Title
What’s Up can be Explained by Language Statistics

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN
1069-7977

Authors
Hutchinson, Sterling
Louwerse, Max

Publication Date
2013

Peer reviewed

What's Up can be Explained by Language Statistics

Sterling Hutchinson (schtchns@memphis.edu)
Department of Psychology/ Institute for Intelligent Systems, University of Memphis
365 Innovation Drive, Memphis, TN 38152 USA

Max M. Louwerse (maxlouwerse@gmail.com)
Department of Psychology/ Institute for Intelligent Systems, University of Memphis
365 Innovation Drive, Memphis, TN 38152 USA

Tilburg Centre for Cognition and Communication (TiCC), Tilburg University
PO Box 90153, 5000 LE, Tilburg, The Netherlands

Abstract

Embodied cognition studies have demonstrated that when words found in high physical locations (e.g., *bird*) are positioned at the top of a screen they are processed faster than when they are positioned at the bottom of the screen. The reverse effect is obtained for words found in low physical locations (e.g., *fish*). This concept-location facilitation effect has been argued to demonstrate that cognitive processing is fundamentally perceptual in nature. However, questions can be raised with regards to the absolute or relative location of these concept-location words We investigated whether semantic judgments were made with respect to an absolute location on the screen (embodied explanation) or with respect to a relative location in comparison to other words included in the experimental session (statistical linguistic explanation). In a response time experiment we presented participants with physical-location words from existing studies at the top or bottom, top or center, and center or bottom of the screen. For animate words we found a concept location facilitation effect for words presented at the top of the screen, at the center of the screen, and at the bottom of the screen. In addition, however, language statistics explained RTs to center words. Findings indicated that participants made judgments relative to other words on the screen and not relative to their absolute location on the screen, lending support to a statistical linguistic explanation of the findings.

Keywords: concepts; embodied cognition; symbolic cognition; concept-location facilitation; perceptual

Introduction

Embodied cognition theories state that language is understood through perceptual representations that are grounded in modality-specific somatosensory experience (Barsalou, 1999; Glenberg, 1997; Semin & Smith, 2008). Words become meaningful only after mentally reenacting external perceptions and experiences associated with that word. Thus, the patterns of neural activity that occur when comprehending a particular word would be similar to those patterns that occur when actually perceiving its referent (Hauk, Johnsrude, & Pulvermüller, 2004). In other words, according to embodied cognition theories mental representations are couched in the physical and perceptual experiences of the body.

There is a wealth of evidence supporting the embodied cognition account, with evidence showing that when experimental tasks cue participants to refer to relevant perceptual representations, language processing is facilitated (Glenberg & Kaschak, 2002; Kaschak et al., 2005; Pecher, van Dantzig, Zwaan, & Zeelenberg, 2009). For example, Zwaan and Yaxley (2003) demonstrated that when word pairs appeared in their expected physical locations on a computer screen (e.g., *ceiling* presented at the top of the screen and *floor* presented at the bottom of the screen), comprehension was faster than when pairs appeared in unexpected physical locations on a computer screen (e.g., *floor* presented at the top of the screen while *ceiling was* presented at the bottom of the screen). That is, it is easier to process a word when the expected physical properties of the word match its actual physical properties. Accumulating research like this tends to suggest that individuals rely on perceptual representations, especially in everyday language comprehension.

This embodied cognition account of semantic representations is often contrasted to an amodal (or symbolic) account of cognition, whereby language is represented amodally. A classical symbolic account of language representation argues that semantic information is seated in language and can be derived from relationships that exist between symbols instead of from the mental reenactment of biomechanical and perceptual experiences. In other words, meaning is represented in a linguistic structure within the brain encoded in a formal abstract language, and words are understood from their natural linguistic context instead of from their perceptual features.

Recently, several studies have argued that an extreme symbolic or an extreme embodied cognition account is untenable, and that a more plausible cognitive model includes both perceptual and symbolic processes in language comprehension (Barsalou, Santos, Simmons, & Wilson, 2009; Louwerse, 2008; 2011a; Paivio, 1986). For instance, Louwerse (2008; 2011a) proposed the Symbol Interdependency Hypothesis. This hypothesis predicts that language encodes the perceptual information we tend to simulate. Consequently, language statistics allows for bootstrapping meaning with only minimal symbol grounding in perceptual experiences. Put differently, according to the idea of symbol interdependency embodied simulations and symbolic relationships are complementary in conceptual processes.

