SOA (400ms). This was evident at both the subject-evaluation and item-ANEW meaning levels. As discussed, the psychological refractory period (PRP) effect could account for this long SOA advantage. Participants cannot begin the response-selection process for the second stimulus (in this case the shape) until a response has been selected, but not necessarily produced for the first stimulus (in this case the word). Thus, the long SOA advantage observed in Study 3 suggests that even though participants were not required to evaluate the word until after the shape task, they were evaluating and selecting their response before they responded to the shapes. However, evaluation of meaning did not seem to activate perceptual shifts congruent with the GOOD IS UP conceptual metaphor.

In Study 2, it was noted that the neutral words were more concrete than the positive and negative words, and that the positive words were more familiar than the neutral and negative words. It is unlikely that such differences could be masking any GOOD IS UP congruent shifts in attention. Rather, it would be expected that greater concreteness and familiarity would enhance any verticality-emotion mappings induced by processing the neutral and positive words respectively. That is, higher concreteness and familiarity could enhance, or indeed be the result of, a stronger link between the perceptuomotor source domain, verticality, and the conceptual target domain, emotion. This is not the case. There was no difference between the upper and lower visual fields for any of the evaluated word types in the direction predicted by Conceptual Metaphor Theory for the GOOD IS UP conceptual metaphor.

STUDY 4: EMOTIONAL PROSODY

Study 4 is the converse of Study 3. The aim was to determine if GOOD IS UP congruent shifts in attention are induced by spoken words that are prosodically emotional but semantically neutral. The auditory cues were semantically neutral words, spoken in negative (sad), neutral, and positive (happy) prosodies. If processing of emotional prosody alone recruits GOOD IS UP mappings then participants' attention should be shifted to metaphorically congruent space. Participants should be faster to respond to targets in the upper visual-field than in the lower visual-field after evaluating happy prosody, and faster to respond to targets in the lower visual-field than in the upper visual-field after evaluating sad prosody.

Method

Participants

Participants were 38 (27 female, 11 male; mean age 20.24 years) undergraduate students. All had normal or corrected-to-normal vision, had no hearing deficits, were right handed (as assessed by the Waterloo Handedness Questionnaire–Revised; Elias et al., 1998), and were in the sub-clinical range (not greater than 56 out of 80) on anxiety or depression (as assessed by the Zung Anxiety, 1965, and Depression Questionnaires, 1971).

Stimuli and Apparatus

See Studies 1, 2, and 3 for details of the computer set up and stimuli used.

Procedure

As in Studies 1 and 3, the participants completed a dual-task paradigm. Participants performed a visual-attention task and a prosodic-evaluation task. Specifically on critical trials, participants heard a word, and then saw a shape. They were first required to make a speeded detection and identification response to the shape and subsequently a non-speeded evaluation of the prosody. Catch trials were included in which no shape was presented.

The specific details of the procedure were almost identical to Study 3. There are three exceptions. First, instead of semantically-emotional words spoken in neutral prosody, participants heard semantically-neutral words spoken in emotional prosody. The tokens ranged from 407 - 1047ms in length. Second, instead of meaning evaluation, there was a prosodic-evaluation component which required participants to evaluate the prosody of the word they heard as sad, neutral, or happy by clicking on the corresponding box. Third, in the practice trials, participants practiced evaluating the valence of the prosody. The break-down of trial numbers by prosodic valence, SOA, shape position, and shape type was the same as in Study 3. See Figure 9 for an illustration of the target-present trial procedure.

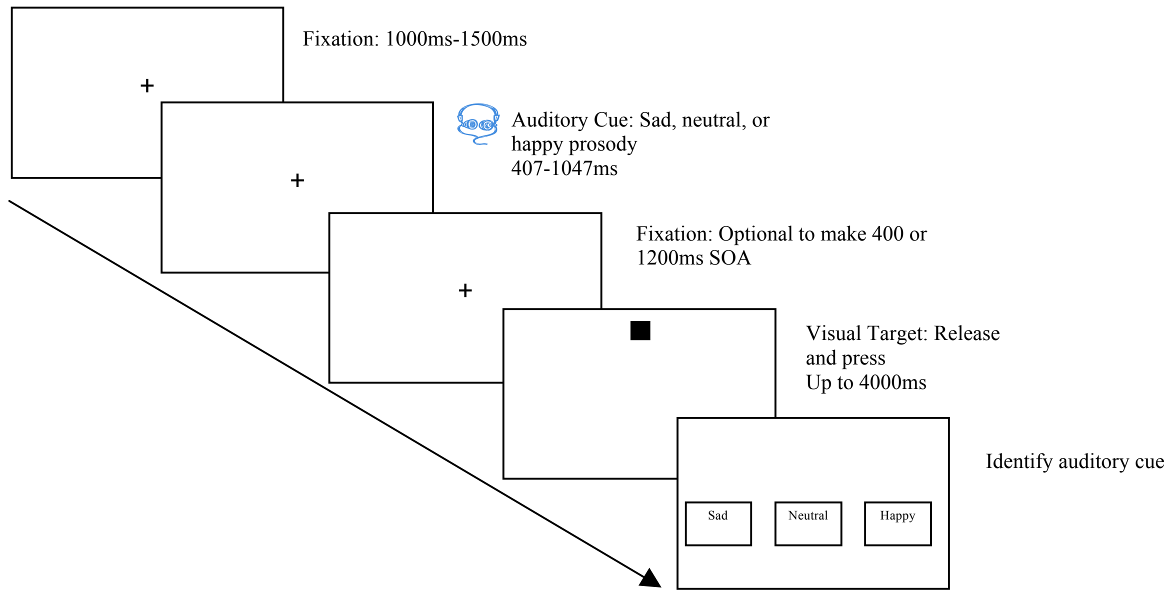

Figure 9. Target-present trial procedure in Study 4.

Results and Discussion

As in Study 3, the evaluation component of the task did not have an objectively correct answer. The “correct” answer was defined as the prosody the majority of participants in Study 2 identified the token as being spoken in. Participants generally agreed with the identification (*M* = 92%, *SD* = 6%), although there was higher agreement for the happy prosody (*M* = 95%, *SD* = 7%) versus the sad prosody (*M* = 89%, *SD* = 10%), $t(27) = -3.567$, $p = .001$. Neutral prosodic agreement (*M* = 91%, *SD* = 7%) did not differ from happy or sad. To be consistent with the subject analyses conducted in Study 3, all subject analyses were conducted using the answer participants gave for the prosodic evaluation rather than the consensus identification.

Ten participants were removed from the analysis (see below), resulting in a sample of 28 participants (20 female, 8 male; mean age 20.64 years)[6]. As in Study 3, there were two reaction time variables: the time to release the 5 key on detection of a target (release time); and the time to press the 4 or 6 key for identification of the shape (press time). As the key release component of the shape task was a go-no-go target detection task, the number of catch trials on which participants responded was inspected. Seven participants responded on more than two (out of 24) catch trials and

[6]Twenty-eight is the same number of participants as in Study 3, and in Meier and Robinson’s (2004) Study 2.

were removed from the analysis. Times below 200ms were deemed anticipatory and times above 1500ms as prolonged detection. Therefore, the release times for the remaining participants were filtered so that only trials on which the release time was greater than 200ms and less than 1500ms were used to calculate a median release time for each SOA, evaluation, and visual-field combination. All participants had at least 91 trials (out of 96; maximum 5% data excluded) with which to calculate a median release time.

The key press component of the shape task was a target discrimination task; participants were required to report whether they saw a square or a circle. The press times excluded trials on which participants identified the shape incorrectly and were filtered so that only trials on which the press time was greater than 200ms were used to calculate a median press time for each SOA, prosody, and visual-field combination. All participants had at least 83 trials (out of 96; maximum 14% data excluded) with which to calculate the median press times. Three participants were removed from the analysis for not meeting this criterion. See Table 7 for a summary of the subject release and press times.

Item analyses were also conducted. Unlike the subject analysis, the item analysis used the averages for each item based on the consensus prosody from Study 2 in order to classify the valence, not the actual evaluations given by participants. The subject and item analyses will be reported together. F_1 denotes the subject analysis with data by evaluation, F_2 denotes the item analysis with data by consensus. See Table 8 for a summary of the item release and press times.

Table 7.

Mean (SD) subject release and press times (ms) for Study 4 by SOA, prosody-evaluation, and visual-field.

	Release Times			
	Short SOA		Long SOA	
Evaluated Prosody	Lower VF *M (SD)*	Upper VF *M (SD)*	Lower VF *M (SD)*	Upper VF *M (SD)*
Happy	502 (112)	490 (106)	460 (91)	446 (84)
Neutral	517 (109)	504 (109)	459 (87)	460 (81)
Sad	528 (115)	513 (92)	442 (68)	459 (89)
	Press Times			
	Lower VF	Upper VF	Lower VF	Upper VF
Happy	829 (231)	822 (213)	733 (178)	753 (196)
Neutral	894 (300)	810 (228)	746 (170)	798 (223)
Sad	856 (222)	853 (249)	748 (197)	728 (175)

Table 8.

Mean (SD) item release and press times (ms) for Study 4 SOA, consensus-prosody, and visual-field.

	Release Times			
	Short SOA		Long SOA	
Consensus Prosody	Lower VF *M (SD)*	Upper VF *M (SD)*	Lower VF *M (SD)*	Upper VF *M (SD)*
Happy	485	494	465	445
	(30)	(31)	(30)	(15)
Neutral	510	494	450	453
	(27)	(30)	(11)	(28)
Sad	508	514	443	451
	(28)	(44)	(21)	(16)
	Press Times			
	Lower VF	Upper VF	Lower VF	Upper VF
Happy	783	753	708	709
	(58)	(29)	(35)	(19)
Neutral	837	785	712	740
	(99)	(36)	(52)	(39)
Sad	822	817	704	703
	(63)	(89)	(32)	(35)

Release Times

The median release times were analysed in a 2 (SOA: 400ms, 1200ms) x 3 (Evaluation F_1/Consensus F_2: negative, neutral, positive) x 2 (Visual-field: upper, lower) repeated-measures ANOVA (F_1) and univariate ANOVA (F_2). There was no evaluation x visual-field interaction $F_1(2, 54) = .727$, $MSE = 1769$, $p = .488$, $\eta_p^2 = .026$, or a consensus x visual-field interaction $F_2(2, 84) = .608$, $MSE = 741$, $p = .547$, $\eta_p^2 = .014$, nor was there a SOA x evaluation x visual-field interaction $F_1(2, 54) = 1.353$, $MSE = 1463$, $p = .267$, $\eta_p^2 = .048$, or a SOA x consensus x visual-field interaction $F_2(2, 84) = 1.630$, $MSE = 741$, $p = .202$, $\eta_p^2 = .037$. See Figures 10 and 11 for the subject and item data displayed by SOA, valence, and visual-field.

