Quick Linguistic Representations and Precise Perceptual Representations: Language Statistics and Perceptual Simulations under Time Constraints

Sterling Hutchinson (S.C.Hutchinson@uvt.nl)
Tilburg Centre for Cognition and Communication (TiCC), Tilburg University
PO Box 90153, 5000 LE, Tilburg, The Netherlands

Richard Tillman (R.N.Tillman@uvt.nl)
Tilburg Centre for Cognition and Communication (TiCC), Tilburg University
PO Box 90153, 5000 LE, Tilburg, The Netherlands

Max M. Louwerse (M.M.Louwerse@uvt.nl)
Tilburg Centre for Cognition and Communication (TiCC), Tilburg University
PO Box 90153, 5000 LE, Tilburg, The Netherlands

Abstract

Many studies have argued that language comprehension requires perceptual simulation. In previous work we have demonstrated that because language encodes perceptual relations, comprehenders can also rely on language statistics to bootstrap meaning through limited grounding. The extent comprehenders do this depends on the nature of the cognitive task, the stimulus, the individual, as well as the speed of processing, with linguistic representations preceding perceptual simulation. In the current study we report results that investigated whether time constraints impacted the use of perceptual and linguistic factors during language processing. Participants made fast or slow speeded judgments about whether pairs of words were semantically related. Subjects were also instructed to either respond as quickly as possible to the words they were presented, or respond as accurately as possible. The perceptual factor was operationalized as an iconicity rating of the stimulus pairs occurring in a particular orientation in the real world and the linguistic factor was operationalized as the frequency of the stimulus pairs in language. The linguistic factor best explained the RTs when subjects had to respond quickly. On the other hand, when given more time to respond, both linguistic and perceptual factors explained response times. These findings support the view that language processing is both linguistic and embodied, with linguistic representations being relevant for quick good-enough representations and perceptual simulations being important for more precise information.

Keywords: embodied cognition; symbolic cognition; symbol interdependency; perceptual simulation; language processing; time course.

Introduction

Embodied representations are concepts in memory that involve modality-specific simulations of all of the actual experiences associated with those concepts. Proponents of embodied cognition have argued that these embodied representations are fundamental to language processing (Barsalou, 1999; Glenberg, 1997; Pecher & Zwaan, 2005; Semin & Smith, 2008; Zwaan, 2004). Prior research has indeed demonstrated that when experimental tasks elicit the consideration of perceptual features of words, processing does seem to be facilitated (Pecher, van Dantzig, Zwaan, & Zeelenberg, 2009; Šetić & Domijan, 2007; Zwaan, Stanfield, & Yaxley, 2002; Zwaan & Yaxley, 2003). For example, response times (RTs) decrease when words are presented in an iconic ordering (i.e., *pan* – *stove* instead of *stove* – *pan*). Sentences implying particular orientations (e.g., pencil in a drawer) also resulted in faster RTs when an image of the object was presented horizontally as opposed to vertically. Furthermore, Glenberg and Kaschak (2002) also demonstrated embodied cognition with the action-sentence compatibility effect, where participants read sentences that implied specific directional movements (e.g., *Liz told you the story*), and responded faster with a congruent motion than with an incongruent motion. Studies like these are taken as evidence that embodied representations facilitate language processing, and that embodiment is central to cognition.

Alternative to embodied theories, symbolic theories suggest that word meaning can be derived from linguistic context (Landauer & Dumais, 1997) and does not rely solely upon embodied representations. In other words, instead of activating perceptual simulations, mental representations can be viewed symbolic connections between a concept and its 'likeness' within memory (Fodor, 1975; Pylyshyn, 1973).

Recently we have argued for a unified account, by proposing the Symbol Interdependency Hypothesis (Louwerse, 2007, 2011; Louwerse & Jeuniaux, 2008). According to the Symbol Interdependency Hypothesis, language comprehension is linguistic through statistical interdependencies within language and is embodied through the references language makes to perceptual experiences or the external world. That is, when a word is processed, a fuzzy meaning of the word is constructed primarily on the basis of language statistics (for instance, the linguistic context of a word). If a more detailed representation is required, perceptual simulations allow for a precise mental representation. In essence, language acts as a sort of shortcut for language users by encoding symbolic and embodied relations in the world. Importantly, this hypothesis does not reject either account but rather suggests that both perceptual and linguistic representations are used during processing and that their prevalence is relative depending on the cognitive task at hand. For instance, the type of stimulus has been shown to be a factor whether linguistic or embodied representations dominate, with pictures eliciting more reliance on perceptual simulations and words eliciting more reliance on linguistic information (Louwerse & Jeuniaux, 2010).

