

From Head to Toe: Embodiment Through Statistical Linguistic Frequencies

Richard Tillman (r.tillman@memphis.edu)

Department of Psychology / Institute for Intelligent Systems, University of Memphis
400 Innovation Drive, Memphis, TN 38152 USA

Vivek Datla (vvdatla@memphis.edu)

Department of Computer Science / Institute for Intelligent Systems, University of Memphis
400 Innovation Drive, Memphis, TN 38152 USA

Sterling Hutchinson (schthns@memphis.edu)

Department of Psychology / Institute for Intelligent Systems, University of Memphis
400 Innovation Drive, Memphis, TN 38152 USA

Max Louwerse (mlouwerse@memphis.edu)

Department of Psychology / Institute for Intelligent Systems, University of Memphis
400 Innovation Drive, Memphis, TN 38152 USA

Abstract

Recent literature in the cognitive sciences has demonstrated that cognition is fundamentally embodied. For instance, various studies have shown that semantic knowledge about the human body correlates with spatial body representations, suggesting that such knowledge is embodied in nature. An alternative explanation for this finding comes from the Symbol Interdependency Hypothesis, which argues that perceptual information is encoded in language. We demonstrated that the findings that can be explained by an embodied cognition account can also be explained through statistical linguistic frequencies. Co-occurrence frequencies of names for common body parts correlated with experimental findings from adults and children. Moreover, the position of the body parts was predicted on the basis of statistical linguistic frequencies. These findings suggest that language encodes embodied information.

Keywords: embodiment; statistical linguistic frequencies; symbolic cognition; embodied cognition; conceptual processing; symbol interdependency.

Introduction

Over the last decade the notion that cognition is fundamentally embodied has dominated the cognitive sciences (Glenberg, 1997; Goldstone, & Barsalou, 1998; Barsalou, 1999; Lakoff & Johnson, 1999; Zwaan, 2004; Pecher & Zwaan, 2005; Semin & Smith, 2008). The central argument in theories of embodied cognition is that our minds co-evolved with our bodies, especially the sensory motor system, and that cognitive processes therefore heavily rely on perceptual simulations. This argument is in sharp contrast with a computational symbolic approach to cognition. Views of symbolic cognition suggest that meaning can be derived from linguistic context (Landauer & Dumais, 1997). In other words, instead of mental reenactment, mental representations can be seen as internal

structures of symbolic concepts and do not necessarily have a direct relation to perceptual states (Fodor, 1975; Pylyshyn, 1984).

There is a large body of literature that finds evidence that cognition is embodied. Studies have shown that processing within modalities is faster than having to map across modalities (e.g., Marques, 2006; Pecher, Zeelenberg & Barsalou, 2003; Spence, Nicholls & Driver, 2000). Language comprehension seems to be influenced by action representations primed in experimental tasks (e.g., McCloskey, Klatzky, & Pellegrino, 1992; Zwaan, Stanfield & Yaxley, 2002), and visual representations get activated during language comprehension. Perceptual feature characteristics that have affected language comprehension include orientation (Stanfield & Zwaan, 2001), temporality (Zwaan, Madden & Whitten, 2000), visibility (Rapp & Horton, 2003), spatial configuration (Louwerse, 2008; Zwaan & Yaxley, 2003), modality (Louwerse & Connell, 2011; van Dantzig et al., 2008), direction (Glenberg & Kaschak, 2002; Kaschak et al., 2005), or location (Šetić & Domijan, 2007).

Several embodied cognition studies have shown a relation between the meaning of words and their spatial configuration when presented on the screen. For instance, when words for concepts in the air, such as birds and insects, are presented in the upper half of a screen, participants respond faster than when the same words are presented in the bottom of the screen, with a reverse effect for words referring to concepts on land or in the ocean (Šetić & Domijan, 2007; Pecher, Van Dantzig, Boot, Zanolie, & Huber, 2010). Similarly, when word pairs such as *attic* and *basement* are presented vertically, one above the other, iconic pairs are processed faster than reverse iconic pairs, presumably because comprehenders perceptually simulate the position of these concepts (Zwaan & Yaxley, 2003).

Other studies have demonstrated that the vertical configuration of words on the screen and the meaning of those words can be extended to concepts we literally embody, such as body parts. For instance, understanding parts of our body is directly linked to the spatial representation of the human body, and that representation contains veridical information about the relative distance between body parts (Smeets et al., 2009; Struiksma, Noordzij, & Postma, 2011; Van Elk & Blanke, 2011). When participants were presented with combinations of concepts that represent body parts, such as *head-neck*, processing time was considerably faster when the embodied distance of those concepts was small, compared to concepts for which the distance is large, such as *head-toe*. Studies like these yet again show that embodiment explains cognition.