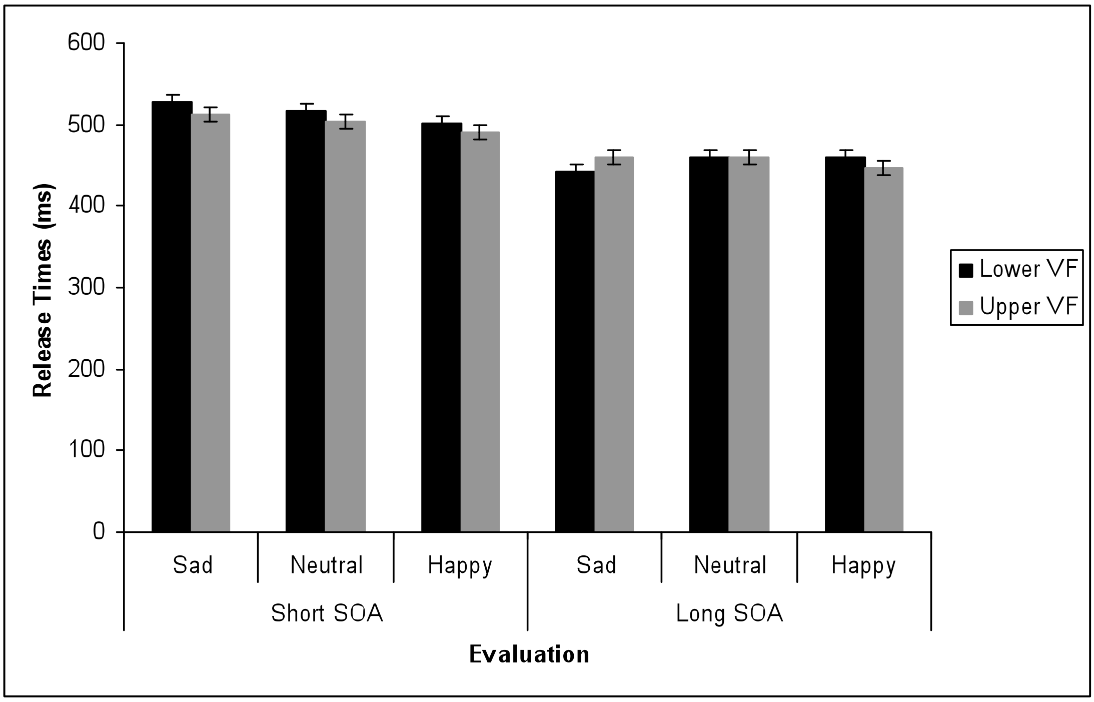

Figure 10. Subject release times for trials evaluated as sad, neutral, and happy for upper and lower visual-field targets at the short and long SOA in Study 4.

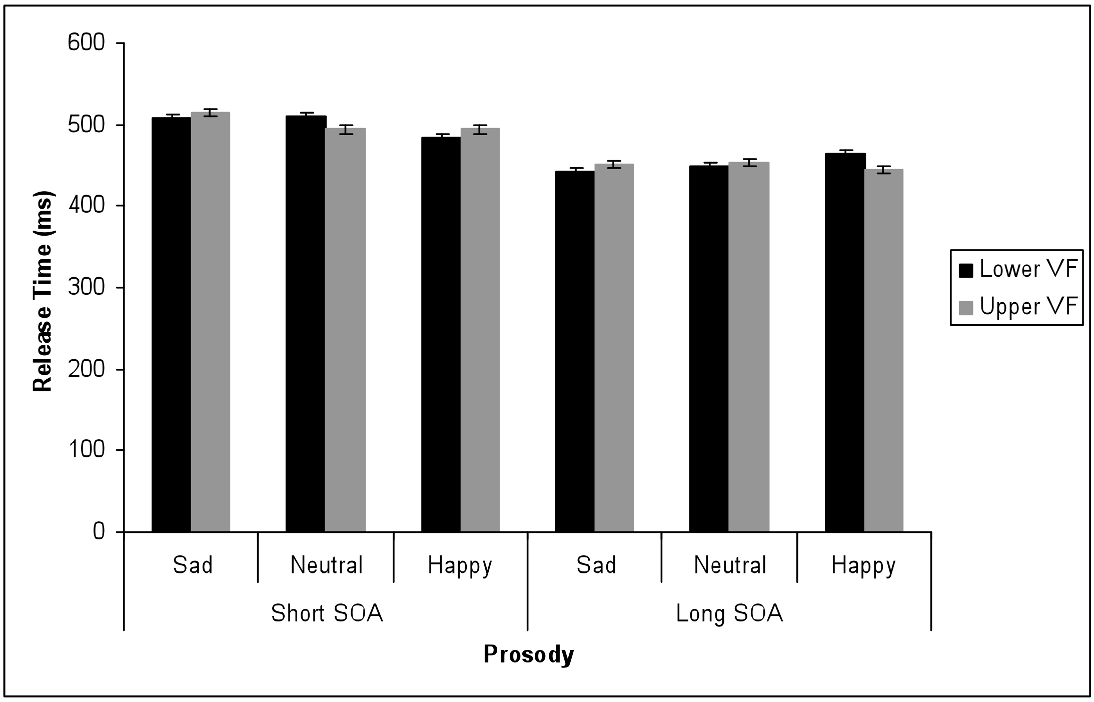

Figure 11. Item release times for sad, neutral, and happy prosody for upper and lower visual-field targets at the short and long SOA in Study 4.

There was a significant effect of SOA at both the subject and item level. Participants were significantly faster to release a key on detecting a shape on the long SOA trials (F_1 *M* = 452ms, *SD* = 76ms; F_2 *M* = 451ms, *SD* = 21ms) than on the short SOA trials (F_1 *M* = 504ms, *SD* = 94ms; F_2 *M* = 501ms, *SD* = 32ms), $F_1(1, 27) = 34.183$, *MSE* = 7433, $p < .001$, $\eta_p^2 = .559$, $F_2\,(1, 84) = 79.526$, *MSE* = 741, $p < .001$, $\eta_p^2 = .486$.

The main effect of SOA was qualified by both evaluation and visual-field. These two interactions were significant at the subject level but not the item level. First, there was a SOA x evaluation interaction $F_1(2, 54) = 4.681$, *MSE* = 1127, $p = .013$, $\eta_p^2 = .148$, see Figure 12, but not a SOA x consensus interaction $F_2(2, 84) = 2.303$, *MSE* = 741, $p = .106$, $\eta_p^2 = .052$. In the subject data the SOA x evaluation interaction was driven by a difference in release times at the short SOA. Participants were significantly faster to release on detection of a target on short SOA trials where they evaluated the prosody as happy (*M* = 491 ms, *SD* = 96 ms) than as sad (*M* = 518 ms, *SD* = 94 ms), $t(27) = 3.218$, $p = .003$, or neutral (*M* = 510 ms, *SD* = 103 ms) $t(27) = 2.395$, $p = .024$. This facilitation of release times for short SOA trials on which the prosody was evaluated as happy could be due to the PRP effect. Participants were more likely to agree with the prosodic identification scores from Study 2 for happy prosody than sad prosody. If prosodic-evaluation consensus (see Table 5) is taken as an index of ease of prosodic-evaluation response selection under dual-task conditions, then participants found the evaluation component easier for words spoken in happy prosody. This could have lead to a reduction in the time needed to select the evaluation response for happy trials, which would have reduced the cognitive-bottleneck at response selection for task 1 (prosodic evaluation) and resulted in quicker release times for task 2 (shape detection). Importantly, the SOA x evaluation interaction was not qualified by visual-field, and is not evidence of GOOD IS UP congruent shifts in attention.

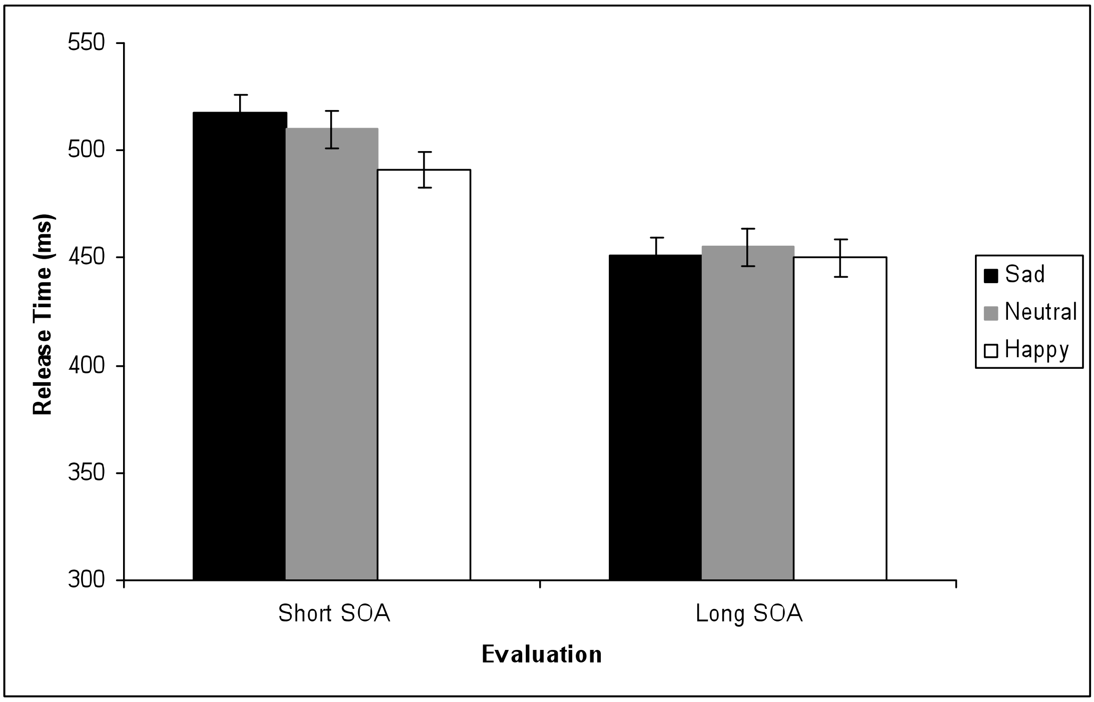

Figure 12. Subject release times by SOA and evaluation in Study 4.

Second, there was a SOA x visual-field interaction at the subject level F_1(1, 27) = 4.846, *MSE* = 887, *p* = .036, η_p^2 = .152, see Figure 13, but not at the item level F_2(1, 84) = .051, *MSE* = 741, *p* = .823, η_p^2 = .001. Though neither comparison was significant, the interaction at the subject level was driven by different visual-field biases at the short and long SOA. At the short SOA participants were non-significantly faster to respond to targets in the upper visual-field (*M* = 502 ms, *SD* = 98 ms) than in the lower visual-field (*M* = 507 ms, *SD* = 93 ms), *t*(27) = .750, *p* = .460. At the long SOA participants were non-significantly faster to respond to targets in the lower-visual field (*M* = 451 ms, *SD* = 78 ms) than in the upper visual-field (*M* = 454 ms, *SD* = 80 ms), *t*(27) = -.589, *p* = .561. Importantly, the SOA x visual-field interaction was not qualified by evaluation, and is not evidence of GOOD IS UP congruent shifts in attention.

As both the SOA x evaluation and SOA x visual-field interactions were significant at the subject but not at the item level, this suggests that the patterns of responding do not generalise well across items.

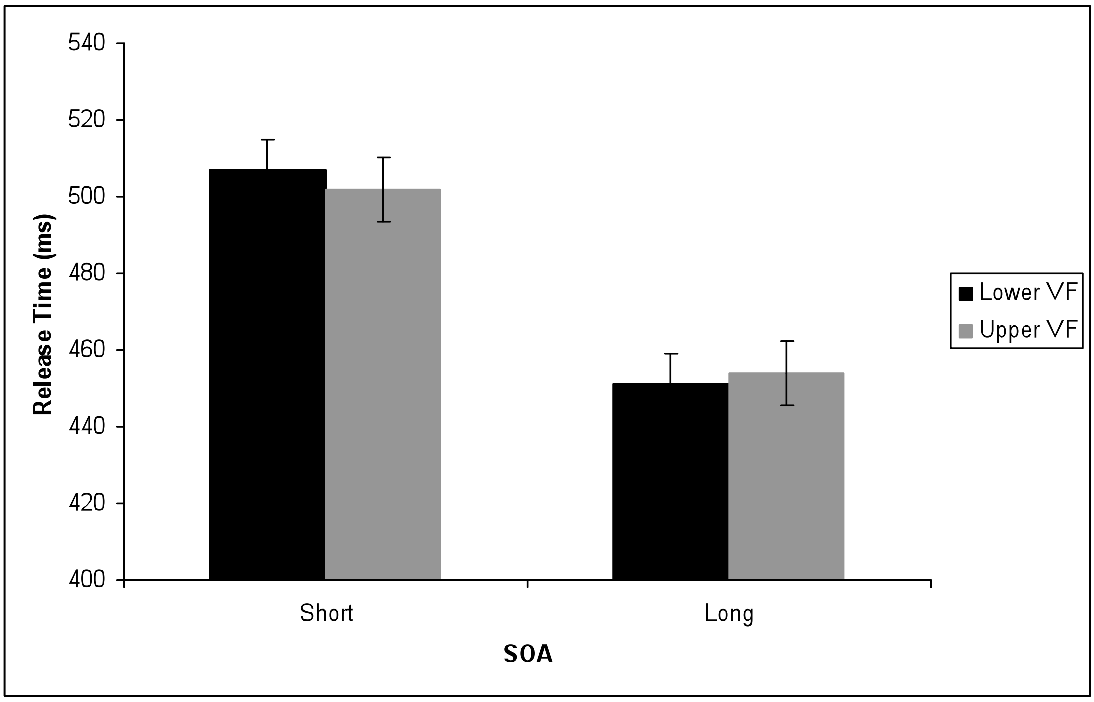

Figure 13. Subject release times by SOA and visual-field in Study 4.

