Complement set reference after implicitly small quantities:

An event-related potentials study

Joanne Ingram[1#] & Heather J. Ferguson
University of Bedfordshire University of Kent

joanne.ingram@uws.ac.uk H.Ferguson@kent.ac.uk

+44 (0)141 848 3894 +44 (0)1227 827120

[1]corresponding author

[#]Joanne Ingram is now based at the University of the West of Scotland.

This work was completed with the support of funding from the University of Bedfordshire's Research Centre for Applied Psychology.

Complement set reference after implicitly small quantities: An event-related potentials study

Abstract

An anaphoric reference to the complement-set is a reference to the set which does not fulfil the predicate of the preceding sentence. Preferred reference to the complement-set has been found in eye movements when a character's implicit desire for a high amount has been denied using a negative emotion. We recorded ERPs to examine if, when a character's desire is denied with a negative emotion, the complement-set is immediately available for reference. Analysis of the N400 over posterior regions showed that while readers favoured the reference-set following a positive emotion, there was no difference in responses between complement-set and reference-set references following a negative emotion. Processing of a complement-set reference did lead to an overall increase in negativity of the N400, suggesting that interpreting a complement-set reference incurred a general processing cost. This study provides novel data on the range of circumstances under which the complement-set is available.

Keywords: complement set, pronoun resolution, negation, N400, emotion words

When processing a plural pronominal reference such as, *They played in the garden*, readers must seek a suitable set to satisfy the ambiguous anaphor, *they*. In many cases the set in question is easily identified by the preceding text. For example, when the above sentence follows, *The children were happy to be outside*, it is clear that *they* refers to the set of children who were happy to be outside. However, this referential assignment is much harder when the preceding anaphoric reference refers to a quantified antecedent (e.g. *Some of the children*…), which increases the possible sets available for resolution of the anaphor (i.e. The children who were happy and the children who weren't). In the current paper we report an event-related potential (ERP) study that examines the processing of this form of pronominal reference in cases where more than one antecedent has been implied by the preceding context.

Previous studies (Moxey & Sanford, 1986; Filik, Leuthold, Moxey & Sanford, 2011; Paterson, Sanford, Moxey & Dawydiak, 1998; Sanford, Moxey & Paterson, 1996) investigating the effect of quantification on pronominal reference have identified a pattern of processing whereby the preferred set for reference changes depending on the polarity of the quantifier. This pattern is demonstrated in sentences (1) and (2).

(1a) Many of the fans went to the match.

(1b) They cheered with vigour when the team scored.

(2a) Not many of the fans went to the match

(2b) They watched on the television at home.

Studies of sentence production have revealed that after a sentence containing a positive quantifier (1a) participants are most likely to provide continuations, beginning with the plural anaphor *They*, which describe the fans who attended the match (as in (1b)). This set is known as the reference set (or refset), as it focuses on the set of people or things referenced by the

material following the quantifier. In contrast, when the same sentence is negatively quantified (as in 2a) continuations are more commonly made to the set of fans who did not attend the match (2b). This set is known as the complement set (or compset), as it focuses on the complement of the referenced set (Moxey & Sanford, 1986; Sanford, Moxey & Paterson, 1996). This preference to produce refset or compset continuations depending on quantifier context has also been revealed in reading comprehension. Specifically, studies using eye-tracking methodology (Paterson, Sanford, Moxey & Dawydiak, 1998) showed that readers experienced pronounced processing difficulties when a positively quantified sentence was followed by an anaphoric reference to the compset. A corresponding difficulty was also reported when a negatively quantified statement was followed with a reference to the refset, suggesting that readers expected a reference to the compset. However, this processing difficulty was weaker and delayed relative to that observed for a refset reference following a negatively quantified statement. This pattern in eye movements could therefore be taken to indicate that the compset is not immediately available after a negatively quantified statement. However, a recent study that recorded ERPs during reading revealed N400 'anomaly detection' responses for both kinds of sentence mismatch, suggesting that readers do access a compset reference following a negative quantifier just as quickly as they access a refset reference following a positive quantifier (Filik, et al., 2011).

Moxey (2006), and Sanford, Dawydiak, and Moxey (2007) outlined a theory to explain the above pattern of preferred pronominal reference after a negative quantifier. The presupposition-denial account, also known as the supposition denial theory (Filik et al., 2011; Sanford et al., 2007), is based on two key assertions which may be made by the reader when processing a negative quantifier. First, that the negative quantifier indicates an amount, and second, that this amount does not reach the amount which was expected or desired. So when *Not many of the fans went to the match* the reader can infer that the amount of fans who went to

the match was not as many as had been expected. Moxey (2006) and Sanford et al., (2007) posit that it is this shortfall between the expected or desired amount and the amount indicated by the negative quantifier which leads to compset reference. This possibility has been empirically tested by assessing processing of references to the compset when a shortfall was made salient through explicit expectations or desires (Moxey, Filik, & Paterson, 2009; Moxey & Filik, 2010). For example, Moxey et al. (2009) examined eye movements during reading and found that when a compset reference followed an initial sentence in which the explicit expectation was intentionally high (e.g. *Robert expected all of the audience to applaud*), processing time and number of regressions were decreased compared to when the expectation was intentionally low (e.g. *Robert expected none of the audience to applaud*). The decreased processing time in the presence of a shortfall between expected and denoted amounts demonstrates participants' recognition of the shortfall and hence rapid availability of the compset for reference.