However, the question can be raised as to what extent the relation between body semantics and spatial body representations can only be explained by an embodied cognition account. This is an important question, particularly if other accounts are complementary to the embodied cognition account.

We have argued for one such account in a number of studies. The Symbol Interdependency Hypothesis argues that language comprehension is both perceptual and linguistic in nature (Louwerse, 2008, 2011; Louwerse & Connell, 2011; Louwerse & Jeuniaux, 2010). That is, language comprehension is linguistic through statistical interdependencies between linguistic units and is perceptual through the references linguistic units make to perceptual representations. The Symbol Interdependency Hypothesis thereby makes an important prediction: language has evolved to become a communicative shortcut for language users and it encodes relations in the world. Accordingly, it is hypothesized that the findings attributed to an embodied cognition account can also be explained through statistical linguistic frequencies.

In a number of studies, we have shown that language indeed encodes perceptual information. Louwerse, Cai, Hu, Ventura, and Jeuniaux (2006) and Louwerse and Zwaan (2009) aimed to determine if language encodes geographical information by comparing city latitude/longitude with how often those cities appeared in a corpus. Louwerse, Cai, and Hutchinson (in press) have shown that these predictions are not limited to English, but can also be found in Chinese (predicting cities in China) and Arabic (predicting cities in the Middle East). Louwerse and Benesh (in press) have recently shown how using the *Lord of the Rings* trilogy the longitude and latitude for cities in the fictional Middle Earth can be predicted. The physical distance between cities was accurately estimated based upon statistical linguistic frequencies of cities, thus suggesting that language does encode (perceptual) geographical information.

The encoding of perceptual information in language goes well beyond geography. Louwerse and Connell (2011) have shown that the modality of a word (e.g., *sour*, *soft*, *loud*) can be predicted on the basis of statistical linguistic frequencies. That is, computational estimates on the modality of a word were less precise (visual/tactile, olfactory/taste, auditory) but equally as accurate as human estimates on the modality of words.

In addition to geographical predictions and modality predictions, Louwerse (2008) investigated whether iconicity of words can be predicted. Analogous to binomials such as *top and bottom*, *high and low*, and *up and down*, this study found that the iconic order of concepts such as *flower-stem* could indeed be predicted by simply looking at the order of the words.

It is relevant here to address the question whether these statistical linguistic cues are in fact used by comprehenders. Louwerse (2008) tested whether word pairs like *flower-stem*, presented vertically, yielded faster response times because participants were perceptually simulating the word pairs, or because of the word order (a linguistic factor). The findings demonstrated that the frequency of word pairs such as flower-stem (a perceptually realistic order) is significantly higher than word pairs stem-flower (a perceptually unrealistic order), and that linguistic frequencies explained response times at least as well as perceptual ratings.

The effect of perceptual and linguistic factors on cognitive processes is modulated by stimulus, cognitive task, and by duration of processing. Louwerse and Jeuniaux (2010) showed that linguistic factors best explained semantic judgments of word pairs, whereas perceptual factors best explained iconicity judgments of picture pairs. Furthermore, they concluded that linguistic factors dominated when participants were involved in shallow cognitive processes, and that perceptual factors dominated in deeper cognitive tasks. Louwerse and Connell (2011) extended these findings, showing that faster response times were best explained by linguistic factors, and slower response times were best explained by perceptual factors. These findings suggest that the relative employment of linguistic or perceptual representations changed as a function of the task, duration of the task, or stimulus.

In the following study, we determined whether embodied information – information about the distance between body parts – was also encoded in language. To test for this possibility, we conducted a computational linguistic study in which we calculated the co-occurrence of body part names and compared the statistical linguistic frequencies with the existing experimental data. We thereby hypothesized that body parts that are perceptually close together are placed in similar linguistic contexts, thereby allowing for accurate computational estimates on the position of the body part.

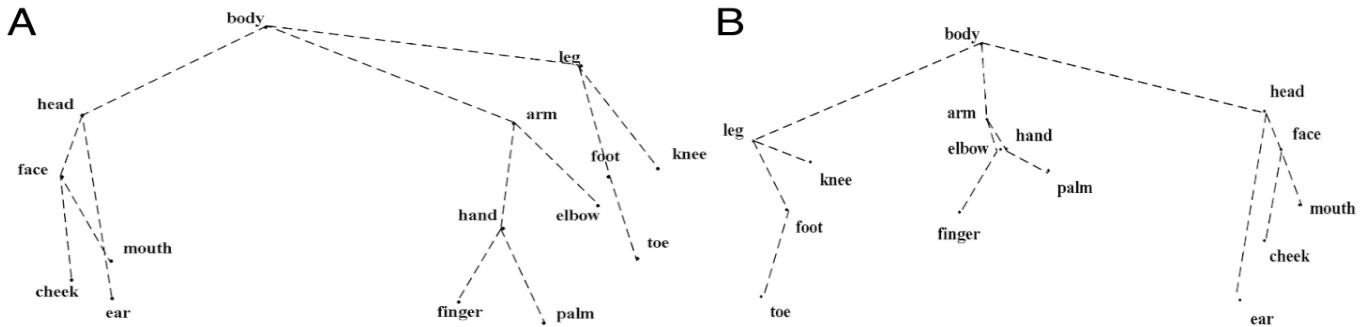


Figure 1. Body part similarity ratings of adults (A) and children (B) (Jacobowitz, 1973).