The cognitive task is also a particularly relevant factor when determining the dominance of linguistic or embodied representations, with subjects depending more on perceptual information when making iconicity judgments and more on linguistic information when making semantic judgments (Louwerse & Jeuniaux, 2010). Because the iconicity judgments imply semantic judgments, Louwerse and Jeuniaux argued that language statistics best explain shallow, good-enough representations, and perceptual simulation best explain detailed full-fledged representations. Louwerse and Connell (2011) tested this hypothesis further and found that when comparing the effect of language statistics and perceptual information on response times, language statistics best explained quick response times, whereas perceptual information best explained slow response times (with both language statistics and perceptual simulation equally contributing to medium response times).

Louwerse and Hutchinson (2012) extended the Louwerse and Connell (2011) study in an EEG experiment, demonstrating that linguistic cortical regions were relatively more active early in a trial and perceptual cortical regions were relatively more active later in a trial. Importantly, both cortical areas were active throughout processing, simply the relative importance of each area varied over time. In sum, processes related to symbolic cognition (i.e., linguistic frequency) are more prominent in the early stages of comprehension whereas processes related to embodied cognition are more prominent later in processing. As subjects glean just enough information from word co-occurrences to understand the relationship between word pairs immediately, while perceptual representations then take over to fill in the rest of the picture. Studies that demonstrate that the relative prominence of linguistic representations precede that of perceptual representations, leave the question open whether RTs actually decrease because perceptual representations are facilitating

Table 1: Related and unrelated iconic word pairs and their LSA cosine values.

Related Word One	Related Word Two	LSA	Unrelated Word One	Unrelated Word Two	LSA
airplane	runway	0.77	belt	shoe	0.23
antenna	radio	0.74	billboard	highway	0.09
attic	basement	0.55	boat	trailer	0.06
car	road	0.43	bouquet	vase	0.28
ceiling	floor	0.72	branch	root	0.23
fender	tire	0.45	bridge	river	0.28
flame	candle	0.59	charcoal	grill	0.19
hat	scarf	0.47	cork	bottle	0.3
head	foot	0.46	cup	saucer	0.4
jockey	horse	0.43	faucet	drain	0.38
kite	string	0.45	flower	stem	0.27
knee	ankle	0.71	fountain	pool	0.31
lid	box	0.57	glass	coaster	0.07
mailbox	post	0.44	handle	bucket	0.2
mane	hoof	0.51	headlight	bumper	0.41
monitor	keyboard	0.51	hiker	trail	0.04
moustache	beard	0.61	hood	engine	0.23
nose	mouth	0.56	lamp	table	0.2
pan	stove	0.51	lighthouse	beach	0.37
pedestrian	sidewalk	0.45	mantle	fireplace	0.06
rocket	launchpad	0.68	penthouse	lobby	0.09
seat	pedal	0.44	pitcher	mound	0.16
smoke	chimney	0.51	plant	pot	0.18
steeple	church	0.52	sheet	mattress	0.12
stirrup	saddle	0.54	sky	ground	0.34
sweater	pants	0.61	sprinkler	lawn	0.11
track	runner	0.63	stoplight	street	0.01
train	railroad	0.66	tractor	field	0.2

processing or rather because the task duration biases subjects to employ embodied representations. In most, if not all, embodied cognition experiments, participants have no pressing time constraints on their responses during a semantic judgment task, other than perhaps several seconds in order to force a decision. It is therefore possible that such a strong effect of perceptual simulation occurs during word processing because subjects are allotted a longer time to process the words they are seeing. Put simply, is it because we are making a slower decision that we rely upon perceptual processes, or is it because we are relying upon perceptual processes do we then make a slower decision? Likewise, is it because we are making a faster decision that we rely upon linguistic processes, or is it because we are relying upon linguistic processes do we then make a faster decision? In the following study we ask whether the RT effects of perceptually related pairs are due to timing constraints in the experiment or if they are indeed due to the perceptual or linguistic relationships between word pairs.

We predicted that subjects might be even more influenced to rely on quick linguistic representations, if the instructions asked them to make responses as quickly as possible, and would be more likely to rely on perceptual representations if they were asked to respond as accurately as possible. Importantly, responses that take longer would then be more likely to encourage the use of perceptual representations, which is consistent with prior research. Therefore, our prediction was that if participants were under a speed time constraint and instructed to be as fast as possible, then they would rely on linguistic information, such as statistical linguistic frequencies. If the participants were given more time or instructed to be more accurate, however, they would access perceptual representations.

Methods

Design

The experiment was a 2x2 design where response speed (fast or slow) and instructions (accuracy-focus or speed-focus) varied.