More recent experiments (Ingram & Moxey, 2011) indicate that a complement pattern of processing is not confined to explicitly negated quantified antecedents. In two experiments that tested sentence production and comprehension (using eye-tracking), the authors demonstrated that a compset preference was also found when a shortfall was *inferred* from the context. Ingram and Moxey manipulated inferred quantity by describing a situation in which a character has a clear, although implicit, desire for a high quantity (as in (3) below). The actual quantity could then be inferred from a positive or negative emotion which the character felt about this quantity. That is, where a character implicitly desires a high amount and that character is described as experiencing a negative emotion in reaction to the situation, it should be clear that the actual amount is not as high as desired, which leads to a shortfall. For example, in (3), the judge wishes all potential jurors to turn up for jury duty. When the emotion is positive this indicates that the judge's desire has been achieved, the amount of jurors turning up

is large and so a continuation to the set who turned up (the refset) is more likely to be produced by participants and is more easily integrated during processing.

(3) The judge was *happy/angry* about the number of people who turned up for jury duty. Their *presence/absence* would affect when the case could commence.

In contrast, when the emotion in the sentence is negative this indicates that the amount of jurors who turned up does not reach the judge's desired amount, creating a shortfall between desired and implied amounts, and so a continuation to the set who did not turn up (the compset) was more likely to be produced. In eye-tracking, an increase in reading times and regressions was seen when a negative emotion was followed by an incongruent reference to the refset. However, in line with previous eye-tracking studies involving quantified antecedents (Paterson et al., 1998), this processing disruption was less pronounced than when a positive emotion was paired with a compset reference. The current paper seeks to extend this research by employing ERP methods, which allow the examination of brain processing during reading with a millisecond by millisecond temporal resolution.

ERPs have been used successfully in recent research to investigate the inferences that readers draw from quantified statements, specifically focusing on the N400 component (e.g. Filik et al., 2011; Kounios & Holcomb, 1992; Nieuwland, 2016; Urbach, DeLong, & Kutas, 2015; Urbach & Kutas, 2010). This component is a negative-going deflection in the ERP waveform, typically centroparietal, and begins around 200ms post word onset and peaks around 400ms. It is thought to index semantic retrieval of a given word, and the degree to which context facilitates this process (Kutas & Hillyard, 1980; Kutas, & Van Petten, 1994; Van Berkum, Hagoort, & Brown, 1999). Crucially, the N400 has been shown to be modulated during anaphoric pronoun resolution to quantified antecedents. Filik et al. (2011) presented

participants with sentences that included positive (e.g. *many*) and negative (e.g. *not many*) quantifiers, followed by a continuation that focused either on the reference set (e.g. *Their presence*…) or the complement set (e.g. *Their absence*…). Results showed that the anaphoric pronoun, *their,* elicited a larger N400 following negative than positive quantifiers. In addition, there was a larger N400 on the disambiguating word (*presence*/*absence*) for complement set reference following a positive quantifier, and for reference set reference following a negative quantifier that implied a shortfall. These processing patterns suggest that readers can assign referents to anaphoric pronouns following positive and negative quantifiers incrementally during comprehension.

The current paper investigates for the first time whether positive and negative quantities that are implied from character desires and emotions elicit a comparable increase in the N400 when the inconsistent set is referenced. Specifically, we aim to establish whether compset references are immediately accessible to readers when they follow a context in which a shortfall must be inferred from an implied desire and quantity, or whether readers experience a delay in processing relative to refset references following a context in which an implicit desire is met. Results of Filik et al., suggest that integration of a compset reference with a shortfall established by a preceding negative quantifier was evident at the N400 component despite delayed integration of the complement set with the preceding context in eye-tracking studies (Paterson et al., 1998). Therefore, if pronominal reference, where the shortfall must be inferred, follows the same pattern as the explicit quantifiers examined in Filik et al. (2011) we should see a larger N400 for incongruent versus congruent sets (*absence/presence*), both when a positive emotion is matched with a compset reference (*e.g. The Judge was happy...Their absence...*) and when a negative emotion is matched with a refset reference (*e.g. The judge was angry...Their presence...*). However, if the process of inferring a shortfall between desired and actual amounts requires extra processing, over and above processing of a shortfall explicitly

defined with a quantifier, and therefore delays access to the compset, as suggested by previous eye-tracking results (Ingram & Moxey, 2011), we should see an increase in the N400 component for compset references, regardless of set congruency.