We also know from previous research that language statistics and perceptual simulations explain cognitive processes to different extents under different conditions. For example, linguistic representations are relatively more prominent early during processing whereas complete perceptual representations take longer to generate (Louwerse & Connell, 2010; Louwerse & Hutchinson, 2012). Louwerse & Jeuniaux (2010) found that both task and stimulus influenced whether participants were more likely to rely on linguistic or perceptual information. Thus, findings reporting effects for word pairs attributed to embodied cognition (e.g., Zwaan & Yaxley, 2003) might likely also be explained by a statistical linguistic account. For example, when participants were asked to make a semantic judgment about word pairs, the statistical linguistic frequency of the word pair best predicted RTs whereas when participants were asked to make an iconic judgment about image pairs, perceptual ratings about the pair better accounted for RTs (Louwerse & Jeuniaux, 2010). Although both the linguistic and perceptual information about the word pair showed to be relevant in both cognitive tasks, with both verbal and non-verbal stimuli, different types of information were more, or less, important across different conditions.

These studies demonstrate that both language statistics and perceptual simulation must be taken into consideration together. After all, the Symbol Interdependency Hypothesis argues that language encodes perceptual information, making it difficult to disentangle the two variables. That is, effects attributed to statistical linguistic frequencies could also be attributed to perceptual simulation and vice versa. Furthermore, studies demonstrating a language statistics effect use word pairs as stimuli (e.g., Louwerse & Hutchinson, 2012; Louwerse & Jeuniaux, 2010; Tse, Kurby, & Du, 2010).

However, evidence supporting an embodied cognition account also comes from single words, presented in different locations on a computer screen. For example, Šetić and Domijan (2007) presented ‘up’ and ‘down’ words one at a time either in an expected physical location or in an unexpected physical location (e.g., *butterfly* would either appear at the top of the screen (expected location) or at the bottom of the screen (unexpected location)). Participants were asked to determine if the word they saw was something animate (living animal) or something inanimate (non-living entity). As expected, patricipants were faster to process concept-location matches (e.g., *butterfly* presented at the top of the screen) than concept-location mismatches (e.g., *butterfly* presented at the bottom of the screen). Unlike experiments comparing word pairs, findings for words in isolation, such as those in Šetić and Domijan (2007), are more difficult to also explain with a statistical linguistic account. That is, unigram word frequency does not explain congruency effects, as the set of ‘up words’ are not all more orless frequent than the set of ‘down words’. In fact, when comparing how frequently the ‘up words’ and ‘down words’ occurred in a massive corpus of the English language (the *Web 1T 5-gram* corpus; Brants & Franz, 2006), no difference was obtained between the frequencies of ‘up words’ and ‘down words’ inform the Šetić and Domijan (2007) study, $t(153.37) = 0.64$, $p = .52$. Consequently, the concept-location word results only seem to support an embodied cognition account and are argued to be due to the congruency of the presentation location and the perceptual features of the word: *butterfly* is processed quickly at the top of the screen because a mental simulation of a butterfly involves perceptual and spatial information about where a butterfly is found in the actual world (above the ground/at the top). This poses a challenge to an account that argues for both linguistic and perceptual simulations factors in conceptual processing, such as proposed by the Symbol Interdependency Hypothesis.