Participants

Ninety-four native English speakers in the United States were recruited through Mechanical Turk (Mean Age = 34.68, SD = 12.12). Forty-five participants were randomly assigned to the fast response condition, 49 to the slow response condition, 43 (25 fast and 18 slow) to the accuracy-focus response condition, and 51 (20 fast and 31 slow) to the speed-focus response condition.

Materials

The experiment consisted of 56 pairs of words that shared an iconic relationship (e.g., *cup* - *saucer*), that is where one item is usually found either above or below another (see Table 1). These word pairs were extracted from prior research (Louwerse & Jeuniaux, 2010). To reduce the likelihood of participants developing expectations about the experiment, 56 filler items consisted of word pairs without an iconic relation, with half of the pairs having a high semantic association and half having a low semantic association as determined by latent semantic analysis (LSA) (Landauer, McNamara, Dennis, & Kintsch, 2007). LSA is a computational linguistic technique that allows for estimating the relationship between words in a corpus while ignoring word order. Each subject saw half of the critical items in their iconic ordering (e.g., *cup* - *saucer*), and the other half in a reverse iconic order (e.g., *saucer* - *cup*), this was counterbalanced throughout the experiment.

Procedure

As in prior studies, participants were asked to judge the semantic relatedness of word pairs presented on a computer screen. Words were presented one above another in a vertical configuration, with the first word appearing at the top of the screen, and the second at the bottom. Upon presentation of a word pair, participants indicated whether the pair was related in meaning by pressing designated counterbalanced yes or no keys. All word pairs were randomly ordered for each participant to negate any order effects and each trial was separated by a '+' fixation symbol.

Subjects were also instructed to either respond as quickly as possible to the words they saw, or respond as accurately as possible to the words they saw. In the fast response condition, subjects were allotted 1000 ms to respond to the stimuli before a message reading 'TOO SLOW' would appear in the center of the screen. In the slow response condition, subjects were allotted 2500 ms to respond to the stimuli before a message reading 'TOO SLOW' would appear. In the accuracy-focus condition, subjects were asked to try to be as accurate as possible in their responses, whereas in the speed-focus condition, subjects were asked to try to be as quick as possible in their responses. Subjects were asked to describe the directions in a few sentences before beginning the task to ensure understanding. They were also asked to write a few sentences describing what they thought the purpose of the experiment was after they completed the session. There were no participants who misunderstood the directions, nor did any participants guess the true purpose of the experiment.

Results

Four participants were removed from the analysis because more than 30% of their responses were incorrect, as measured by incorrect responses to lexical items. The

Table 2: Mean and *SD* values of RTs for accuracy, speed, fast, and slow conditions.

	Accuracy	Speed
Fast	*M* = 806.27 *SD* = 252.22	*M* = 787.12 *SD* = 258.81
Slow	*M* = 1328.47 *SD* = 473.27	*M* = 1247.89 *SD* = 454.96

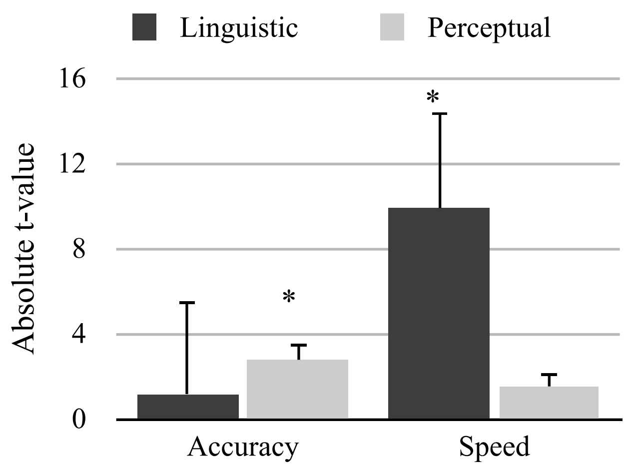

Figure 1: Strength of the effect in absolute t values of perceptual and linguistic factors for accuracy and speed conditions. * denotes p < .05

linguistic and perceptual factors were operationalized as in Louwerse and Jeuniaux (2010), with the linguistic factor being calculated as the log frequency of each word pair, in both orders. 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). The perceptual factor was operationalized as an iconicity rating of the stimulus pairs whereby a set of different participants from the University of Memphis were asked to estimate the likelihood that the word pairs appeared above one another in in the real world. Ratings were made on a scale of 1–6, with 1 being extremely unlikely and 6 being extremely likely.