Method

Participants

Twenty-four participants (M_{age} = 22.9 years; 11 females; 20 right-handed) were recruited from the University of Kent. All participants were native speakers of British English, had normal or corrected to normal vision and did not have any reading disorders e.g. dyslexia. Each participant was paid for taking part in the study.

Materials & Design

The experiment employed a 2 (emotion word: positive *vs.* negative) x 2 (set: refset *vs*. compset) within participants design. The materials were 160 passages, each containing 2 sentences. Each item was made up of an initial sentence that described a character's positive or negative emotion in response to the number of a set who fulfilled the predicate of the sentence (e.g. "The judge was happy/angry about the number of people who turned up for jury duty"). The second sentence was the critical sentence; it began with the pronoun *Their* and referred to either the refset or the compset (e.g. "Their presence/absence would affect when the case could commence"). The critical word was always the second word in this sentence. An example item in each condition can be seen in Table 1, where the critical word is marked in bold, and further experiemental items can be seen in Appendix 1.

Given that the amplitude of the N400 have been shown to be related to word predictability, experimental items were pretested using latent semantic analysis (landauer, Foltz

& Dumais) to compare the strength of the lexical-semantic relationship between thecritical word (e.g. presence/absence) and the preceding test in each item/condition; strong relationships elicit higher LSA values, which are associated with reduced N400 amplitudes. Statistical analysis performed using an ANOVA crossing emotion word and set[1]. Results revealed an interaction between emotion and set, F (1, 146) = 21.51, MSE = 1.76, $p < .001$. Post hoc comparisons, with Bonferroni adjustments, showed that LSAs were higher for refset than compset critical words following both a positive (.31 vs. .29), F (1, 146) = 6.99, $p = .009$ and a negative emotion (.31 *vs*. .29), F (1, 146) = 5.50, $p = .02$. Critical words referring to the refset had higher LSAs when they followed a positive as opposed to negative emotion word, F (1, 146) = 10.44, $p = .002$, but there was no difference for critical words referring to the compset, F (1, 146) = 2.35, $p = .13$.

Experimental items were presented over four lists, with each list containing 160 experimental items, 40 in each of the four conditions. Each list also contained 80 unrelated filler sentences, which were randomly distributed among the experimental sentences. Thus, each participant saw each experimental sentence once, in one of the four conditions. Six participants were randomly assigned to read each list.

---------- Table 1 about here ----------

Procedure

Participants were informed about the EEG procedure and experimental task. After electrode application they were seated in a booth where they read the materials from a computer screen. There were four practice trials to familiarize them with the procedure. Each trial began with presentation of a single centrally-located red fixation cross for 500ms. After this time, a white

[1] Thirteen items were excluded from the analysis as they contained words not found in the corpus (http://lsa.colorado.edu/).

fixation cross appeared for 500ms. Next, the initial sentence was presented on the screen, and participants were instructed to read this sentence and press the spacebar on a keyboard to continue when ready. A blank screen appeared for 500ms, followed by a fixation cross (500ms). The target sentence was then presented word-by-word, with each word appearing at the centre of the screen for 300ms, with a 200ms blank-screen between words. Target words were always the second word presented and the preceding word was always the plural pronoun *Their*. A 2500ms blank-screen interval followed each item. Trials appeared in ten blocks of twenty-four trials. Each block was separated by a self-paced break. Thus, participants were tested in a single session that lasted approximately one hour, during which they were seated in a comfortable chair located in an isolated room.

Electrophysiological Measures

A Brain Vision Quickamp amplifier system was used with an ActiCap cap for continuous recording of electroencephalographic (EEG) activity from 62 active electrodes over midline electrodes Fz, Cz, CPz, Pz, POz, and Oz, over the left hemisphere from electrodes Fp1, AF3, AF7, F1, F3, F5, F7, FC1, FC3, FC5, FC7, C1 C3, C5, T7, CP1, CP3, CP5, TP7, TP9, P1, P3, P5, P7, PO3, PO7, PO9, O1, and from the homologue electrodes over the right hemisphere. EEG and EOG recordings were sampled at 1000 Hz, and electrode impedance was kept below 10kΩ. Off-line, all EEG channels were recalculated to an average mastoid reference.

Prior to segmentation, EEG and EOG activity was band-pass filtered (0.01-30 Hz, 12 dB/oct), and EEG activity containing blinks was corrected using a semi-automatic ocular ICA correction approach (Brain Vision Analyzer 2). The continuous EEG record was then segmented into epochs of 1700ms, starting 200ms prior to the onset of the target word. Thus, the post-stimulus epoch lasted for a total duration of 1500ms. Semi-automatic artifact detection software (Brain Vision Analyzer 2) was run, to identify and discard trials with non-ocular

artifacts (drifts, channel blockings, EEG activity exceeding ± 75μV). This procedure resulted in an average of 32 segments per condition.