Press Times

The median press times were analysed in a 2 (SOA: 400ms, 1200ms) x 3 (Evaluation F_1/Prosody F_2: negative, neutral, positive) x 2 (Visual-field: upper, lower) repeated-measures ANOVA (F_1) and univariate ANOVA (F_2). There was no evaluation x visual-field interaction $F_1(2, 54) = .503$, $MSE = 7893$, $p = .607$, $\eta_p^2 = .018$, or a prosody x visual-field interaction $F_2(2, 84) = .097$, $MSE = 2954$, $p = .907$, $\eta_p^2 = .002$, nor was there a SOA x evaluation x visual-field interaction $F_1(1.480, 39.963) = 3.118$, $MSE = 18732$, $p = .069$, $\eta_p^2 = .104$ (Greenhouse-Geisser corrected)[7], or a SOA x prosody x visual-field interaction $F_2(2, 84) = .979$, $MSE = 2954$, $p = .380$, $\eta_p^2 = .023$. See Figures 14 and 15 for the subject and item data displayed by SOA, valence, and visual-field.

[7]That this interaction approaches significance is probably due to shifts in attention on trials evaluated as neutral at the short and long SOA, see Figure 14. At the short SOA there was an upper visual-field advantage and at the long SOA a lower visual-field advantage for trials evaluated as neutral prosody. This is not consistent with a GOOD IS UP shift in attention.

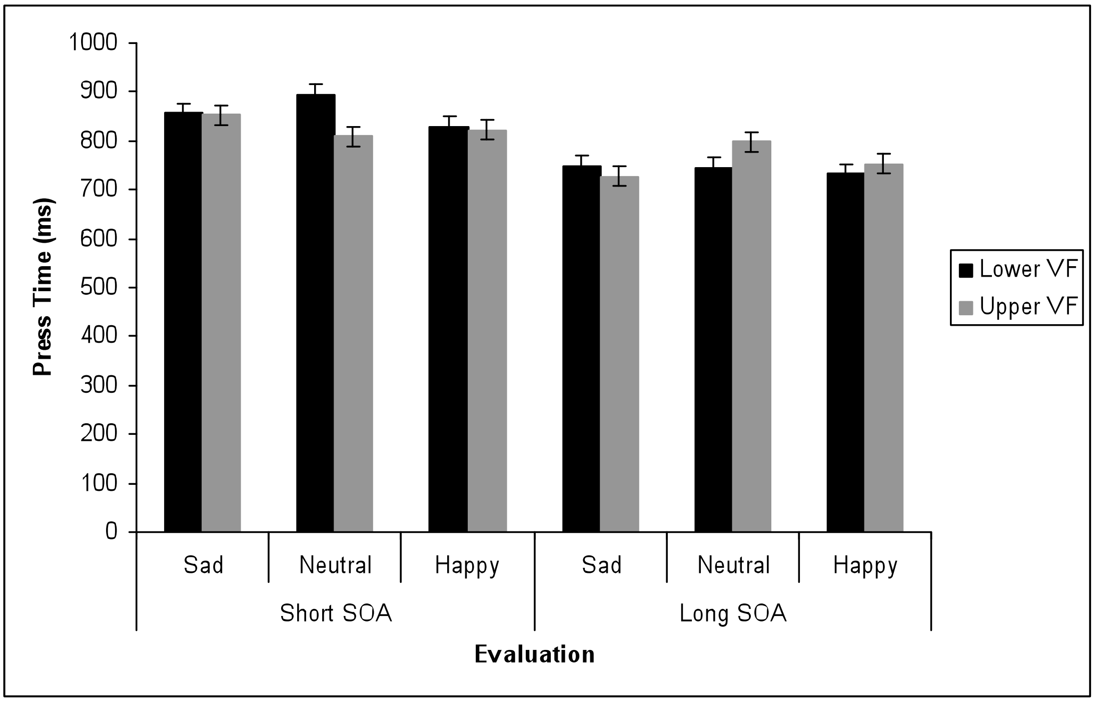

Figure 14. Subject press times for trials evaluated as sad, neutral, and happy prosody for upper and lower visual-field targets at the short and long SOA in Study 4.

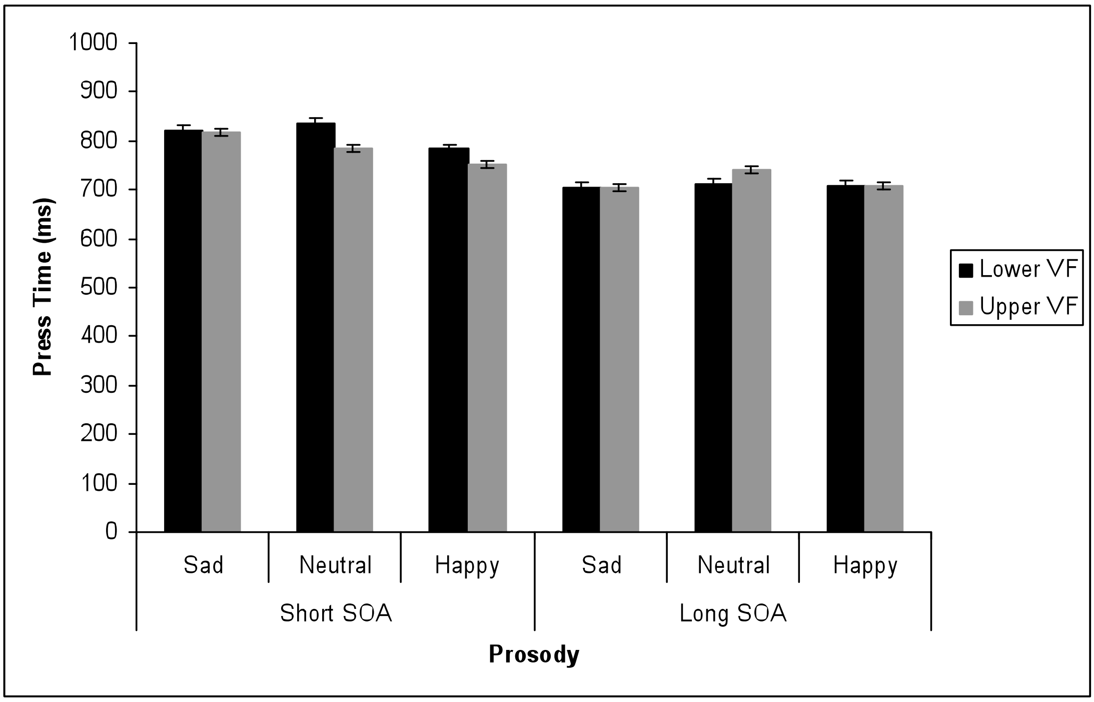

Figure 15. Item press times for sad, neutral, and happy prosody for upper and lower visual-field targets at the short and long SOA in Study 4.

There was a significant effect of SOA at both the subject and item level. Participants were significantly faster to press a key on identifying a shape at the long SOA (F_1 *M* = 740 ms, *SD* = 181 ms; F_2 *M* = 713, *SD* = 37) than at the short SOA (F_1 *M* = 826 ms, *SD* = 214 ms; F_2 *M* = 800 ms, *SD* = 70 ms), $F_1(1, 27) = 47.117$, $MSE = 15422$, $p < .001$, $\eta_p^2 = .636$, $F_2(1, 84) = 61.503$, $MSE = 2893$, $p < .001$, $\eta_p^2 = .423$.

The only other significant effect in the press times was that, as in the release times, the main effect of SOA was qualified by visual-field. While the SOA x visual-field interaction was significant at the subject level, $F_1(1, 27) = 9.679$, $MSE = 5047$, $p = .004$, $\eta_p^2 = .264$ (see Figure 16), it was not significant at the item level, $F_2(1, 84) = 2.895$, $MSE = 2954$, $p = .093$, $\eta_p^2 = .033$. The interaction at the subject level is driven by different visual-field biases at the short and long SOA. At the short SOA participants were significantly faster to identify targets in the upper visual-field (*M* = 816, *SD* = 219) than in the lower visual-field (*M* = 841 ms, *SD* = 206 ms), $t(27) = 2.188$, $p = .037$. At the long SOA participants were non-significantly faster to identify targets in the lower-visual field (*M* = 738 ms, *SD* = 177 ms) than in the upper visual-field (*M* = 745, *SD* = 181), $t(27) = -.674$, $p = .506$. Importantly, as in the release times, the SOA x visual-field interaction was not qualified by evaluation, and is not suggestive of a GOOD IS UP congruent shift in attention.

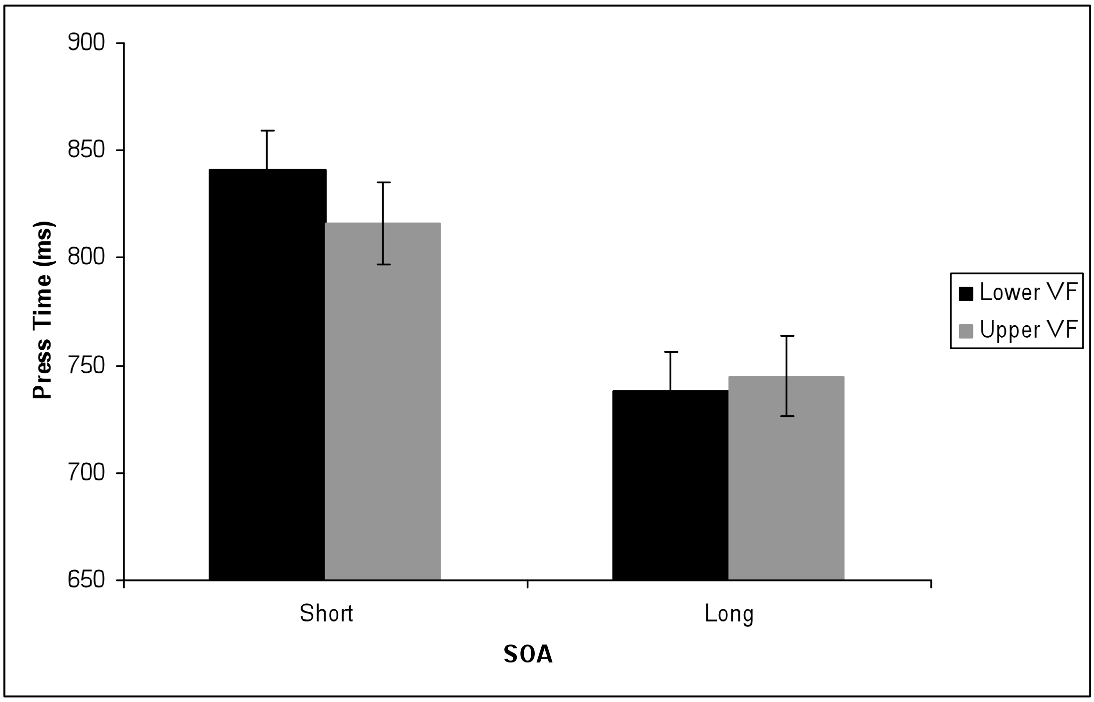

Figure 16. Subject press times by SOA and visual-field in Study 4.