All analyses were mixed models that specified subjects and items as random factors (Baayen, Davidson, & Bates, 2008) and RT as the dependent variable. To ensure participants correctly performed their task we tested for and found a main effect of speed, $F(1, 9378) = 38.43$, $p < .01$, with faster RTs in the fast time constraint. We also found a main effect of task, $F(1, 9378) = 20.95$, $p < .01$, with faster RTs in condition where subjects were asked to respond as quickly as possible. Furthermore, we found a significant interaction between time constraint speed and task, $F(1, 22.45) = 26.96$, $p < .01$ (see Table 2). These findings suggest that participants were faster to respond to word pairs during a semantic judgment task when a 1000 ms time constraint was imposed than when they were allotted 2500 ms to respond. In addition, subjects were faster to respond to word pairs when they were asked to focus on responding quickly than when they were asked to focus on responding accurately.

To determine if linguistic and/or perceptual factors impacted processing during a semantic judgment task, for all correct critical trials, we ran a linear mixed effect model with word frequency and perceptual ratings as fixed factors and subject and item as random factors. Word frequency best explained resulting RTs, $F(1, 2349) = 26.96$, $p < .01$, although the perceptual factor also contributed (albeit not significantly), $F(1, 2349) = 2.93$, $p = .08$. These findings suggest that the linguistic factor accounts for processing during a semantic judgment task. However, past research found that both linguistic and perceptual factors explained RTs (Louwerse & Jeuniaux, 2010). Perhaps the perceptual factor here failed to reach significance for all subjects due to the fact that many of the subjects were under time constraints.

In fact, when subjects were asked to focus on accuracy, word frequency did not explain RTs $F(1, 1177) = 1.75$, $p =$ 19, while the perceptual factor did, $F(1, 1177) = 8.08$, $p <$. 01. This is in line with the idea that perceptual simulations are more relevant later during processing. However, when subjects were asked to focus on speed, word frequency explained RTs, $F(1, 1169) = 97.42$, $p < .00$, but the perceptual factor did not, $F(1, 1169) = 2.36$, $p = .12$. These results suggest that linguistic factors might play a more important role early, and perceptual information becomes important later in processing.

To determine if this is the case, we conducted a linear mixed effect regression where word frequency and a perceptual factor were fixed factors and subjects and items were random factors for analyses of both the short and long periods of time. When subjects were only given a short time period to respond, word frequency accounted for RTs, $F(1, 1125) = 7.01$, $p < .01$, whereas the perceptual factor remained irrelevant $F(1, 1125) = 0.31$, $p = .57$. In contrast, when subjects were given a longer time period to respond, both factors explained RTs, with word frequency $F(1, 1221) = 5.92$, $p = .02$, and the perceptual ratings $F(1, 1221) = 3.67$, $p =.05$, both reaching significance. These results are in line with prior research that suggest that when participants are given enough time to respond to word pairs in a semantic judgment task, linguistic factors and perceptual factors are relevant for processing, with linguistic representations preceding perceptual representations (as the linguistic factor was significant for the short time period,

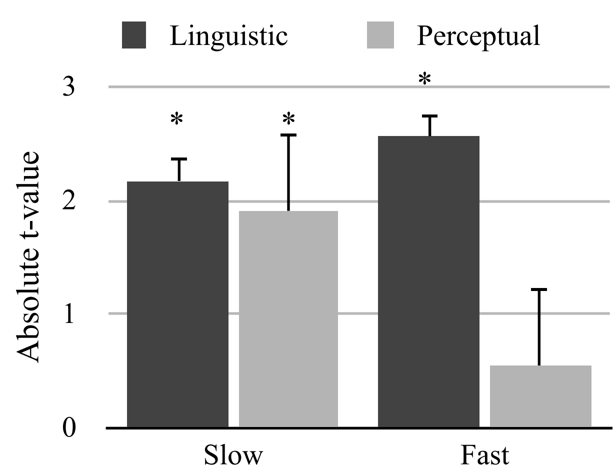

Figure 2: Strength of the effect absolute t values of perceptual and linguistic factors for accuracy and speed conditions. * denotes p < .05

whereas the perceptual factor only became significant for the longer time period).

General Discussion

In the current study our objective was to determine if time constraints on a semantic judgment task could influence how much a subject relied on linguistic and perceptual factors during processing. The Symbol Interdependency Hypothesis predicted that linguistic factors are important immediately during processing, preceding a deeper simulation system. Results from a RT study found exactly that: participants relied more on a linguistic factor during processing when subjects given strict time constraints, or when they were told to focus on responding quickly. When given more time to respond, both linguistic and perceptual factors explained response times. These findings are in line with the findings from Louwerse and Jeuniaux (2010), Louwerse and Connell (2011), and Louwerse and Hutchinson (2012) that suggest the linguistic representations are more relevant early on, and that perceptual representations are more relevant as time progresses.