Study 1

In previous research, Jacobowitz (1973) explored the development of language by comparing body part similarity ratings of five-year-old children, and adults. The 15 body parts used were: *Arm, body, cheek, ear, elbow, face, finger, foot, head, hand, knee, leg, mouth, palm, and toe*. Jacobowitz conducted four replicated multi-dimensional scaling analyses (RMDS), which simultaneously analyzed multiple matrices. The dimensional scaling illustrated that the five-year-olds grouped the head items, arm items, and leg items more similarly (see Figure 1B). The adults, on the other hand, grouped head terms together, but the other extremities were grouped by function (e.g., arm and leg (limbs) were more similar, finger and toe (digits) were more similar, etc.) (see Figure 1A). Jacobowitz found that the similarity ratings for body parts were hierarchical for both the children and adults.

In the current study, Jacobowitz's (1973) data was compared with findings from statistical linguistic frequencies. We calculated the frequency of first-order co-occurrences in the *Web 1T 5-gram* corpus (Brants & Franz, 2006). This corpus consists of one trillion word tokens (13,588,391 word types) from 95,119,665,584 sentences. The volume of the corpus allows for an extensive analysis of patterns in the English language. The frequency of co-occurrences of the 15 words was computed in bigrams, trigrams, 4-grams and 5-grams. For instance, the frequency of the words *{head, toe}* was determined by considering these words next to one another *{head, toe}*, with one word in between *{head w1 toe}*, with two *{head w1 w2 toe}* or with three intervening words *{head w1 w2 w3 toe}*. This method is identical to the one used in Louwerse (2008), Louwerse and Jeuniaux (2010), and Louwerse and Connell (2011).

The result of these computations was a 15 x 15 matrix of raw frequencies of co-occurrences, from which log frequencies were obtained. This matrix was submitted to an MDS analysis using the ALSCAL algorithm (see Young, Takane, & Lewyckyj, 1978). For purposes of mapping the relative location of body parts, it is insufficient to simply obtain the co-occurrence frequencies in the Google corpus. The frequencies must be converted to x and y coordinates, and then a mathematical analysis performed to find the relative spatial location of the body parts. Multidimensional scaling (MDS) is a series of mathematical operations that can illuminate patterns within data that may not be

immediately recognizable with standard numerical output (Kruskal, & Wish, 1977; Blake, Schulze, & Hughes, 2003). MDS has been utilized to not only analyze similarity, but also to provide a graphical representation of those similarities. A Euclidean distance measure transformed the semantic similarities into dissimilarities, such that the higher the value, the longer the distance. Default MDS criteria were used with an S-stress convergence of .001, a minimum stress value of .005, and a maximum of 30 iterations. The fitting on a two-dimensional scale was moderate, with a Stress value = .21 and an $R^2 = .86$.

To do justice to the geometry of the 2D variables in Jacobowitz (1973), we used bidimensional regression analyses to compare the participants' estimates with the actual coordinates of the body parts. Tobler (1994) and Friedman and Kohler (2003) introduced bidimensional regressions in order to compute the mapping of any two planes under consideration. Whereas in a unidimensional regression each data point is shifted by intercept and slope, each actual and predicted value of the dependent variable are presented by a point in space, whereby vectors represent intercept and slope.

A bidimensional regression yielded a significant correlation between the frequency estimates and Jacobowitz's (1973) loadings on a two-dimensional plane for both the adult study, $r = .66$, $p < .01$, $n = 15$, and the child study, $r = .63$, $p = .01$, $n = 15$. To ascertain that these findings could not be attributed to accidental pairings of coordinates, we conducted a Monte Carlo simulation, randomly sampling each dataset 1000 times. The findings solidified the results, with no bidimensional relation between the statistical linguistic frequencies and Jacobowitz's (1973) adult data, average $r = .23$ ($SD = .12$), $n = 15$ or child data, average $r = .24$ ($SD = .12$), $n = 15$. These findings suggest that statistical linguistic frequencies can explain data obtained from human participants.