Contrary to the hypothesis, evaluation of emotional-prosody did not induce shifts in visual attention consistent with the GOOD IS UP conceptual metaphor. There was no significant evaluation/prosody x visual-field or SOA x evaluation/prosody x visual-field interactions in the release or press times when examining data by subject or by item. Participants were not faster to respond to targets in the upper visual-field than in the lower visual-field after evaluating happy prosody, and were not faster to respond to targets in the lower visual-field than in the upper visual-field after evaluating sad prosody. Thus, Study 4 demonstrates that evaluation of prosodically emotional words does not seem to activate a GOOD IS UP conceptual metaphor. The lack of shifts in attention in Study 3, with emotional semantics, and in Study 4, with emotional prosody, suggests that serious thought needs to be given as to whether a GOOD IS UP conceptual metaphor is recruited to process emotional semantics or prosody in spoken language.

As in Studies 1 and 3, participants were faster to make both detection and identification responses to shapes on trials with a long SOA than with a short SOA. As discussed, the psychological refractory period (PRP) effect could account for this long SOA advantage. Participants cannot begin the response-selection process for the second stimulus (in this case the shape) until a response has been selected, but not necessarily produced for the first stimulus (in this case the evaluation of the prosody). Thus, the long SOA advantage observed in Study 4 suggests that even though participants were not required to respond to the prosody until after the shape task, they were most probably evaluating it and selecting their response before they responded to the shape. The presence of a PRP effect in this paradigm is also supported by the SOA x evaluation interaction in the subject release-times. It appears that participants were facilitated on response-selection for the prosodic-evaluation task for happy-prosody, perhaps due to the greater salience or ease of processing of the happy-prosody, resulting in a reduction of the response-selection bottleneck, which was reflected in faster release times for the for happy-prosody trials at the short SOA.

In Study 4, there was also a significant SOA x visual-field interaction in the subject release and press times. Participants were biased to attend to the upper visual-field at the short SOA, and to the lower visual-field at the long SOA. One explanation for this effect is that attention sweeps up and down, over the time course of evaluative processing. It is important to note that the SOA x evaluation and the SOA x visual-field interactions were only present at the subject level. Furthermore, the SOA x evaluation interaction was not qualified by visual-field and the SOA x visual-field interaction was not qualified by evaluation. Also note, that the SOA x evaluation and SOA x visual-field interactions were not consistent with activation of a GOOD IS UP conceptual metaphor and were not observed in Study 3 with semantically emotional spoken words.

General Discussion

If representation is achieved through conceptual metaphors, then activation of the GOOD IS UP conceptual metaphor should be required for the evaluation of emotional words, and metaphor congruent shifts in attention should be observed. In their 2004 study, Meier and Robinson reported GOOD IS UP congruent shifts in attention. Participants read written semantically positive and negative words, evaluated them, and then responded to a target in the upper or lower visual-field. Participants were faster to respond to targets in the upper visual-field than in the lower visual-field after evaluating positive words, and were faster to respond to targets in the lower visual-field than in the upper visual-field after evaluating negative words. Such shifts in attention suggest that processing emotional words activates a GOOD IS UP conceptual metaphoric representation.

GOOD IS UP congruent shifts in attention were not replicated in this thesis with spoken emotional words. The paradigm required participants to listen to spoken words which were emotional in terms of their semantics or in terms of their prosody, then to detect and identify a visual target presented in the upper or lower visual-field, and then to evaluate the word. In Study 3, in which participants evaluated spoken semantically-positive and negative words, no shifts in attention were observed at the short (chosen to index automatic processing stages) or the long SOA (chosen to index controlled processing stages). Participants were not faster to detect targets in the upper visual-field than in the lower visual-field when evaluating words as semantically positive, and were not faster to detect targets in the lower visual-field than in the upper visual-field when evaluating words as semantically negative. In Study 4, in which participants evaluated happy and sad prosodies, no shifts in attention were observed at the short or the long SOA. Participants were not faster to detect targets in the upper visual-field than in the lower visual-field when evaluating prosody as happy, and were not faster to detect targets in the lower visual-field than in the upper visual-field when evaluating prosody as sad. The lack of GOOD IS UP congruent shifts in attention at the short or long SOA is contrary to Meier and Robinson's predictions of conceptual metaphor congruent perceptual processing and automaticity.

If metaphor congruent shifts in attention are consistently not observed during spoken word processing, this would suggest that conceptual metaphor representation is not obligatory and a main prediction of Conceptual Metaphor Theory is violated. It is interesting that GOOD IS UP congruent shifts in attention, which seem to be robust when induced with written words (Brookshire et al., 2010; Casasanto, 2008, as cited in Brookshire et al., 2010; Meier & Hauser, 2008, as cited in Crawford, 2009; Meier & Robinson, 2004, 2006; Weger et al., 2007), were not observed with a paradigm using spoken words. Given the claim repeatedly made by grounded cognition theorists (e.g. Barsalou, 1999; Lakoff & Johnson, 1999), that the evolutionarily older perceptuomotor system underlies representation, and given that spoken language is an evolutionarily older cognitive process than writing, it was expected that evidence of conceptual metaphoric representation should be easily observed in spoken language processing. Perhaps source-target domain mappings are not activated universally across language modalities. If that is the case, Conceptual Metaphor Theory needs to be revised. However, before addressing the theoretical issues raised by the use of spoken language, it is necessary to scrutinise the methodology used in this thesis.

Methodology

Unlikely methodological explanations.

It is not likely that the paradigm used in this thesis was ill suited for assessing metaphoric shifts in attention. First, lack of statistical power is unlikely to have contributed to the failure to find GOOD IS UP congruent shifts in attention in Studies 3 and 4. The number of participants was sufficient. Metaphor congruent shifts were observed in Study 1 with 20 participants and there were 28 participants in each of Studies 3 and 4; the same number of participants Meier and Robinson (2004) recruited for their Study 2, which mirrored the design in this thesis. There were fewer trials for each of the valences in Studies 3 and 4 (32 trials for each of positive/happy, neutral, and negative/sad) than for the two tone types in Study 1 (50 trials), however, 32 is still a reasonably high trial count with which to calculate a mean reaction time. Although Meier and Robinson had trial counts of 50 for each of the positive and negative valences, GOOD IS UP congruent perceptual processing has been observed with as few as 30 trials per valence type (see Crawford et al., 2006). Importantly, a visual inspection of the standard errors of the valence effect in Meier and Robinson's Figure 2 revealed that the standard errors of the valence effect for Study 3 and 4 are smaller than theirs. Thus, the analyses reported in this thesis are in fact more powerful than those reported by Meier and Robinson.

Second, it is evident in Study 1 that the paradigm used in this thesis was sensitive to metaphoric shifts in attention. The pattern of responding in the release times, when the cues were high and low tones, was congruent with the HIGH PITCH IS UP conceptual metaphor (e.g. Evans & Treisman, 2010). Participants were significantly faster to respond to targets in the upper visual-field than in the lower visual-field after identifying high pitched tones, and were non-significantly faster to respond to targets in the lower visual-field than in the upper visual-field after identifying low pitched tones. This indicates that the attention task was sensitive to metaphoric congruent shifts in attention.

Third, although the data was analysed in such a way that the trial number in each cell was sometimes reduced, it is unlikely that the choice of data analysis masked any GOOD IS UP shifts in attention. Data was excluded at a trial level for each participant. A small number of trials were removed for the release times, a maximum of five percent in Study 4. Three percent was the highest proportion of trials for the release times removed for a participant in Study 1 and 3. Yet shifts in attention were observed in Study 1, but not in Study 3. A greater proportion of trials were removed for the press times. However, given the shifts in attention observed in the release times for Study 1 and the susceptibility of the press times to stimulus-response compatibility effects, it was expected that if shifts in attention were induced they would be observed in the release times. In Studies 3 and 4 the data was examined by evaluation which may have resulted in some cells with a small number of trials. Importantly, metaphor congruent processing has been observed by Crawford et al. (2009) when analysing data by evaluation. Furthermore, in this thesis no shifts in attention were observed when the data was examined by item; and the item analyses kept the number of trials in each SOA x semantics/prosody x visual-field cell evenly at eight.

As the statistical power is sufficient, the paradigm is sensitive to shifts in attention, and the choice of data analysis is unlikely to be masking shifts in attention, the lack of GOOD IS UP congruent shifts in attention in this thesis seems to be reliable.

Comparison to Meier and Robinson (2004).

In order to perform a more stringent test of Conceptual Metaphor Theory I made a number of changes from the paradigm used by Meier and Robinson (2004) in their Study 2. Although making a large number of design changes at once goes against conventional wisdom in experimental design, the goal was to eliminate as many potential confounds as possible, and conduct a clean set of studies. Identifying the change (or changes) that were potentially responsible for the difference in patterns of responding induced by written and spoken emotional words would help to identify the boundary conditions under which conceptual metaphors play a role in language processing. In fact, identifying the change (or changes) which resulted in the null results of this thesis may reveal that the shifts in attention observed in Meier and Robinson (2004) were artifactual. Most of these changes would not have been expected to eliminate the activation of the GOOD IS UP conceptual metaphor and associated perceptual processes. However, the possibility that these changes are important boundary conditions for observing GOOD IS UP congruent shifts in attention should be explored in future studies.

First, the reversal of the order of the component tasks was changed. In Meier and Robinson (2004) participants evaluated the words immediately after their presentation, before the target was presented. The change in task component order was necessary to conduct a more stringent test of Meier and Robinson's (2005) prediction of automaticity; that shifts in attention should be seen at automatic processing stages if representation is achieved with conceptual metaphors. Meier and Robinson's (2004) design did not allow a robust test of the automaticity prediction. To allow controlled manipulation of SOA, in the current studies participants evaluated the words at the end of the trial, after a response was made to the visual target. Therefore it is possible that participants were not immediately evaluating the word but were delaying meaning access until after presentation of the target. If participants were delaying the evaluation process attentional shifts would not be observed.

However, there is evidence that participants were evaluating the word when it was presented. Evaluation agreement was fairly high so comprehension must have taken place at some point in the trial. Furthermore, in studies 1, 3, and 4, participants were faster to respond to visual targets at the long SOA than at the short SOA.

Although the facilitation of response times for the shape task at the long SOA could be due to any number of reasons, it is consistent with a PRP explanation (a dual task bottleneck at response selection; Pashler, 1992, 1993), suggesting that participants had to select their evaluation response for stimulus one, the auditory cue, before selecting their response for stimulus two, the visual target. Presumably the same delaying strategy would have been present in Study 1. In Study 1 HIGH PITCH IS UP congruent attentional shifts were observed and a PRP effect was also present. Thus if the GOOD IS UP conceptual metaphor was recruited during spoken word processing, the timing of evaluation is probably not a critical boundary condition. To completely rule out evaluation timing as a boundary condition, a study should be conducted in which participants evaluate the spoken word immediately after it is presented. If the immediate evaluation of the spoken word is necessary to observe GOOD IS UP congruent shits in attention, GOOD IS UP shifts in attention should be observed with this design.

Second, it is also worth considering the role that explicitly producing the words ‘positive' and ‘negative’ may have on activation of the GOOD IS UP conceptual metaphor. Meier and Robinson’s (2004) paradigm may have exaggerated the role that the GOOD IS UP conceptual metaphor plays in written word processing. In Meier and Robinson's (2004) study participants were required to say out loud their evaluation of the word. In the paradigm used in this thesis participants were required to click on the appropriate label; positive, negative, neutral in Study 3, and happy, sad, neutral in Study 4. Actively saying the word would have activated the motor program for positive and negative semantics which may on its own have resulted in feedback activation to the mapping between the source domain (verticality) and the target domain (valence), resulting in GOOD IS UP congruent shifts of attention. In fact the labels used by Meier and Robinson, positive and negative, are the target dimensions mapped to upper and lower space in the GOOD IS UP conceptual metaphor. Therefore the shifts in attention described by Meier and Robinson may be a result of the explicit vocal response, not the evaluation. However, Brookshire et al. (2010) did observe GOOD IS UP congruent shifts of attention with no explicit evaluation of the words, aloud or otherwise. Thus, if activation of motor programs does result in feedback activation to the conceptual metaphor system, this was probably not producing the totality of the effect.