Although it seems straightforward to conclude that these effects must be due to the mental simulation of words, there are alternative explanations. Lakens (2011a; 2011b) argues that such effects might instead be due to polarity correspondence. Proctor and Cho (2006) found that in binary classification tasks, concepts can be processed faster when their polarity matches the response polarity. In other words, when a stimulus and a response are coded as either both positive or both negative, processing is facilitated, e.g., *butterfly* is processed quickly at the top of the screen because its location is positive (up), as is the response to whether or not it is found in the sky (yes). In order to rule out a polarity correspondence explanation for the results, in a similar experiment, Pecher, van Dantzig, Boot, Zanzolie, and Huber (2010) asked participants to respond to the question *Is it usually found in the ocean?* or to the question *Is it usually found in the sky?*. They argued that for a polarity correspondence explanation to be valid, *yes* responses would be expected to be processed faster at the top of the screen, regardless of the question being asked, and regardless of word meaning. For instance, when being asked if an animal is found in the ocean, one would expect *butterfly* to be processed faster at the bottom of the screen because it is not found in the ocean, a hypothesis contrary to an embodied cognition explanation and a hypothesis that was not supported. Instead, the results showed just the opposite, i.e., when being asked if the animal is found in the ocean, *butterfly* was still processed faster at the top of the screen. In a response, Lakens (2011b) still suggested that perhaps *butterfly* is processed faster at the top of the screen, even when participants are making an ocean judgment because the judgment becomes a relative assessment with down as the default response (as all comparisons are made with reference to the ocean, which is down).

Lakens (2011b) goes further to point out that alternative explanations for data explained solely by perceptual simulations should not be overlooked. In addition, Lakens (2011b) and Louwerse (2011b) both suggest that results from Pecher et al. (2010) might likely also be explained by a statistical linguistic account. That is, although Pecher et al. (2010) concludes that mental simulation accounts for responses in the sky/ocean task, linguistic frequencies do contribute to word meaning and should also be considered. To illustrate, Louwerse (2011b) found that ocean animal names paired with the word *ocean* occur more frequently than ocean animal names paired with the word *sky* (and vice versa for sky animal names) and that these frequencies account for subject RTs. In sum, findings previously attributed to mental simulation accounts can also be explained by a statistical linguistic account, as was also demonstrated in earlier research (Louwerse & Jeuniaux, 2010). These findings illustrate that task instructions might influence response times because *ocean* and *sky* are more or less linguistically associated with the stimuli. In other words, linguistic explanations for these findings should also be explored.

But it remains difficult to offer a linguistic explanation for results when words are presented in isolation. Although task instructions might influence the speeded responses, the frequency of *butterfly – sky* is only able to account for faster RTs for congruent word categories and tasks while still leaving mental simulations to offer the only explanation for the facilitative effect of the congruency of the presentation location and the perceptual features of the word (as unigram word frequency cannot account for these RTs). Perhaps linguistic information might play a role explaining these concept-location effects for isolated words after all. Although words are presented in isolation on the screen (i.e., one word is presented at a time), it is possible that decisions might be made relative to the other words presented in the other trials of the experiment. Such an explanation would suggest that instead of making judgments relative to the congruency between the concept and the absolute position of the word on the screen (i.e., top of the screen or the bottom of the screen), participants are making judgments relative to the other words in the experiment. That is, participants might show a concept-location facilitation effect not because the words are presented on the top and bottom of the screen, but rather because words are asynchronously presented relatively above and below one another throughout the duration of the experiment.

To explore this possibility, in this study we presented participants with isolated words at either the top or bottom (to replicate the original results), top or center, or center or bottom of the screen. According to an embodied cognition account, if responses are faster because word meaning and world location are congruent, we would expect the same high and low words, presented in the center of the screen to show no concept-location facilitation effect because the presentation location is not congruent with the physical and spatial properties of the simulated word. In other words, when *butterfly* is presented in the center of the screen, processing should not be facilitated.

Alternatively, if decisions are based on the relationship between one word relative to the other words in the experiment (as opposed to being relative to the presentation location of the word; a linguistic explanation), then we might find that high words presented in the center of the screen (concept-location mismatch) will still show a concept-location facilitation effect if low words are presented at the bottom of the screen. That is, when *butterfly* is presented in the center of the screen, processing will be facilitated if other words in the experiment are 'below' a butterfly. Similarly, we might find that low words presented at the center of the screen would show a concept-location facilitation effect if high words are presented at the top of the screen. In essence, if concept-location facilitation is found when words are presented in relative positions on the screen (i.e., above/below one another) as opposed to absolute positions on the screen (i.e., at the top/bottom of the screen), it might be the case that perceptual simulation (concept-location facilitation effect) is not entirely accounting for RTs but rather, participants are making decisions about words presented in isolation by comparing those words to the group of words included in the experiment.