ERP Data Analysis

For analysis of the EEG data, the signal at each electrode site was averaged separately for each experimental condition time-locked to the onset of the target word. Before the measurement of ERP parameters, the waveforms were aligned to a 200ms baseline prior to the onset of the target word. To analyze experimental effects on the N400, mean ERP amplitude was calculated in the time interval from 300-450ms relative to target word onset.

ERP amplitudes were analysed separately for midline and lateral electrode sites. Midline analyses included electrodes Fz, Cz, CPz, Pz, POz, and Oz. Lateral electrodes were analyses using four regions of interest (ROIs). Electrodes were divided along a left-right dimension, and an anterior-posterior dimension. The two ROIs over the left hemisphere were: left-anterior (Fp1, AF3, AF7, F1, F3, F5, F7, FC1, FC3, FC5, FT7), and left-posterior (CP1, CP3, CP5, TP7, P1, P3, P5, P7, PO3, PO7, PO9, O1); two homologue ROIs were defined for the right hemisphere.

Results

Behavioural Analysis of Initial Sentence

Reading times for the initial sentence (i.e. from sentence presentation to participant pressing the space bar) revealed significantly faster reading times when the sentence described a positive emotion (M = 4468ms) compared to when it described a negative emotion (M = 4689ms), [$t(22) = 2.64$, $p = .01$][2].

[2] Note that one behavioural data file was corrupted and therefore not available for inclusion in this reading time analysis.

N400 Analysis

Figure 1 displays the grand-average waveforms for the four conditions crossing emotion (positive *vs*. negative) and set (compset *vs*. refset), as well as topographic maps of the ERP difference waveforms (compset *minus* refset, separately for positive and negative conditions) over the N400 time period. These plots show that a compset reference, regardless of the preceding emotion word, triggered a more negative-going deflection in the N400 wave (beginning at approximately 300ms after target word onset).

---------- Figure 1 about here ----------

Statistical analysis of the N400 effect over midline sites employed a 6 (electrode) × 2 (emotion word: positive *vs.* negative) × 2 (set: refset *vs.* compset) repeated-measures ANOVA. This revealed a significant main effect of set, $F(1, 23) = 6.90$, MSE = 17.31, $p = .01$, with compset references leading to a more negative N400 wave (-1.17 μV) than refset references (-.26 μV). None of the remaining effects reached significance (*p*s > .2).

Statistical analysis over lateral electrode sites employed a 2 (hemisphere: left *vs.* right) × 2 (ant-pos: anterior *vs.* posterior) × 2 (emotion word: positive *vs*. negative) × 2 (set: refset *vs.* compset) repeated-measures ANOVA. This analysis revealed a main effect of set, $F(1, 23) = 8.28$, MSE = 32.81, $p = .01$, with compset references leading to a more negative N400 wave (-.5 μV) than refset references (-.08 μV). In addition, the interaction between ant-pos × hemisphere × emotion word was significant, $F(1, 23) = 5.17$, MSE = 1.45, $p = .03$. however none of the post-hoc comparisons reached significance once the Bonferroni correction was applied. Moreover, we found a significant interaction between ant-pos × emotion word × set, $F(1, 23) = 6.54$, MSE = 2.01, $p = .02$, reflecting a significant emotion word × set interaction over posterior sites, $F(1, 23) = 4.38$, MSE = 1.38, $p = .05$. In line with our predictions, planned

comparisons were conducted, which revealeed that the N400 was significantly larger to compset references than to refset references only when they had been preceded by a positive emotion word (-1.22 μV *versus* -.03 μV respectively; $t(23) = 3.29$, $p = .003$), but not when they had been preceded by a negative emotion word (-.86 μV *versus* -.67 μV respectively; $t(23) = .60$, $p = .55$). There was no difference in the size of the N400 between positive and negative emotions elicited by a refset reference (-.04 μV *versus* -.67 μV respectively; $t(23) = 1.90$, $p = .07$), or a compset reference (-.86 μV *versus* -1.22 μV respectively; $t(23) = -1.06$, $p = .30$). These results suggest that a compset reference may elicit processing difficulty regardless of the polarity of the preceding emotion word; the inconsistent set (i.e. compset) was only detected immediately following a positive emotion word.