So it might be the case that RT effects are simply influenced by the amount of time a subject used to respond to a word pair, as time constraints influenced a subject's reliance on perceptual or linguistic factors. Put differently, it seems as if we are using perceptual simulations when we make slower decisions and likewise for linguistic representations and fast decisions. In this study, RT effects of perceptually related pairs were indeed influenced by the timing constraints in the experiment and not simply from the perceptual or linguistic relationships between word pairs. Specifically, subjects were more likely to utilize perceptual information when they had more time to process word pairs on a screen. Conversely, subjects relied more on linguistic information (and less on perceptual information) when they had more stringent timing constraints. In comparison, if it was instead the case that embodied (or linguistic) effects are found only because of an embodied (or linguistic) relationship between word pairs, we would not have expected timing constraints to impact the effect of either perceptual or linguistic representations. Importantly, both factors are relevant throughout the time course of processing word pairs, with the linguistic factor being relatively *more* important early during processing and perceptual information being relatively *more* important later. In fact, the linguistic factor significantly explains RTs in both the slow and fast conditions. Since the linguistic representations precede perceptual simulation during processing (Louwerse & Hutchinson, 2012), it is logical that the linguistic factor would still remain relevant in later processing, with the perceptual factor becoming relatively more relevant.

Importantly, the previous findings that have given support to perceptual simulation (e.g., Barsalou, 1999; Glenberg, 1997; Pecher & Zwaan, 2005; Semin & Smith, 2008; Zwaan, 2004) are not invalidated through these findings. In those instances, perceptual simulation was more suited for the task and presumably time constraint was not a studied factor. However, the results from the current study show that perceptual simulation does not always win the struggle for the most efficient type of processing. The results of the current study show that when people need to be accurate, and have enough time to do so, they will more often rely on perceptual simulation. But that comes at a cost. Intuitively, in order to activate and process all those connected concepts, processing cannot be completed as quickly. However, people rely on the linguistic associations when time is less available. In these instances, utilizing distributional semantics from language statistics is a more efficient route. Symbol interdependency (Louwerse 2007, 2011), argues that these symbols have been encoded with the grounded referents, so there is no need to activate all simulations when one symbol can easily and directly lead to another symbol. It stands to reason that over the course of a person's life where they make these symbolic connections over and over again, they can allow those shortcuts to make the connections for them. These findings support the view that language processing is both linguistic and embodied, with linguistic representations being more prevalent in quick processing and perceptual simulations being more important for more precise information.

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. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645.

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

Brants, T., & Franz, A. (2006). *Web 1T 5-gram Version 1*. Philadelphia, USA: 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.

Fodor, J. (1975). *The language of thought*. New York, USA: Crowell.

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (Eds.). (2007). *Handbook of latent semantic analysis.* Mahwah, USA: Erlbaum.

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211-240.

Louwerse, M. M. (2007). Symbolic or embodied representations: A case for symbol interdependency. In T. Landauer, W. Kintsch, S. Dennis, & D. S. McNamara (Eds.), *Handbook of Latent Semantic Analysis* (pp. 102–120). Mahwah, NJ: Erlbaum.

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

Louwerse, M. M. (2010). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science, 3*, 273–302.
Louwerse, M. M., & Connell, L. (2011). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science, 35*, 381–398.
Louwerse, M. 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.
Louwerse, M. M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. de Vega, A. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning* (pp. 309–326). Oxford, USA: Oxford University Press.
Pecher, D. van Dantzig, S.; Zwaan, R. A.; Zeelenberg, R (2009). Language Comprehenders Retain Implied Shape and Orientation of Objects. *Quarterly Journal of Experimental Psychology*, *62*, 1108-1114.
Pecher, D. Zwaan, R.A. (2005). *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*. Cambridge, USA: Cambridge University Press.
Pylyshyn, Z. (1984). *Computation and cognition: Towards a foundation for cognitive science*. Cambridge, USA: MIT Press.
Semin, G. R., & Smith, E. R. (Eds.). (2008). *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*. New York, USA: Cambridge University Press.
Šetić, M., & Domijan, D. (2007). The influence of vertical spatial orientation on property verification. *Language and Cognitive Processes, 22,* 297–312.
Stanfield, R. A., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science, 12,* 153–156.
Zwaan, R. A., Stanfield, R. A., & Yaxley, R. H. (2002). Do language comprehenders routinely represent the shapes of objects? *Psychological Science, 13*, 168–171.
Zwaan, R. A., & Yaxley, R. H. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin & Review, 10*, 954–958.