Method

Participants

Eighty-seven undergraduate native English speakers at the University of Memphis participated for extra credit in a Psychology course. Participants were randomly assigned to each of the three conditions (words presented at either a) the top of the screen and the center of the screen, b) the center of the screen and the bottom of the screen, or c) the top of the screen and the bottom of the screen).

Materials

The experiment consisted of 48 living animal words that could be found in a low spatial location, (such as the ground or ocean, n=24) or found in the sky (a high spatial location, n=24). The remaining 48 words consisted of non-living objects that could also be found in either high (n=24) or low (n=24) physical locations. Words were extracted from both Pecher et al. (2010) and Šetić and Domijan (2007).

Procedure

The procedure was almost identical to Pecher et al. (2010) and Šetić and Domijan (2007). Participants were asked if words presented on a 1280x1024 computer screen were either living or nonliving. This task has the advantage that it does not bias participants to consciously judge the physical location of a word. The center of the screen was positioned at eye level. Similar to Pecher et al. (2010) and Šetić and Domijan (2007), each trial began with the presentation of three fixation crosses appearing on the screen for 300ms. Fixation crosses were presented either at the top, center, or bottom of the screen, depending on where the proceeding word would appear on the screen. This occurred in order to notify participants where the next word would appear.

Words were presented at either the top and the center of the screen, the center and bottom of the screen, or – as in the original Šetić and Domijan (2007) study the top and bottom of the screen, depending upon the between participants condition. Upon presentation of a word, participants indicated whether the word was living or not living by pressing designed counterbalanced keys on the keyboard (*f* and *j* keys). All words were seen once and were counterbalanced for each participant where half the high spatial location words were presented in the upper position (relative to the other presentation location, i.e., top relative to center/bottom or center relative to bottom) and half in the lower position (i.e., bottom relative to center/top or center relative to top), likewise for the low spatial location words.

If responses were slower than 2,500 ms a message reading 'TOO SLOW' would appear. Participants were asked to try to be as quick and as accurate as possible in their responses. The next trial began immediately after the subject's response or after the feedback message.

Results and Discussion

Eleven participants were removed from the analysis because >40% of their answers were incorrect. All remaining participants were split evenly between conditions. In all analyses, we used the parameters found in Pecher et al. (2010) for outlier identification and removal. Outliers were identified as those correct responses greater than three standard deviations from the mean per subject per item. Outlier removal (as described above) resulted in a loss of 2.8% of the data. All error trials were removed, resulting in a loss of an additional 8.7% of the data.

A mixed-effect regression analysis was conducted on RTs with match/mismatch (match or mismatch between word category (low or high spatial location word) and relative presentation location (relatively high location of top or center or relatively low location of center or bottom)) as a fixed factor and participants and items as random factors (Baayen, Davidson, & Bates, 2008). The model was fitted using the restricted maximum likelihood estimation (REML) for the continuous variable (RT). F-test denominator degrees of freedom were estimated using the Kenward-Roger's degrees of freedom adjustment to reduce the chances of Type I error (Littell, Stroup, & Freund, 2002).

In addition to the location presentation manipulation, we investigated the source of the RT differences in this task, linguistic or embodied. An embodied account would be predict a concept-location facilitation effect, whereas a linguistic account would suggest these same effects are driven by language statistics. To further explore if participants were relying on language statistics, we ran analyses using word frequency as a fixed factor to determine if a possible additional explanation for any concept-location facilitation effects may exist. The word frequency factor was calculated as the log frequency of each word being presented obtained using the Web 1T 5-gram corpus (Brants & Franz, 2006).