Discussion

This paper provides novel data on the range of circumstances under which anaphoric reference to the complement set is immediately available. We investigated whether positive and negative quantities that are implied from character desires and emotions elicit a comparable increase in the N400 when the inconsistent set is referenced (i.e. the reference set (e.g. Their *presence*…) or the complement set (e.g. Their *absence*…)). Previous ERP research by Filik et al. (2011) examined comprehension when positive (e.g. *many*) and negative (e.g. *not many*) quantifiers explicitly established whether a shortfall was present. Results revealed a comparable N400 mismatch effect when a positive emotion word (e.g. X was *happy*, indicating there is no shortfall) was followed by a compset reference, and when a negative emotion word (e.g. X was *angry*, making a shortfall salient) was followed by a refset reference. In contrast, recent eye-tracking evidence suggests that extra processing is required to infer a shortfall between desired and actual amounts which delays access to the compset, regardless of set congruency (Ingram & Moxey, 2011; Patterson et al., 1998).

The current results revealed a typical N400 effect over posterior sites when a compset reference followed a positive emotion word (which implied a large quantity). This pattern reflects previous studies in showing difficulty integrating compset references following both explicit positively quantified antecedents (Filik et al., 2011; Paterson et al., 1998), and positive quantities that are implied from emotion words (Ingram & Moxey 2011). As such, the N400 effect for the positive-compset condition could reflect either the mismatch between the implied positive quantity and the critical word (*absence*), or a general processing difficulty in inferring a shortfall between desired and actual amounts. Results from the conditions following a negative emotion word (implying a small quantity) provide further insights into these possibilities, as no difference in N400 amplitude was found between compset and refset references over posterior sites. This suggests that both types of reference were equally acceptable, and that participants did not have immediate access to the compset in situations where a shortfall must be inferred between a character's implicitly desired amount and a low amount implied by a negative emotion word.

We note that this pattern of effects cannot be explained by simple lexical-semantic word predictability, since LSA analyses showed equivalent refset/compset differences following positive and negative contexts. However, it is possible that the processing benefit for a refset reference following a positive emotion word is a result of emotional congruency. Consideration of the stimuli (Appendix 1) suggests that critical words referring to the refset have a positive valence whilst those referring to the compset have a negative valence (e.g. success vs. failure). The emotion word and critical word may be matched or mismatched in terms of valence with positive emotion words being congruent with positive (refset) critical words, and negative emotion words being congruent with negative (compset) critical words. Research suggests that when there is an emotional mismatch in stimuli, as would be the case in the positive-compset and negative-refset conditions, this may lead to larger N400 effects (Eder,

Leuthold, Rothermund & Schweinberger, 2012; Zhang, Lawson, Guo & Jiang, 2006). It is possible that emotional congruency has affected the current results as a potentially negative compset reference led to a larger N400 than a potentially positive refset reference after a positive emotion over posterior sites. Mean amplitudes for other comparisons between conditions support an effect of emotional congruency although these differences were not statistically significant.

Overall, the N400 component was more negative-going following a compset reference than a refset reference, regardless of the prior emotion word, suggesting that integration of the compset was generally more difficult. Results from Ingram and Moxey's (2011) eye-tracking study suggested that despite an initial delay, readers did eventually infer the shortfall between desired and actual amounts (i.e. on later measures on reading), and a reference to the compset became preferred over a reference to the refset. Analysis of the N400 window in the current data does not replicate this finding since a compset reference was never preferred over a refset reference following a negative emotion word. Thus, we suggest that inferring a shortfall based on a character's negative emotion has led to compset and refset references being equally acceptable within the N400 time period. However, we concede that the overall main effects of set suggest that compset references are unlikely to be available to readers within 400ms of the critical word onset. That is, at the time of encountering the critical word the shortfall has not yet been inferred, thus the compset is not yet the preferred set for reference.

It is likely that differences in paradigm sensitivity account for the differences in sensitivity to implied quantities. Specifically, ERP paradigms typically present linguistic stimuli as a stream of individual words and effects are time-locked to the onset of a critical word. While these procedures are important for time-locking ERP effects, they also limit the degree to which later influences on processing (e.g. reanalysis of text) are observed (since earlier ERP effects influence downstream waveforms). Visual inspection of the waveforms in

Figure 1 does not indicate differences between compset and refset following a negative emotion at later time points (though note that these may emerge if different time-locking points, following the critical word, were analysed). In contrast, eye-tracking paradigms allow analysis of natural reading over longer narratives, and thus allow readers to increase processing time on words that are difficult to integrate and revisit important parts of the text to make sense of processing difficulties. As such, eye-tracking paradigms are able to capture both the initial difficulty inferring a shortfall following a negative emotion (demonstrated by Ingram and Moxey as a lack of compset/refset difference), as well as the later compset preference (either on later reading measures or on post-critical regions of text). Similar complementary findings between ERP and eye-tracking paradigms have been observed by Ferguson, Sanford and Leuthold (2008), who showed that context-level negations were not assimilated immediately in ERP data, but were fully integrated on later processing measures in an eye-tracking reading study (see also Filik & Leuthold, 2013). The results presented here are important in demonstrating the processing delay following a negative emotion with millisecond accuracy, and in showing for the first time an overall difficulty integrating a reference to the compset.