The paradigm used in this thesis could be modified to test the contribution of explicit labelling using the positive and negative dimensions. Participants could complete the evaluation task using explicit naming instead of mouse clicks. Given the above comments on the timing of the evaluation component, two versions should be conducted, one where the spoken evaluation occurs immediately after presentation of the word and one where the spoken evaluation occurs at the end of each trial. If explicit production of the valence labels produces GOOD IS UP congruent shifts in attention, then shifts in attention should be observed with the use of a naming rather than mouse click response. In fact, if the change to a spoken response restores GOOD IS UP shifts in attention then a study should be conducted in which the task on each trial is to say aloud “positive” or “negative” before completing a visual attention component. This would allow pure assessment of the contribution of explicit valence label production to GOOD IS UP shifts in attention, without the confounding influence of an evaluation task.

The most likely methodological reason I did not observe GOOD IS UP congruent shifts in attention in this thesis is the addition of neutral semantics and prosody to the task. Theoretically, the addition of neutral may have changed the experimental context (compared to Meier & Robinson, 2004) and thus this thesis may demonstrate the dependency of grounded cognition on context.

Contextually dependent grounded cognition.

In this thesis the words used were only emotional in a maximum of one channel. The words could be semantically emotional or prosodically emotional, but were never both semantically and prosodically emotional. Neutral semantics and prosody were included for two reasons. First, I wished to look at the recruitment of the GOOD IS UP conceptual metaphor during processing of emotional semantics and emotional prosody separately. This entailed that one channel, either semantics or prosody, was neutral in each study. Second, naturalistic speech is not exclusively emotional. Generally, in emotional language research, words are selected which are very high in valence and arousal. It was thought that the inclusion of neutral stimuli would increase the ecological validity of the test of the cognitive reality of the GOOD IS UP conceptual metaphor in spoken language. Thus, a condition was included in Studies 3 and 4 in which semantically-neutral words were spoken in neutral prosody.

However, by including neutral semantics and prosody, the contrast between the two emotional valences used in each study was reduced, which could have affected the degree to which conceptual metaphoric representations were activated. In the written GOOD IS UP conceptual metaphor studies with only positive and negative words (e.g. Brookshire et al., 2010; Meier & Robinson, 2004, 2006) the valence of a word would have been very salient to the participant, and the evaluation required, positive or negative, would have been relatively easy. However, in the current studies it may not have been as clear to the participants which semantic valence a word carried, and to a greater degree in which prosodic valence a word was spoken. Therefore, the evaluation task would have been harder than in previous studies. Indeed there is evidence, in the modulation of the PRP effect in the prosody study, that participants found evaluation of some prosodies easier than others. In Study 4 participants appeared to find it easier, as assessed with reaction time for the visual target (which can give an indication of ease to select an evaluation response), to select their evaluation response as happy rather than as sad or neutral. Perhaps the difficult evaluation task forced recruitment of a strategy of deeper processing, and other non-metaphoric knowledge was used to make the evaluation judgement. If the inclusion of neutral stimuli does lessen the contribution of metaphoric mapping, the strong view of Conceptual Metaphor Theory outlined in the introduction is not supported. In fact, according to grounded cognition theory, a strategy of deeper processing (as when evaluation is hard) should have resulted in greater (not lesser) recruitment of conceptual metaphoric representations (Barsalou et al., 2008; Brookshire et al., 2010).

The measure in this thesis of activation of conceptual metaphors was shifts in attention. No GOOD IS UP shifts in attention were observed, thus there was no evidence of activation of the GOOD IS UP conceptual metaphor. Yet participants in the current studies could clearly complete the evaluation task which suggests that conceptual metaphorical mapping may not be all there is to representation of emotion concepts. In order to evaluate the word another kind of representation must have been activated. Proponents of grounded cognition theory, and Conceptual Metaphor Theory more specifically, are focused on the importance of grounding to the detriment of evidence to the contrary. For instance, as pointed out by Louwerse and Jeuniaux (2008), experiments that provide evidence *for* grounded cognition representations do not provide evidence *against* abstract, symbolic representation. Analogous to the black swan problem (Popper, 1959), if researchers only test for the presence of conceptual metaphor representations, they will not find evidence for non-grounded representation. There is also the problem of comparison. Grounded cognition theorists (for example Johnson, 2007 and Lakoff & Johnson, 1999, in the conceptual metaphor literature) criticise “traditional theories”. All non-grounded theories cannot be lumped together (Murphy, 1996) and indeed aspects of “traditional theories” may be valid descriptions of representation. A less extreme view of Conceptual Metaphor Theory in conjunction with aspects of “traditional theories” may be able to incorporate the findings of this thesis more parsimoniously than either perspective alone. The neutral context in which participants evaluated the emotional stimuli may have reduced activation in the grounded, conceptual metaphor representation system. Reduced activation of conceptual metaphor mappings would have resulted in the lack of GOOD IS UP shifts in attention. However, participants still evaluated the emotional linguistic stimuli correctly. Thus, there must be another representational system, which does not consist of metaphoric source-target domain mappings, access to which allowed participants to complete the evaluation task. This is the multiple systems view of grounded cognition.

Multiple systems.

In a similar vein to Dual-Coding Theory (Paivio, 1971; 1986), the less extreme, multiple systems, version of grounded cognition is that there are at least two systems of representation; one that recruits the perceptuomotor system and one that is separate from the perceptuomotor system (e.g. Barsalou et al. 2008). The non-grounded system stores abstract-linguistic information. Both the linguistic and grounded systems underlie representation of emotional concepts; however recruitment of the systems is contextually determined.

Grounded cognition evidence.

Evidence up to the current date for contextually-activated representation systems mainly comes from the general grounded cognition literature, not specifically from the conceptual metaphor literature. Recent thinking in the grounded cognition literature strongly advocates a role for context in simulation, and while embodiment definitely seems to be part of the representation of emotional concepts, grounding is not seen as automatic (Winkielman et al., 2008). There are an increasing number of recent studies in the general grounded cognition literature which point to the existence of more than one representation system; one grounded, one non-grounded.

An example from the emotion literature illustrates that facial muscle activation seems to be contextually determined. In a series of experiments, Niedenthal et al. (2009) examined facial muscle activation by emotional concrete and abstract concepts. Electromyographic (EMG) activity congruent with the emotion of the word was only observed when the context was appropriate. In one experiment all participants were required to complete a property generation task to emotional words. In a clever manipulation half of the participants were required to imagine they were generating features of the words for a close friend (the hot audience context condition), the other half were required to imagine they were generating features of the words for a supervisor with whom the participant had a formal work level association (the cold audience context condition). The hypothesis was that the first group would employ a simulation strategy while the second group would employ a lexical association strategy. EMG activity measured in the two groups was consistent with the hypothesis. While there was no difference in the amount or properties of the words participants generated in the two conditions (both groups completed the task to the same level of performance), facial EMG activity congruent with the emotion of the properties being generated was observed to a greater extent in the hot audience group, suggesting they were simulating emotional experience. The difference between groups indicated that context can modulate the processes used to access emotion-related representations. Implicit in such a conclusion is that there is more than one representational system, one grounded, one non-grounded.

All one system grounded cognition theories state that the grounded representation process is obligatory. Yet an increasing number of studies point to a non-obligatory role for the grounded conceptual system. See Havas et al. (2007) for a study illustrating that lexical processing level is an important boundary condition to define for observing embodied effects (but see van Dam, Rüschemeyer, Lindemann, & Bekkering, 2010 for a counter example). The grounded conceptual system seems not to be activated automatically. It seems to only be engaged when the context accentuates the perceptuomotor nature of the concept referred to by the word. See Shintel and Nusbaum (2008) for an example of contextual constraining of embodied effects in spoken language and see Raposo, Moss, Stamatakis, and Tyler (2009) and Rüschemeyer, Brass, and Friedericic (2007) for neuroimaging studies demonstrating contextual embodiment.

Multiple systems in Conceptual Metaphor Theory.

The multiple systems view of grounded cognition has been taken up more slowly by conceptual metaphor theorists than in other grounded cognition literatures. Meier and Robinson's (2005) predictions, derived from Lakoff and Johnson (1999) who are absolute in the view of conceptual metaphors in representation, test the strong version of Conceptual Metaphor Theory. Murphy (1996) states that the strong view of Conceptual Metaphor Theory is not an accurate description of our representation system. A weaker view, in which metaphorical mappings shape our representations but are not the totality of them, is suggested as an alternative by Murphy.

The weaker view removes serious problems with Conceptual Metaphor Theory, such as the problem of multiple metaphorical mappings for a target domain. Good is not only up, but also close and warm. Sometimes the source domains may be contradictory for a given target domain. For example, anger is negative so should activate the BAD IS DOWN metaphor. Anger is not always distant though. Some forms of anger may activate an ANGER IS CLOSE metaphor (see Harmon-Jones, 2003 for a discussion of whether anger is associated with approach or withdrawal motivation). A central component of Conceptual Metaphor Theory is that the metaphorical mappings are necessary and are obligatorily activated. If that is the case, then for concepts with conflicting metaphors, multiple metaphors should be problem for understanding spoken and written language, yet normally functioning people do not have confused representational systems. A weaker view of Conceptual Metaphor Theory, in which there is more to representation than metaphoric mappings, allows for flexible representations. Certain source-target domain mappings may play more of a role in some situations than in others with context determining which mappings are activated.

Different questions need to be asked and different predictions need to be tested to reveal a more realistic, non-absolute, view of the nature of conceptual representation. Fourteen years after Murphy (1996), Brookshire et al. (2010) also posit that we need to start establishing a different view of the role of conceptual metaphors in representation. Rather than test whether metaphorical mappings are a necessary component of representation, we should test under what conditions metaphorical mappings are activated and whether the mappings contribute to representation. Rather than test whether metaphorical mappings are recruited at automatic processing stages, we should test the limits of automaticity and explore the contexts in which stronger and weaker recruitment of metaphorical mappings are observed. Rather than all or nothing, automaticity in conceptual metaphor recruitment may be a continuum.

As an example of how researchers could conduct research with these revised aims, Brookshire et al. (2010) explored the effect that practice and context have on conceptual metaphor processing. Practice was operationalised as the difference in recruitment of the GOOD IS UP metaphor for the first and second presentations of emotional words. Participants were required to identify the colour, purple or green, in which positive and negative words were displayed. The valence of the words was not central to the task, and the task could be completed with participants ignoring the valence of the words. The verticality aspect of the task was in the positioning of the response keys for the colour task. If the word was in one colour participants released a centralised key and pressed a key positioned in the upper position, if the word was in the other colour participants pressed a key positioned in the lower position. The metaphorical mapping between the valence of the word and the button pressed could be congruent with respect to the GOOD IS UP conceptual metaphor, positive word-upper key/ negative word-lower key, or incongruent, positive word-lower key/ negative word-upper key.

For the first presentation of words, a congruency effect was observed. Participants were faster to identify the colour of words when the metaphorical mapping was congruent than when it was incongruent. That this GOOD IS UP congruency effect was observed even when the word meaning was not central to the task, suggested to Brookshire et al. (2010) that the GOOD IS UP conceptual metaphor is strongly activated at automatic processing stages. For the second presentation of words, no congruency effect was observed. Participants were not faster to identify the colour of words when the metaphorical mapping was incongruent.

In order to explore what could be contributing to the modulation of the congruency effect, Brookshire et al. (2010) conducted another experiment in which the nature of filler trials was manipulated. The experimental trials were identical to the first experiment but rather than the colour task performed on the experimental emotion words, the filler trials were presented in a white font which cued participants that they had to perform a semantic or visual judgement on these words. Half of the participants were presented with filler trials for which the task was to decide whether the word was an animate or inanimate object. The other half of the participants were presented with filler trials for which the task was to decide whether a red X was present in a grid of grey squares. Thus, the animacy task oriented participants to attend to the meaning of the stimuli, while the red X task oriented participants to attend to a perceptual feature of the stimuli. Participants were faster to identify the font colour on experimental trials when pressing the upper key after positive words and when pressing the lower key after negative words than vice versa. Importantly though, this GOOD IS UP congruency effect was only observed in the data of the participants whose attention was oriented to the meaning of the words.