Unlike Šetić and Domijan (2007), no significant concept-location facilitation effect was found for words appearing at the top of the screen, $F(1, 2330)=1.46$, $p=.23$, at the center of the screen, $F(1, 1599)=.10$, $p=.75$, nor at the bottom of the screen, $F(1, 2395)=1.76$, $p=.19$. Just as in Pecher et al., (2010) these findings also fail to replicate the concept-location facilitation effect found in Šetić and Domijan (2007). In fact, there was no interaction between location and word category for any of the three word presentation locations and experimental conditions. Pecher et al. (2010) offered the explanation that the concept location facilitation effect is not well understood, with some factors causing facilitation and others causing interference. The linguistic frequency factor did not explain

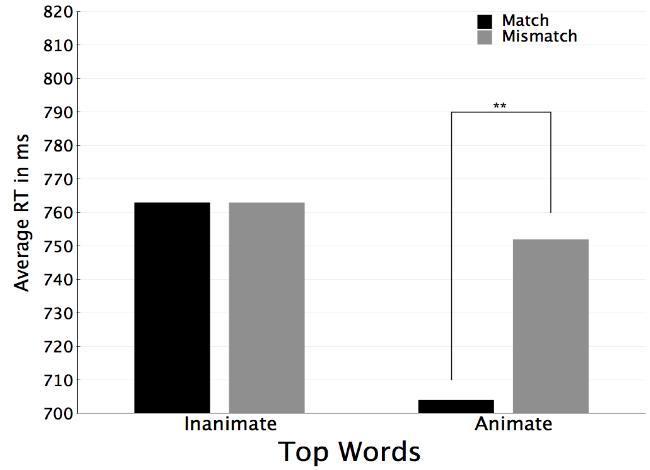

Figure 1: Average RTs in ms for the words appearing at the top of the screen.

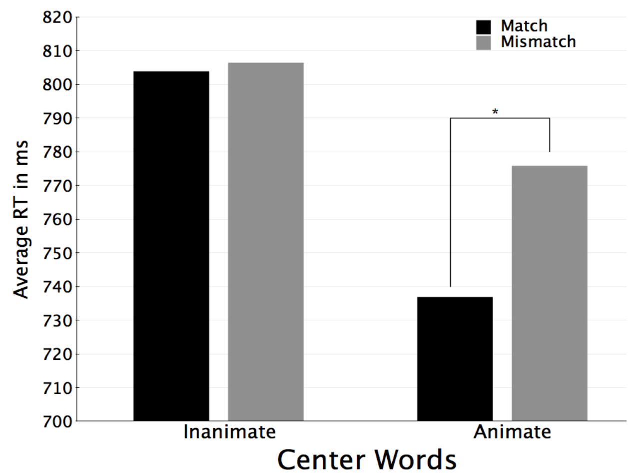

Figure 2: Average RTs in ms for the words appearing at the center of the screen.

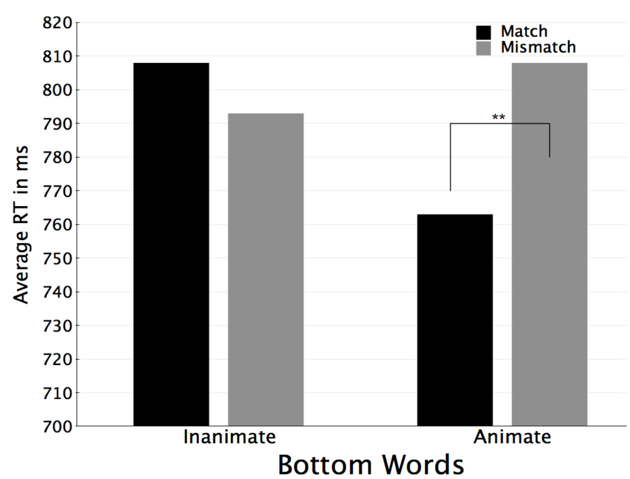

Figure 3: Average RTs in ms for the words appearing at the bottom of the screen.

the results either, with no significant main effects for words appearing at the top of the screen, $F(1, 2330)=.0001$, $p=.99$, the center of the screen, $F(1, 1599)=.19$, $p=.66$, nor the bottom of the screen, $F(1, 2395)=.11$, $p=.74$. These current results seem to support neither an embodied cognition account (as there was no concept-location facilitation for the top-bottom condition) nor an alternative linguistic account (as there was no concept-location facilitation for either condition including the center location nor was linguistic frequency significant). In the absence of a replication in both the current study and in Pecher et al. (2010), perhaps the effects reported in Šetić and Domijan (2007) might be attributed to linguistic differences in the Hungarian stimuli. Alternatively, such concept-location facilitation effects might simply be relevant for certain groups of words and not others.