In addition to the effects of set, results also showed that the valence of the emotion word led to differences in processing. Reading times for the initial sentence were slower when the emotion word was negative. This pattern reflects results of Ingram and Moxey (2011) where there were more regressions into the emotion word when it was negative rather than positive. Research into the processing of emotion words suggests that words with a negative valence are responded to more slowly than those with a positive valence (see Kuperman, Estes, Brysbaert & Warriner, 2014, for a review). Differences between processing of emotion words have also been found during fluent reading, where there is a processing advantage for positive emotion words (Scott, O'Donnell & Sereno, 2012), and differences in the processing of emotion words may account for the differences in reading times for our initial sentence.Using

ERPs high frequency negative emotion words have been shown to incur additional cognitive resources in early components (Scott, O'Donnell, Leuthold & Sereno, 2009). Differences in the processing of emotion words may account for the differences in reading times for our initial sentence. However as previous studies focused on effects of emotion on the emotion word itself, it is difficult to interpret how encountering a negative emotion word in the preceding context has affected processing of the critical word independent of other factors.

Taken together, our results could be interpreted as supporting semantic accounts of anaphoric reference (e.g. Nouwen, 2003). Such accounts would suggest that, although possible, compset references are only available after the resolution of a pronoun to the refset has failed. According to this account, the refset reference should initially be favoured over the compset even when the preceding context makes reference to the compset more appropriate. However, the interaction over posterior electrodes in the current data, and the absence of a set effect on the N400 following a negative emotion, goes against this proposal suggesting that refset is not the default inference when interpreting a shortfall between the desired and actual amounts. As an alternative, the pre-supposition denial account (Moxey, 2006; Sanford, et al., 2007) suggests that compset reference is reliant on reader's recognition of a shortfall between expected and denoted amounts. To support this, Ingram and Moxey (2011) found that using implicit desire and negative emotion words was sufficient to allow compset references to be become preferable to refset references in later processing measures. As well as the methodological differences between studies described above, it is possible that this pattern is not reflected in the current study since the manipulation of character's desire and negative emotion word was not sufficient to make the shortfall set immediately available to participants. Participants were only provided with one introductory sentence, containing the positive or negative emotion word, and from this information were required to infer both that the character would desire a high amount and the amount the emotion word suggested before focusing their attention on a

possible shortfall set. In contrast, Ingram and Moxey's study included an additional introductory sentence, which gave neutral but important contextual details which may have helped establish the high amounts expected by the character (e.g. *"The birthday party had been in full swing for hours")*. This foreword may have given time for participants to access the schema relevant for birthday parties consequently assisting them in inferring the character's desire, in this case the desire to receive a large number of gifts. Related research suggests that the shortfall set may become available for reference through establishing an upper-bound (Breheny, Katsos & Williams, 2006). Thus, when the character's desires imply the maximum available (e.g. that the judge wanted all of the people to show up for jury duty), it is then easier to identify the shortfall when a smaller actual amount is denoted or implied. It is likely that participants in this study did not have sufficient time or information to recognise the upper-bound implied by the implicit character desire and hence the shortfall set, or compset, did not become available when this desire was denied by a negative emotion word. Thus, our results partially support the pre-supposition denial account in showing that successful integration of the compset is reliant on readers making the inference about a shortfall.

In conclusion these results extend the body of literature on complement anaphora suggesting that in situations where a shortfall must be inferred from implicit desire and quantity information a compset reference is not immediately available. This contrasts with similar studies involving explicitly quantified antecedents where a reference to the compset is available on processing of the critical word (Filik et al., 2011), but is in line with previous eye-tracking results where a shortfall has been implicitly created and references to the compset are preferred only in later processing (Ingram & Moxey, 2011).

References

Breheny, R., Katsos, N., & Williams, J. (2006). Are generalized scalar implicatures generated by default? An online investigation into the role of context in generating pragmatic inferences. *Cognition, 100*, 434-463. doi:10.1016/j.cognition.2005.07.003

Eder, A. B., Leuthold, H., Rothermund, K., & Schweinberger, S. S. (2012). Automatic response activation in sequential affective priming: An ERP study. *SCAN*, *7*, 436-445. doi: 10.1093/scan/nsr033

Ferguson, H. J., Sanford, A.J., & Leuthold, H. (2008). Eye-movements and ERPs reveal the time-course of processing negation and remitting counterfactual worlds. *Brain Research*, 1236, 113-125. doi:10.1016/j.brainres.2008.07.099

Filik, R., & Leuthold, H. (2013). The role of character-based knowledge in online narrative comprehension: Evidence from eye movements and ERPs. *Brain Research, 1506*, 94-104. doi:10.1016/j.brainres.2013.02.017

Filik, R., Leuthold, H., Moxey, L. M., & Sanford. A. J. (2011). Anaphoric reference to quantified antecedents: An event-related brain potential study. *Neuropsychologia, 49*, 3786-3794. doi:10.1016/j.neuropsychologia.2011.09.043

Kounios, J., & Holcomb, P. J. (1992). Structure and process in semantic memory: evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General*, *121*(4), 459-479.