Brookshire et al. (2010) have begun to test the relevance of a weaker view of Conceptual Metaphor Theory and to probe the boundary conditions under which metaphorical mappings play a role in conceptual representation. The Brookshire et al. study suggests that context, in the from of task demands, specifically the level to which words are processed, is an important boundary condition for observing metaphor congruent perceptual processes. Thus, conceptual metaphors such as GOOD IS UP may not be activated obligatorily. By asking more specific research questions than whether or not representation is grounded, research can be conducted which is more informative as to the nature of representation.

Explaining the current results.

I have presented examples of studies which suggest that there is more to representation than only a grounded system utilising the perceptuomotor systems of the brain. There must also be a non-grounded system, characterised by abstract, amodal, symbolic representations, which allows semantic tasks to be completed when the context is inappropriate for activating the grounded system. While not referencing Conceptual Metaphor Theory specifically, a recent theory, Linguistic and Situated Simulation (LASS; Barsalou et al., 2008) has been developed which includes two systems of representation in the mind. Barsalou et al. state that they are open to there being more than two systems. The point is that there is more than one system; at least one which is embodied, is grounded in the perceptuomotor systems (termed situated simulations), and at least one which relies on abstract, amodal linguistic statistical information. Both the linguistic system and the grounded system are activated when processing a linguistic form (Barsalou et al. focus on words). The two systems are not modular and they interact; activation in one system modulates activity in the other system. The time course of activation in the systems is a key part of LASS theory. Given past experimental evidence, Barsalou et al. (2008) assume that the situated simulation system is activated early and automatically, at least by 200ms. However, the linguistic system reaches peak activation before the situated simulation system. Furthermore, under certain circumstances the central executive can prolong the majority role of the linguistic system. See Barsalou et al. for a summary of evidence supporting the LASS theory.

The LASS theory is just one example of a multiple systems theory with at least one grounded system (see Dove, 2011; Louwerse & Connell, 2010; and Louwerse & Jeuniaux, 2008, 2010 for additional examples of multiple systems theories). LASS may not be an accurate, or a complete picture of how the mind achieves representation of emotional concepts. The point is that multiple systems theories, in conjunction with experimental evidence, provide a compelling argument for the cognitive reality of more than one representational system for processing emotion related stimuli (semantically or prosodically), at least one of which is non-grounded. Regardless of the specific details of the systems, the presence of two systems could also allow for the role of context to be elucidated in the activation of conceptual metaphors. There is no equivalent to the LASS theory in the conceptual metaphor literature. However, it is conceivable that a similar multiple systems architecture is valid for Conceptual Metaphor Theory. There could be both a grounded system, utilising metaphoric mappings between perceptuomotor source domains and conceptual target domains, and a linguistic system, utilising some type of non-grounded (statistical, abstract, amodal) knowledge, which underlie representation of emotion concepts.

In fact, the nature of the stimuli used in the studies reported in this thesis could have increased reliance on the non-grounded linguistic system. It was expected that the conceptual metaphor GOOD IS UP should only be recruited during evaluation of emotional words, either semantic or prosodic, and not during evaluation of words which were both semantically and prosodically neutral. However, the spoken emotional words in these studies were never both semantically and prosodically emotional. Analogous to studies which demonstrate that grounded cognition congruent effects were only observed when the context is appropriate, perhaps in this thesis the presence of at least either neutral semantics or prosody on each trial and the inclusion of completely (semantically and prosodically) neutral words set an inappropriate context and the conceptual metaphor GOOD IS UP was not activated or only minimally activated. Instead the non-grounded system may have been more dominant, sufficient to complete the evaluation task. If the conceptual-metaphor system was not activated of course no GOOD IS UP congruent shifts in attention would be observed.

It is possible to test the role of context in activating the GOOD IS UP conceptual metaphor with a small modification of the studies in this thesis. Participants would listen to semantically positive, negative, and neutral words spoken in congruent emotional prosodies. That is, semantically-positive words spoken in happy prosody, semantically-negative words spoken in sad prosody, and semantically-neutral words spoken in neutral prosody. Given that the contextual boundary conditions for spoken language are unknown, it would be prudent to conduct two versions of the proposed study, one with semantically-neutral words spoken in neutral prosody included and one with them excluded. If an appropriate context is necessary for the conceptual metaphor system to be reliably recruited, then GOOD IS UP metaphoric congruent shifts in attention should be observed when the spoken words are both semantically and prosodically emotional. Following positive words participants should be faster to respond to targets in the upper visual-field than in the lower visual-field, and following negative words participants should be faster to respond to targets in the lower visual-field than in the upper visual-field. In conjunction with this thesis' results, affirmative evidence of this predication would indicate 1) that at least two systems, one grounded utilising conceptual metaphoric mappings, and one non-grounded, underlie representation of emotion concepts in spoken language; and 2) that recruitment of the representational systems is contextually determined.

Spoken Language

If with future scrutiny it is determined that none of the other changes from the Meier and Robinson (2004) paradigm, including those described above (task component order, response modality, or context), are responsible for the lack of GOOD IS UP congruent shifts in attention in this thesis, the change in stimulus modality is the most likely the cause. The aim of this thesis was to add to theoretical understanding of Conceptual Metaphor Theory. To that end, the words in this thesis were presented in the spoken modality. In contrast, all of the previous studies which reported GOOD IS UP congruent perceptual effects with linguistic stimuli (Brookshire et al., 2010; Casasanto, 2008, as cited in Brookshire et al., 2010; Meier & Hauser, 2008, as cited in Crawford, 2009; Meier & Robinson, 2004, 2006; Weger et al., 2007) used written words. As laid out in the introduction investigation of the role of the GOOD IS UP conceptual metaphor in spoken word processing is interesting for evolutionary and complexity reasons.

Evolution.

Simply put, the rationale for expecting GOOD IS UP attentional shifts during processing of spoken emotional words was that:

a) according to all grounded cognition theorists, representation is achieved using the evolutionarily older perceptuomotor system.

b) Speech is evolutionarily older than writing.

c) As GOOD IS UP congruent shifts in attention (which are assumed to index access to the verticality-emotion mapping) have been observed during the evolutionarily younger reading process, shifts in attention should have definitely be observed during spoken language processing.

See the introduction for a full summary of the argument. Laying aside all the other possible reasons for the lack of GOOD IS UP congruent attentional shifts, that this evolution argument is not supported is concerning for Conceptual Metaphor Theory. Lakoff and Johnson (2009) in fact say that conceptual metaphoric representation is obligatory. In conjunction with the context and multiple systems literature, this thesis would suggest that conceptual metaphoric representation is not obligatory. This violates the third, automaticity, prediction of Meier and Robinson.

Complexity.

As described in the introduction, spoken language is more complex than written language; information can be conveyed in the semantic and prosodic channels, which do not have to be congruent. It is important to remember that emotion related conceptual metaphors will not be the only conceptual metaphors relevant to representation of a word during spoken language processing. Multiple metaphors (Murphy, 1996) are especially likely to be activated by spoken words in which two channels of information are salient. For example, the HIGH PITCH IS UP metaphor used in Study 1 as a test of the suitability of the paradigm to observe metaphor congruent shift in attention, may also have been activated when processing emotional prosody. The prosodies used in Study 4 were consistent in their verticality mappings for both emotional and pitch target domains. Happy prosody is both positive and higher in pitch, target domains that map on to upper space, and sad prosody is both negative and lower in pitch, target domains that map on to lower space. However, if the HIGH PITCH IS UP metaphor was activated when processing prosody in this thesis, the doubling up of the emotion and pitch mappings should have increased the likelihood of observing faster response times to targets in the upper-visual field than in the lower-visual field after happy prosody and of observing faster response times to targets in the lower visual-field than in the upper visual-field after sad prosody.

If conceptual metaphors are not activated, or only minimally activated, during spoken language processing, then how large a role could conceptual metaphoric representation have in even more complex linguistic processing, such as processing of sentences or discourse? If future research rules out all alternative explanations (aside from the change to spoken word stimuli) for the lack of GOOD IS UP shifts in attention during spoken word processing, Conceptual Metaphor Theory must be reconsidered as a theory of emotional concept representation.

Other Considerations

If conceptual metaphoric mappings are a cognitively real form of representation, the paradigm used in this thesis and the lack of GOOD IS UP congruent shifts in attention point to two other factors that should be considered: time course of conceptual metaphor activation and the relevance of dimensional versus categorical emotion.

Time course.

The time course of the recruitment of the two representational systems, one grounded and one-non grounded, will be a key component of any valid theory of representation (van Dam et al., 2010). Even the LASS theory has not yet elucidated the exact time course of the activation of the linguistic and situated simulation system. The SOAs used in this thesis may have allowed context to play a large role and precluded any possibility of observing automatic activation of the GOOD IS UP conceptual metaphor. According to Kotz and Paulmann's (2011) multistep theory of emotional speech processing, context comes into play at later stages of cognition, which Kotz and Paulmann define as from approximately 400ms from the onset of a word. The shortest SOA was 400ms in the paradigm used in this thesis. Thus, the neutral context could have constrained processing on all trials. Obligatory activation of the GOOD IS UP conceptual metaphor may have been missed by the paradigm. Future studies should use shorter SOAs to explore the automaticity of the GOOD IS UP conceptual metaphor; although word length will limit the shortest SOA that can be used with spoken words. For example, if using the same stimuli as in this thesis (with the shortest word duration of 313ms), GOOD IS UP shifts in attention would probably not be observed with a 100ms SOA. Participants need to hear enough of the word to activate emotional evaluation processes.

Using a shorter SOA (such as 300ms) may not reveal that the GOOD IS UP conceptual metaphor is activated automatically. Chiou and Rich (2011) conducted the first thorough exploration of the automaticity of the HIGH PITCH IS UP conceptual metaphor. The series of studies conducted by Chiou and Rich demonstrated first, that the HIGH PITCH IS UP conceptual metaphor is only activated after context determines whether a pitch is high or low (relative pitch). Second, HIGH PITCH IS UP congruent shifts in attention do not appear until at least 300ms after the onset of the pitch cue. Third, the mapping between pitch and vertical location is not automatic; it is susceptible to top down control. Participants with training demonstrated the opposite shifts in attention; high pitch- lower visual field, low pitch- upper visual field. Chiou and Rich concluded that the activation of the HIGH PITCH IS UP conceptual metaphor lies between exogenous and endogenous processing. If a robust conceptual metaphor like the HIGH PITCH IS UP metaphor is not activated automatically, then it is unlikely that a less robust metaphor such as the GOOD IS UP metaphor will be activated automatically.

The Chiou and Rich (2011) studies suggest that the robust HIGH PITCH IS UP conceptual metaphor is not obligatorily activated but a similar thorough exploration of the time course of activation of the GOOD IS UP conceptual metaphor is needed before a similar conclusion can be made. The current paradigm could be useful for determining the time course of activation of systems of representation. An advantage of the paradigm used in this thesis, compared to Meier and Robinson (2004) is that SOA between the word and visual target is easily manipulated. The current paradigm would be suitable with the addition of more SOAs (and consequently more trials). A range of SOAs from 200ms (see Hauk & Pulvermüeller, 2004) to 1200ms would cover the range of automatic and controlled processing stages.

Dimensional versus categorical emotion.