To further explore the results of the current experiment, and the possibility that words are processed relative to the words around them, we analyzed our findings mixed effects model but for animate versus inanimate words. Words that were inanimate again showed no interactions for words appearing at the top of the screen, $F(1, 1172)=.003$, $p=.96$ (see Figure 1), the center of the screen, $F(1, 787)=.07$, $p=.80$ (see Figure 2), or the bottom of the screen, $F(1, 1072)=.92$, $p=.34$ (see Figure 3). Linguistic frequency was also not significant for words appearing at the top of the screen, $F(1, 1172)=1.53$, $p=.22$, the center of the screen, $F(1, 787)=.62$, $p=.43$, nor the bottom of the screen, $F(1, 1072)=.002$, $p=.96$.

However, words that were animate did show significant interactions. Words appearing in any given location (top, center, and bottom) were processed faster when that location was relatively the same as the word category. 'Up words' presented in the center were processed faster in the center-bottom condition, whereas 'down words' presented in the center were processed faster in the top-center condition, $F(1, 789)=6.10$, $p<.02$. Figure 2 clearly illustrates RTs for matched and mismatched up and down words presented in the center of the screen, showing that words with a concept-location match are processed faster than words with a concept-location mismatch. Similarly, 'up words' presented in the top of the screen were processed faster in both the top-bottom and top-center conditions, $F(1, 1134)=6.80$, $p<.01$, (see Figure 1). Finally, 'down words' presented in the bottom of the screen were processed faster in both the top-bottom and center-bottom conditions, $F(1, 1067)=10.97$, $p=.001$, (see Figure 3).

In addition, to further explore the impact of linguistic frequency also significantly explained RTs to words presented at the bottom of the screen, $F(1, 1067)=5.08$, $p=.02$, but only marginally for words presented in the center of the screen, $F(1, 789)=3.22$, $p=.07$, with no effects for words presented at the top of the screen, $F(1, 1134)=2.58$, $p=.10$. These findings seem to be consistent with the idea that decisions are based on the relationship between one word relative to the other words in the experiment, as 'up words' presented relatively above 'down words' still showed a concept-location facilitation effect despite these words being presented in the center of the screen.

In addition, in all conditions, words appearing relatively below other words (M= 767.45, SD=267.40) were processed significantly slower than words appearing relatively above other words (M= 889.36, SD=421.41), $t(4926) = 15.36$, $p<.001$. That is, regardless of the absolute location of the word on the screen, where-ever the bottom position was (i.e., center of the screen or bottom of the screen), words presented in that location were processed slower than the same words presented in a relatively higher location. Consider the case of the center presentation location: when words were presented in either the center of the screen or the bottom of the screen, words took longer to process at the bottom and less time to process at the center. However, when those same words were presented in the center or the top, they took longer to process in the center and less time to process at the top. This means that the same words presented in the same location are processed faster or slower simply due to whether other words are appearing above or below them. This at least suggests that comparisons between high and low positions are biased given that the center represents both the relative top and bottom in different conditions.

Finally, to explore whether participants indeed made comparative judgments for words, we assessed whether bigram frequencies were able to account for the response times of center words. As in previous studies (Louwerse, 2008) we operationalized the bigram linguistic frequencies as the log frequency of a-b (e.g., *owl-lizard*) or b-a (e.g., *lizard-owl*) order of word pairs. Because words were presented individually on the screen, pairs were determined by the randomized presentation order. The bigram frequency of each pair was assigned to the second word in the randomly presented pair. The order frequency of all word pairs within 3-5 word grams was obtained using the large *Web 1T 5-gram* corpus (Brants & Franz, 2006). A mixed-effect regression analysis was conducted on RTs to center words with the bigram frequency as a fixed factor and participants and items as random factors (Baayen, Davidson, & Bates, 2008). Bigram frequency was a significant predictor of RTs for center words only, $F(1,906)=3.99$, $p=.05$. This was true for all center words regardless of experimental condition, implying that participants consider past trials while making judgments about the current word in question, and implying that a linguistic frequencies explain RTs during a concept-location facilitation task.