Kuperman, V., Estes, Z., Brysbaert, M., & Warriner, A. B. (2014). Emotion and language: Valence and arousal affect word recognition. *Journal of Experimental Psychology: General, 143,* 1065-1081. doi:10.1037/a0035669

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203-205.

Kutas, M., & Van Petten, C.K. (1994). Psycholinguistics electrified: Event-related brain potential investigations. In: Gernsbacher MA, editor. *Handbook of Psycholinguistics*. San Diego: Academic Press, 83–143.

Landauer, T.K. & Dumais, S.T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-140.

Moxey, L. M. (2006). Effects of what is expected on the focussing properties of quantifiers: A test of the presupposition-denial account. *Journal of Memory and Language*, *55*, 422–439. doi:10.1016/j.jml.2006.05.006

Moxey, L. M., Filik, R., & Paterson, K. (2009). On-line effects of what is expected on the resolution of pronouns. *Language and Cognitive Processes, 24 (6).* doi:10.1080/01690960802634038

Moxey, L. M. & Sanford, A. J. (1986). Quantifiers and focus. *Journal of Semantics*, *5*, 189–206. doi:10.1093/jos/5.3.189

Nouwen, R. (2003). Complement anaphora and interpretation. *Journal of Semantics. 20*, 73-113. doi:10.1093/jos/20.1.73

Paterson K.B., Sanford A.J., Moxey L.M., & Dawydiak E.J. (1998). Quantifier polarity and referential focus during reading. *Journal of Memory and Language,* 39 (2), *290-306*. doi:10.1006/jmla.1998.2561

Sanford, A. J., Dawydiak, E., & Moxey, L. (2007). A unified account of quantifier perspective effects in discourse. *Discourse Processes*, 44, 1-32. doi**:**10.1080/01638530701285556

Sanford, A. J., Moxey, L. M., & Paterson, K. B. (1996). Attentional focussing with quantifiers in production and comprehension. *Memory and Cognition*, *24*, 144–155. doi:10.3758/BF03200877

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

Scott, G. G., O'Donnell, P. J., & Sereno, S. C. (2012). Emotion words affect eye fixations during reading. *Journal of Experimental Psychology: Learning, Memory & Cognition, 38*, 783-792. doi:10.1037/a0027209

Urbach, T. P., DeLong, K. A., & Kutas, M. (2015). Quantifiers are incrementally interpreted in context, more than less. *Journal of Memory and Language*, *83*, 79-96.

Urbach, T. P., & Kutas, M. (2010). Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation. *Journal of Memory and Language*, *63*(2), 158-179.