Researchers should also consider that there may not be a correspondence between all positive valenced emotions and upper space and all negative valenced emotions and lower space. All past psychological research into the cognitive reality of the GOOD IS UP metaphor treats emotion as two dimensional: positive and negative (Brookshire et al., 2010; Casasanto, 2008, as cited in Brookshire et al., 2010; Meier & Hauser, 2008, as cited in Crawford, 2009; Meier & Robinson, 2004, 2006; Weger et al., 2007). This thesis treats emotion as both dimensional and categorical. In Study 3 words of positive and negative semantic valence (dimensional emotion) were presented to participants, whereas in Study 4 happy and sad prosodies (categorical emotion) were presented to participants. The assumption is made that all positive emotions map onto upper space and all negative emotions onto lower space. However it may be that some categorical emotional valences map more robustly onto upper and lower space than others.

The consideration of variation in source domain recruitment by categorical valences is especially prudent for negative emotions. There are many more categories of negative emotions, including sad, fear, disgust, and anger, than positive ones. Happy is usually the one categorical emotion included as a positive emotion in psychological experiments. Linguistic evidence (i.e. use of corpus and dictionary data) suggests that the GOOD IS UP conceptual metaphor may be most relevant for happy and sad valences (Köveces, 2000) than for other valences such as anger and fear. Kövecses, who considers the mappings between positive valence and upper space, and between negative valence and lower space separately, lists fifteen conceptual metaphors underlying the representation of happiness including HAPPY IS UP, HAPPY IS LIGHT, HAPPY IS WARM, HAPPINESS IS HEALTH, and HAPPINESS IS FLUID IN A CONTAINER and fourteen for sadness including SAD IS DOWN, SAD IS DARK, SADNESS IS A LACK OF HEAT, SADNESS IS AN ILLNESS. There are more variations on the HAPPY IS UP conceptual metaphor than for the other conceptual metaphors of emotion (i.e. HAPPY IS LIGHT). There are also more variations on the HAPPY IS UP conceptual metaphor than for the converse conceptual metaphor SAD IS DOWN, for example HAPPINESS IS BEING OFF THE GROUND and HAPPINESS IS BEING IN HEAVEN do not have a complementary SAD version. Which categorical emotions the GOOD IS UP metaphor applies to is an important boundary condition to define in both written and spoken word processing.

Conclusions

This thesis tested for evidence of activation of the GOOD IS UP conceptual metaphor in processing of spoken emotional words. The aim was to learn more about the nature of conceptual representations activated during processing of spoken language, and emotional semantics and emotional prosody were considered separately. If evaluation of spoken emotional words activated metaphorical representation, then GOOD IS UP consistent shifts in attention should have been observed in response times to targets in the upper and lower visual field. No shifts in attention were observed when participants evaluated semantically-emotional words in Study 3, or when participants evaluated prosodically-emotional words in Study 4. A multiple systems view of Conceptual Metaphor Theory in which there are at least two contextually activated systems of representation, one involving grounded source-target domain metaphorical mappings, and one involving non-grounded linguistic information, may explain the lack of attentional shifts observed. Future research should explore the boundary conditions on automaticity and recruitment of conceptual metaphorical mappings.

The majority of experiments conducted with the aim of understanding representation seem to rely on written stimuli. This thesis demonstrates that it is important not to construct theories of conceptual representation only on the basis of evidence from written language processing. Given that speech is evolutionarily older than writing and that grounded cognition theorists claim that representation is achieved using the evolutionarily older perceptuomotor system (e.g. Lakoff & Johnson, 1999) it was expected that stronger evidence for Conceptual Metaphor Theory would be observed with spoken words. This was not the case. Spoken language is also more complex. Information can be conveyed through the semantic and prosodic channels. Thus, by overly relying on written stimuli, we may have created a biased or unrealistic view of the nature of representation and even cognition in general (Wurm et al., 2001, 2004). Any valid theory of representation must be able to account for findings in both written and spoken language processing.

Recent research in Conceptual Metaphor Theory (Brookshire et al., 2010) is driven by more complex questions than “is representation embodied or not?”. By asking fine grained questions we may gain more information about the nature of representation. This strategy of refining boundary conditions needs to be extended further into Conceptual Metaphor Theory. The previous studies using written emotional words may have serendipitously selected the necessary boundary conditions for GOOD IS UP congruent perceptual processing to be observed. This thesis has taken the first step in exploring metaphoric representation during processing of spoken emotional words. From the four studies conducted in this thesis I can only conclude that representation of emotion-related concepts is not solely achieved with conceptual metaphor mappings. Exploration of the boundary conditions under which conceptual metaphors play a role in emotional language processing in written versus spoken language will shed more light on the role of the GOOD IS UP conceptual metaphor in representation of emotional concepts.

References

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

Banse, R. & Scherer, K.R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*, 614-636. doi: 10.1037/0022-3514.70.3.614

Barnard, P.J., Duke, D.J., Byrne, R.W., & Davidson, I. (2007). Differentiation in cognitive and emotional meanings: An evolutionary analysis. *Cognition and Emotion, 21,* 1155-1183. doi: 10.1080/02699930701437477

Barsalou, L.W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22,* 577-660. doi: 10.1017/S0140525X99002149

Barsalou, L.W. (2008). Grounded cognition. *The Annual Review of Psychology, 59*, 617-645. doi: 10.1146/annurev.psych.59.103006.093639

Barsalou, L.W. (2010). Grounded cognition: Past, present, and future. *Topics in Cognitive Science, 2,* 716-724. doi: 10.1111/j.1756-8765.2010.01115.x

Barsalou, L.W., Santos, A., Simmons, W.K., & Wilson, C.D. (2008). Language and simulation in conceptual processing. In M. De Vega, A.M. Glenberg, & A.C. Graesser (Eds.). *Symbols, Embodiment, and Meaning.* Oxford: Oxford University Press.

Ben-Artzi, E. & Marks, L.E. (1999). Processing linguistic and perceptual dimensions of speech: Interactions in speeded classification. *Journal of Experimental Psychology, 25,* 579-585. doi: 10.1037/0096-1523.25.3.579

Bernstein, I.H. & Edelstein, B.A. (1971). Effects of some variations in auditory input upon visual choice reaction time. *Journal of Experimental Psychology, 87,* 241-247. doi: 10.1037/h0030524

Boersma, P., & Weenink, D. (2007). Praat: Doing phonetics by computer (Version 4.6.06).

Bradley, M.M. & Lang, P.J. (1999). *Affective Norms for English words (ANEW): Stimuli, Instruction Manual and Affective ratings.* Technical report C-1, Gainesville, Florida. The Center for Research in Psychophysiology, University of Florida.

Brookshire, G., Ivry, R., & Casasanto, D. (2010). Modulation of motor-meaning congruity effects for valenced words. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 1940-1945). Austin, TX: Cognitive Science Society.

Casasanto, D. (2009). Embodiment of abstract concepts: Good and bad in right- and left-handers. *Journal of Experimental Psychology: General, 138,* 351-367. doi: 10.1037/a0015854

Chiou, R., & Rich, A.N. (2011). *When cross-modal metaphor meets attention: Auditory pitch orients visual attention to high and low location*. Paper presented at the 38th meeting of the Australian Society for Experimental Psychology, Auckland, New Zealand.

Collins, A.M. & Quillian, M.R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 240-247. doi: 10.1016/S0022-5371(69)80069-1

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33A*, 497-505.

Cook, N.D. (2002). *Tone of Voice and Mind: The Connections between Intonation, Emotion, Cognition, and Consciousness.* Amsterdam, Philadelphia: J. Benjamins Publishing.

Crawford, L.E. (2009). Conceptual metaphors of affect. *Emotion Review, 1,* 129-139. doi: 10.1177/1754073908100438

Crawford, L.E., Margolies, S.M., Drake, J.T., & Murphy, M.E. (2006). Affect biases memory of location: Evidence for spatial representation of affect. *Cognition and Emotion, 20,* 1153-1169. doi: 10.1080/0269993050034779

Dove, G. (2011). On the need for embodied and dis-embodied cognition. *Frontiers in Psychology, 1,* 1-13. doi: 10.3389/fpsyg.2010.00242

Duncan, S. & Feldman-Barrett, L. (2007). Affect is a form of cognition: A neurobiological analysis. *Cognition and Emotion, 21,* 1184-1211. doi: 10.1080/02699930701437931

Elias, L. J., Bryden, M. P., & Bulman-Fleming, M. B. (1998). Footedness is a better predictor than is handedness of emotional lateralisation. *Neuropsychologia, 36*, 37-43. doi:10.1016/S0028-3932(97)00107-3

Evans, K.K. & Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *Journal of Vision, 10,* 1-12. doi: 10.1167/10.1.6

Fodor, J.A. (1985). Prècis of the modularity of mind. *Behavioral and Brain Science, 8,* 1-42. doi: 10.1017/S0140525X0001921X

Gallese, V. & Lakoff, G. (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge. *Cognitive Neuropsychology, 22,* 455-479. doi: 10.1080/02643290442000310

Gibbs, R.W. (2006). Metaphor interpretation as embodied simulation. *Mind and Language, 21,* 434-458. doi: 10.1111/j.1468-0017.2006.00285.x

Gibbs, R.W. & Matlock, T. (2008). Metaphor, imagination, and simulation: Psycholinguistic evidence. *The Cambridge Handbook of Metaphor and Thought.* Leiden: Cambridge University Press.

Gray, J.R., Braver, T.S., & Raichle, M.E. (2002). Integration of emotion and cognition in the lateral prefrontal cortex. *Proceedings of the National Academy of Science, USA, 19,* 4115-4120. doi: 10.1073/pnas.062381899

Grush, R. (2003). In defense of some 'Cartesian' assumptions concerning the brain and its operation. *Biology and Philosophy, 18,* 53-93.

Harley, T. (2001). *The Psychology of Language: From Data to Theory.* New York: Taylor and Francis.

Harmon-Jones, E. (2003). Anger and the behavioral approach system. *Personality and Individual Differences, 35,* 995-1005. doi: 10.1016/S0191-8869(02)00313-6

Haselarger, P., de Groot, A., & van Rappard, H. (2003). Representationalism vs. anti-representationalism: A debate for the sake of appearance. *Philosophical Psychology, 16,* 5-23. doi: 10.1080/0951508032000067761

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Nueron, 41,* 301-307. doi: 10.1016/S0896-6273(03)00838-9

Hauk, O. & Pulvermüller (2004). Neurophysiological distinction of action words in the fronto-central cortex. *Human Brain Mapping, 21,* 191-201. doi: 10.1002/hbm.10157

Havas, D.A., Glenberg, A.M., & Rinck, M. (2007). Emotion simulation during language comprehension. *Psychonomic Bulletin & Review, 14,* 436-441.

Havas, D.A., Glenberg, A.M., Gutowski, K.A., Lucarelli, M.J., & Davidson, R.J. (2010). Cosmetic use of Botulinum Toxin-A affects processing of emotional language. *Psychological Science, 21,* 895-900. doi: 10.1177/0956797610374742

Johnson, M. (2007). *The Meaning of the Body: Aesthetics of Human Understanding.* Chicago: The University of Chicago Press.

Katz, J.J. & Fodor, J.A. (1983). The structure of a semantic theory. *Language, 39,* 170-210.

Kotz, S.A. & Paulmann, S. (2011). Emotion, language, and the brain. *Language and Linguistic Compass, 5,* 108-125. doi: 10.1111/j.1749-818X.2010.00267.x

Kövecses, Z. (2000). *Metaphor and Emotion: Language, Culture, and Body in Human Feeling.* New York: Cambridge University Press.

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

Ladd, D.R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Lakoff, G. & Johnson, M. (1980) *Metaphors We Live By*. Chicago: University of Chicago Press.

Lakoff, G. & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought.* New York: Basic Books.

Lazarus, R.S. (1984). On the primacy of cognition. *American Psychologist, 39,* 124-129. doi: 10.1037/0003-066X.39.2.124

Leventhal, H. & Scherer, K. (1987). The relationship of emotion to cognition: A functional approach to a semantic controversy. *Cognition and Emotion, 1,* 3-28. doi: 10.1080/02699938708408361

Levy-Drori, S. & Henik, A. (2006). Concreteness and context availability in lexical decision tasks. *The American Journal of Psychology, 119,* 45-65.