General Discussion

In three presentation location conditions (top and center, bottom and center, or top and bottom) we failed to replicate a concept-location facilitation effect as found in Šetić and Domijan (2007) for inanimate words. However, when considering animate words, words matched between the relative presentation location and word category resulted in faster RTs than words with a mismatch. This finding suggests that participants make judgments about individual words they see on the screen with respect to other words they see throughout the duration of an experiment. The absolute location of a word on a screen does not seem to impact the concept-location facilitation effect, but rather the relative location appears to be what is important. This finding suggests that decisions are based on the relationship between one word relative to the other words in the experiment, not only based on the relationship between one word and the embodied physical and spatial properties of that simulated word. In addition, across all three conditions, we found a main effect of location, such that words presented below other words were processed slower. This finding suggested that participants made judgments relative to other words, not only relative to their location on the screen. To further determine whether participants made comparative judgments between words presented asynchronously over the duration of an experiment we also showed that bigram frequencies can predict subject RTs. These findings together indicate that it might be the case that participants are making decisions about words presented in isolation by comparing those words to the group of words included in the experiment, suggesting that findings that are easily attributed to embodied cognition (Pecher et al., 2010; Šetić & Domijan, 2007) can also be attributed to language statistics.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, *22*, 577–660.

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* (pp. 245–283). Oxford, England: Oxford University Press.

Brants, T., & Franz, A. (2006). *Web 1T 5-gram Version 1*. Philadelphia, PA: Linguistic Data Consortium.

Glenberg, A. M. (1997). What memory is for. *Behavioral and Brain Sciences*, *20*, 1–55.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic Bulletin & Review*, *9*, 558–565.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron*, *41*, 301–307.

Kaschak, M. P., Madden, C. J., Therriault, D. J., Yaxley, R. H., Aveyard, M. E., Blanchard, A. A., & Zwaan, R. A. (2005). Perception of motion affects language processing. *Cognition*, *94*, B79–B89.

Lakens, D. (2011a). High skies and oceans deep: Polarity benefits or mental simulation? *Frontiers in Psychology*, *2,* 1-2.

Lakens, D. (2011b). Polarity correspondence in metaphor congruency effects: Structural overlap predicts categorization times for bi-polar concepts presented in vertical space, *Journal of Experimental Psychology, 38,* 726-736.

Littell, R. C., Stroup, W. W., & Freund, R. J. (2002). *SAS for linear models*. Cary, NC: SAS Publishing.

Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review*, *15*, 838–844.

Louwerse, M. M., (2011a). Symbol interdependency in symbolic and embodied cognition. *TopiCS in Cognitive Science, 3*, 273-302.

Louwerse, M. M. (2011b). Stormy seas and cloudy skies: conceptual processing is (still) linguistic and perceptual. *Frontiers in Psychology*, *2,* 1-4.

Louwerse, M. M., & Connell, L. (2010). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science*, *35*, 381–398.

Louwerse, M., & Hutchinson, S. (2012). Neurological evidence linguistic processes precede perceptual simulation in conceptual processing. *Frontiers in psychology*, *3*: 385.

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition*, *114*, 96–104.

Paivio, A. (1986). *Mental representations: A dual coding approach*. New York, NY: Oxford University Press.

Pecher, D., van Dantzig, S., Boot, I., Zanolie, K., & Huber, D. E. (2010). Congruency between word position and meaning is caused by task Induced spatial attention. *Frontiers in Psychology*, *1*, 1–8.

Pecher, D., van Dantzig, S., Zwaan, R. A., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *The Quarterly Journal of Experimental Psychology*, *62*, 1108–1114.

Proctor, R. W., & Cho, Y. S. (2006). Polarity correspondence: A general principle for performance of speeded binary classification tasks. *Psychological Bulletin*, *132*, 416.

Semin, G. R., & Smith, E. R. (Eds.). (2008). *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*. New York, NY: Cambridge University Press.

Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes*, *22*, 297–312.

Tse, C., Kurby, C.A., & Du, F. (2010). Perceptual simulations and linguistic representations have differential effects on speeded relatedness judgements and recognition memory. *The Quarterly Journal of Experimental Psychology, 63*, 928-941.

Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review*, *10*, 954–958.