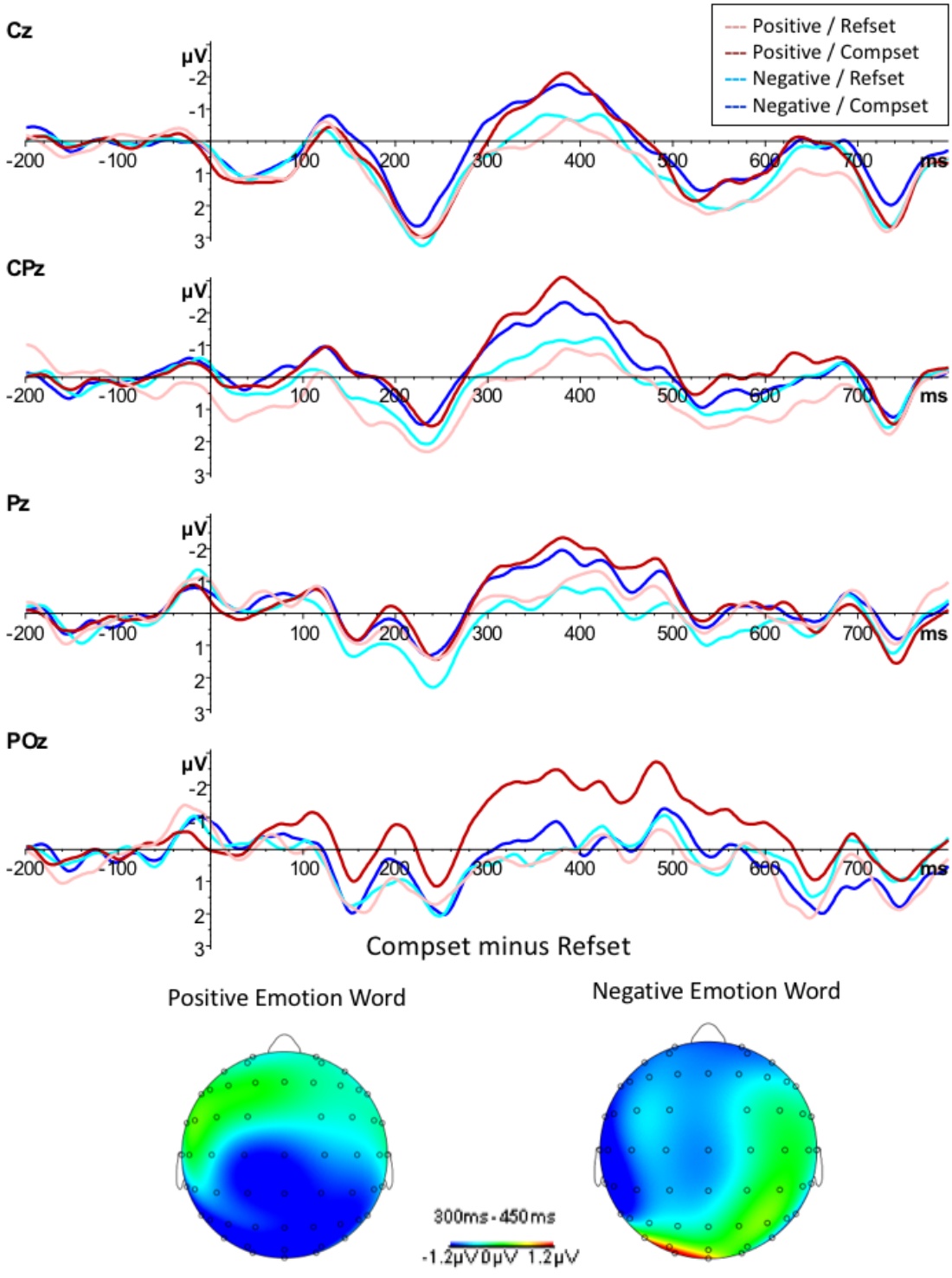
Van Berkum, J. J. A., Hagoort, P., & Brown, C.M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*, 657–671. doi:10.1162/089892999563724

Zhang, Q., Lawson, A., Guo, C., & Jiang, Y. (2006). Electrophysiological correlates of visual affective priming. *Brain Research Bulletin*, *71*, 316-323. doi:10.1016/j.brainresbull.2006.09.023

Table 1: Example material in each of the four conditions.

Condition	Material
Positive emotion / Refset	The judge was happy about the number of people who turned up for jury duty. Their **presence** would affect when the case could commence.
Positive emotion / Compset	The judge was happy about the number of people who turned up for jury duty. Their **absence** would affect when the case could commence.
Negative emotion / Refset	The judge was angry about the number of people who turned up for jury duty. Their **presence** would affect when the case could commence.
Negative emotion / Compset	The judge was angry about the number of people who turned up for jury duty. Their **absence** would affect when the case could commence.

Figure 1: Top: Grand average event-related brain potentials elicited by the critical word for each of the four conditions at electrodes Cz, CPz, Pz and POz. Bottom: Topographic maps of the set effect for positive and negative emotion words for the time period 300-450ms post critical word onset.



Appendix 1. Selection of experimental items. *Positive/negative* emotion words and *refset/compset* critical words shown in italics.

The football coach was *pleased/annoyed* about the number of players who came to training.

Their *presence/absence* would be reflected in Saturday's performance.

The actor was *pleased/ashamed* about the number of people in the audience that applauded.

Their *cheers/silence* would be noted in the newspaper review.

The committee were *happy/angry* about the number of people who contributed to the charity.

Their *kindness/meanness* would affect those in need.

Marion was *pleased/annoyed* about the number of giant pandas remaining in the world.

Their *survival/decline* would be affected by breeding habits.

The headmaster was *delighted/disappointed* about the number of pupils who went on to Cambridge.

Their *brilliance/stupidity* would surely affect the next intake.

Mr Smith was *happy/angry* about the number of his guests who enjoyed the dinner party.

Their *compliments/criticisms* would be repeated to his wife.

The politician was *happy/angry* about the number of supporters who had confidence in his election victory.

Their *optimism/pessimism* would inspire him to keep running.

The surgeon was *pleased/annoyed* about the number of people willing to donate organs.

Their *consent/refusal* would affect a patient's survival.

The golfer was *impressed/unimpressed* by the number of his friends who made it to the competition finals.

Their *success/failure* would mean lots of whiskey.

The judge was *impressed/unimpressed* with the number of people willing to testify against the notorious gangster.

Their *cooperation/reluctance* would determine if a conviction was possible.