Louwerse, M. & Connell, L. (2010). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science, 35,* 381-398. doi: 10.1111/j.1551-6709.2010.01157.x

Louwerse, M. & Jeuniaus, P. (2008). Language comprehension is both embodied and symbolic. In M. De Vega, A.M. Glenberg, & A.C. Graesser (Eds.). *Symbols, Embodiment, and Meaning.* Oxford: Oxford University Press.

Louwerse, M.M. & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition, 114,* 96-104. doi: 10.1016/j.cognition.2009.09.002

Luodonpää-Manni, M. & Viimaranta, J. (2010). Metaphoric expressions on vertical axis revisited: An empirical study of Russian and French material. *Metaphor and Symbol, 25,* 74-92. doi: 10.1080/10926481003715994

Maeda, F., Kanai, R., & Shimojo (2004). Changing pitch induced visual motion illusion. *Current Biology, 14,* 990-991. doi: 10.1016/j.cub.2004.11.018

Meier, B.P. (2005). *Using metaphor to promote happiness: Will directing people to attend up make them "feel up"?* PhD dissertation, Department of Psychology, North Dakota State University of Agriculture and Applied Science, Fargo, North Dakota.

Meier, B.P. & Robinson, M.D. (2004). Why the sunny side is up: Associations between affect and vertical position. *Psychological Science, 15,* 243-247. doi: 10.1111/j.0956-7976.2004.00659.x

Meier, B.P. & Robinson, M.D. (2005). The metaphorical representation of affect. *Metaphor and Symbol, 20,* 239-257. doi: 10.1207/s15327868ms2004_1

Meier, B.P. & Robinson, M.D. (2006). Does "feeling down" mean seeing down? Depressive symptoms and vertical selective attention. *Journal of Research in Personality, 40,* 451-461. doi: 10.1016/j.jrp.2005.03.001

Meier, B.P., Robinson, M.D., & Clore, G.L. (2004). Why good guys wear white: Automatic inferences about stimulus valence based on brightness. *Psychological Science, 15,* 82-87. doi: 10.1111/j.0963-7214.2004.01502002.x

Meier, B.P., Robinson, M.D., Crawford, L.E., & Ahlvers, W.J. (2007). When "light" and "dark" thoughts become light and dark responses: Affect biases brightness judgements. *Emotion,* 7, 366-376. doi: 10.1037/1528-3542.7.2.366

Melara, R.D. & O'Brien, T.P. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General, 116,* 323-336. doi: 10.1037/0096-3445.116.4.323

Michalak, J., Troje, N.F., Fischer, J., Vollmar, P., Heidenreich, T., & Shulte, D. (2009). Embodiment of sadness and depression-Gait patterns associated with dysphoric mood. *Psychosomatic Medicine, 71,* 580-587. doi: 10.1097/PSY.0b013e3181a2515c

Murphy, G.L. (1996). On metaphoric representation. *Cognition, 60,* 173-204. doi: 10.1016/0010-0277(96)00711-1

Navarrete, E., Basagni, B., Alario, F.X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production. *The Quarterly Journal of Experimental Psychology, 59,* 1681-1690. doi: 10.1080/17470210600750558

Niedenthal, P.M. (2007). Embodying emotion. *Science, 316*, 1002-1005. doi: 10.1126/science.1136930

Niedenthal, P.M. (2008). Emotion concepts. In M. Lewis, J.M. Haviland-Jones, & L.F. Barrett (Eds.). *The Handbook of Emotion* (pp. 587-600). New York: Guilford.

Niedenthal, P.M., Krauth-Gruber, S., & Ric, F. (2004). A role for emotion in lexical access. In P. Bonin (Ed.) *Mental Lexicon: Some words to talk about words* (pp. 147-166). New York: Nova Science Publishers.

Niedenthal, P.M., Winkielman, P., Mondillon, L., Vermeulen, N. (2009). Embodiment of emotion concepts. *Journal of Personality and Social Psychology, 96,* 1120-1136. doi: 10.1037/a0015574

Nygaard, L.C. & Lunders, E.R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory and Cognition, 30,* 583-593. doi: 10.3758/BF03194959

Nygaard, L.C. & Queen, J.S. (2008). Communicating emotion: Linking affective prosody and word meaning. *Journal of Experimental Psychology, 34,* 1017-1030. doi: 10.1037/0096-1523.34.4.1017

Occelli, V., Spence, C., & Zampini, M. (2009). Compatibility effects between sound frequency and tactile elevation. *NeuroReport, 20,* 793-797. doi: 10.1097/WNR.0b013e32832b8069

Pashler, H. (1992). Attentional limitations in doing two tasks at the same time. *Current Directions in Psychological Science, 1,* 44-48. doi: 10.1111/1467-8721.ep11509734

Pashler, H. (1993). Doing two things at the same time. *American Scientist, 81,* 48-55.

Paivio, A. (1971). *Imagery and Verbal Processes*. New York: Holt, Rinehart & Winston.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.

Popper, K.R. (1959). *The Logic of Scientific Discovery,* English translation of *Logik der Forschung*. London: Hutchinson.

Posner, M.I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32,* 3-25. doi: 10.1080/00335558008248231

Posner, M.I. & Snyder, C.R.R. (2004). Attention and cognitive control. In D.A. Balota & E.J. Marsh (Eds.), *Key Readings in Cognitive Psychology* (pp. 205-223). New York: Taylor & Francis Books, Inc.

Raposo, A., Moss, H.E., Stamatakis, E.A., & Tyler, L.K. (2009). Modulation of motor and premotor cortices by actions, action words and action sentences. *Neuropsychologia, 47,* 388-396. doi: 10.1016/j.neuropsychologia.2008.09.017

Robinson, M.D., Zabelina, D.L., Ode, S., & Moeller, S.K. (2008). The vertical nature of dominance-submission: Individual differences in vertical attention. *Journal of Research in Personality, 42,* 933-948. doi: 10.1016/j.jrp.2007.12.002

Rüeschemeyer, S.A., Brass, M., & Friederici, A.D. (2007). Comprehending prehending: Neural correlates of processing verbs with motor stems. *Journal of Cognitive Neuroscience, 19,* 855-865. doi: 10.1162/jocn.2007.19.5.855

Sadaghiani, S., Maier, J.X., & Noppeney, U. (2009). Natural, metaphoric, and linguistic auditory direction signals have distinct influences on visual motion processing. *The Journal of Neuroscience, 29,* 6490-6499. doi: 10.1523/JNEUROSCI.5437-08.2009

Samson, D. & Pillon, A. (2004). Orthographic neighbourhood and concreteness effects in the lexical decision task. *Brain and Language, 91,* 252-264. doi: 10.1016/j.bandl.2004.03.001

Scherer, K.R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40,* 227-256. doi: 10.1016/S0167-6393(02)00084-5

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime Suite (Version 1.1). Pittsburgh: Psychology Software Tools Inc.

Shintel, H., Nusbaum, H.C., & Okrent, A. (2006). Analog acoustic expression in speech communication. *Journal of Memory and Language, 55,* 167-177. doi: 10.1016/j.jml.2006.03.002

Shintel, H. & Nusbaum, H.C. (2008). Moving to the speed of sound: Context modulation of the effect of acoustic properties of speech. *Cognitive Science, 32,* 1063-1074. doi: 10.1080/03640210801897831

Steckler, L. & Young, C. (2009). Depression and body psychotherapy. *International Journal of Psychotherapy, 13,* 32-41.

Tolaas, J. (1991). Notes of the origin of some spatialization metaphors. *Metaphor and Symbolic Activity, 6,* 203-218. doi: 10.1207/s15327868ms0603_4

van Dam, W.O., Rüeschemeyer, S.A., Lindemann, O., & Bekkering, H. (2010). Context effects in embodied lexical-semantic processing. *Frontiers in Cognition, 1,* 1-6. doi: 10.3389/fpsyg.2010.00150

Wagner, S., Winner, E., Cicchetti, D., & Gardner, H. (1981). “Metaphorical” mapping in human infants. *Child Development, 52,* 728-731.

Weeks, D.J. & Proctor, R.W. (1990). Salient-features coding in the translation between orthogonal stimulus and response dimensions. *Journal of Experimental Psychology: General, 119,* 355-366. doi: 10.1037/0096-3445.119.4.355

Weger, U.W., Meier, B.P., Robinson, M.D., & Inhoff, A.W. (2007). Things are sounding up: Affective influences on auditory tone perception. *Psychonomic Bulletin & Review, 14,* 517-521. doi: 10.3758/BF03194100

Whaley, C.P. (1978). Word-nonword classification time. *Journal of Verbal Learning and Verbal Behavior, 17,* 143-154. doi: 10.1016/S0022-5371(78)90110-X

Willems, R.M., Labruna, L., D’Esposito, M., Ivry, R., & Casasanto, D. (2011). A functional role for the motor system in language understanding: Evidence from theta burst TMS. *Psychological Science.* In press.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9,* 625-636.

Wilson, M.D. (1988). The MRC Psycholinguistic Database: Machine Readable Dictionary, Version 2. *Behavioural Research Methods, Instruments and Computers*, *20*, 6-11.

Winkielman, P., Niedenthal, P.M., & Oberman, L. (2008). The embodied emotional mind. In G.R. Semin & E.R. Smith (Eds.) *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches* (pp. 263-288). New York: Cambridge University Press.

Wurm, L.H., Vakoch, D.A., & Seaman, S.R. (2004). Recognition of spoken words: Semantic effects in lexical access. *Language and Speech, 47,* 175-204. doi: 10.1177/00238309040470020401

Wurm, L.H., Vakoch, D.A., Strasser, M.R., Calin-Jageman, R., & Ross, S.E. (2001). Speech perception and vocal expression of emotion. *Cognition and Emotion, 15,* 831-852. doi: 10.1080/02699930143000086

Zajonc, R.B. (1984). On the primacy of affect. *American Psychologist, 39,* 117-123. doi: 10.1037//0003-066X .39.2.117

Zbikowski, L.M. (1998). Metaphor and music theory: Reflections from cognitive science. *Music Theory Online, 4.* Retrieved from http://www.mtosmt.org/issues/mto.98.4.1/mto.98.4.1.zbikowski_frames.html

Zung, W.W. (1965). A self-rating depression scale. *Archives of General Psychiatry, 12,* 63-70.

Zung, W.W. (1971). A rating instrument for anxiety disorders. *Psychosomatics, 12,* 371-379.

Appendix A: Target-present Word Lists

Target-present word lists used in Studies 3 and 4

	Positive-semantics List	Negative-semantics list	Neutral List	Happy-prosody List	Sad-prosody List
1	baby	afraid	doll	avenue	basket
2	beautiful	assault	appliance	barrel	bowl
3	brave	burial	autumn	book	butter
4	cake	cancer	cabinet	building	chair
5	comfort	crisis	chance	chin	clothing
6	diamond	dead	city	clock	column
7	elegant	devil	coast	coin	cork
8	excitement	disaster	context	cord	custom
9	fantasy	failure	cottage	elbow	dress
10	gentle	gloom	fish	fabric	gender
11	gift	grief	foot	fork	green
12	heart	hardship	fur	highway	hand
13	holiday	hate	glacier	hotel	history
14	hope	hurt	hairpin	inhabitant	jelly
15	joke	infection	hat	item	key
16	kiss	injury	industry	kettle	lantern
17	love	insult	journal	locker	machine
18	luxury	jail	lawn	market	milk
19	miracle	misery	metal	method	name
20	music	morgue	news	month	office
21	passion	neglect	opinion	paper	part
22	peace	poverty	paint	patent	patient
23	pillow	rape	pencil	phase	poster
24	rescue	sick	plant	quart	quiet
25	respect	slave	rattle	salute	revolver
26	reward	thief	salad	sphere	scissors
27	sleep	torture	seat	teacher	ship
28	truth	tragedy	statue	tower	street
29	warmth	ulcer	table	trunk	tennis
30	wedding	venom	theory	utensil	truck
31	win	victim	unit	violin	trust
32	wish	war	whistle	window	watch