In addition to the comparison between Jacobowitz's (1973) two-dimensional fitting, we compared a one-dimensional solution, using the first dimension of the MDS solution, with the location of the body part terms. The correlation between the location of the body part words and the computational estimates was again high, $r = .6$, $p < .001$, $n = 15$. The linear fitting between the computational estimates and the actual position is presented in Figure 2.

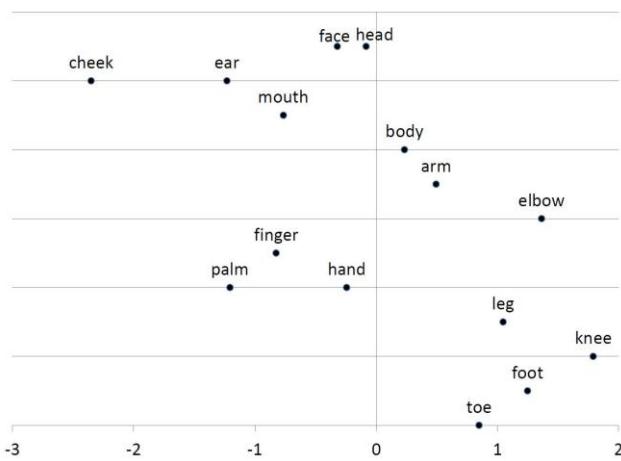


Figure 2. Multidimensional Scaling of 15 body parts from Jacobowitz (1973).

Study 2

Van Elk and Blanke (2011) established a relationship for spatial position of body parts as well as the relative distance between them for native French speakers (see Table 1). In experiment 1, 38 body parts were assigned to nine categories dependent upon the distance from each other on the body (e.g., forehead/toe = 9; forehead/eye = 1). The words were then presented vertically in a congruent or incongruent spatial position (forehead/toe; toe/forehead). Subjects demonstrated increased RTs for larger distances, while position congruency did not seem to have an effect. Experiment 2 consisted of an iconicity judgment also using relative distance and congruency. However, in this experiment the words were not in the center of the screen as in Experiment 1, but arranged in varying distances from each other. There were significant main effects, as well as

an interaction, for the error rates. The RTs revealed there were main effects of congruency and distance, but no interaction was found.

We computed the log frequency of all combinations of the English body part words and compared the data with the Van Elk and Blanke (2011) distance data. Because the algorithm functions best with single words in a 2-5 gram window, we removed all words that require two words in English (*under arm*, *ring finger*, *index finger*, and *middle finger*). Moreover, no frequencies were found for *instep* and *pinkie* combinations, these words were removed from the analysis.

The correlation of the 32×32 word pair frequencies and the distances was significant, $r = .35$, $p < .001$, $n = 1024$, with higher frequencies yielding lower physical distances. This finding suggests that embodiment is encoded in language, such that the relative location of body parts can be estimated using statistical linguistic frequencies.

Next, we conducted analyses similar to the first study, whereby we did not use the raw frequency comparisons, but instead entered the $n \times n$ matrix in an MDS algorithm and used the loadings of the body parts names as a comparison. To do justice to the one-dimensional plane Van Elk and Blanke (2011) used, the MDS solution was restricted to a one-dimensional solution. The fitting was moderate, *Stress* = .47, $R^2 = .50$. When the loadings of the 32 body parts were compared with their physical distances, a strong correlation was found, $r = -.76$, $p < .001$, $n = 32$.

To determine whether these findings could in any way be attributed to accidental pairings of variables, we again conducted a Monte Carlo simulation, whereby correlations of the 1000 randomizations of the data were computed. The average correlation did not come close to the correlation obtained for the actual data, average $r = .15$, $p = .41$, $n = 1000$. As before, we plotted the position of the body parts and their corresponding words (Figure 3).

Table 1: Body part positions and categories (Van Elk & Blanke, 2011).

Word	Position	Loading	Cat.	Word	Position	Loading	Cat.	Word	Position	Loading	Cat.
hair	1	0.79	1	back	3	-1.15	10	palm	6	1.05	13
eye	1	1.01	4	shoulder	3	-0.42	9	thigh	7	-0.83	17
ear	1	1.13	4	chest	3	0.31	10	leg	7	-0.56	18
forehead	1	1.35	2	elbow	4	-0.81	11	knee	8	-0.95	19
eyebrow	1	2.09	3	wrist	5	-0.82	13	calf	8	-0.88	20
neck	2	-0.14	8	forearm	5	-0.75	12	ankle	9	-1.36	21
throat	2	0.91	8	butt	5	-0.58	16	shin	9	-1.24	20
chin	2	0.92	7	thumb	5	0.13	14	heel	9	-1.09	22
nose	2	1.24	5	stomach	5	0.86	15	foot	10	-0.9	21
lip	2	1.35	6	hand	6	-0.89	13	toe	10	-0.61	23
cheek	2	1.56	5	hip	6	-0.7	16	palm	6	1.05	13

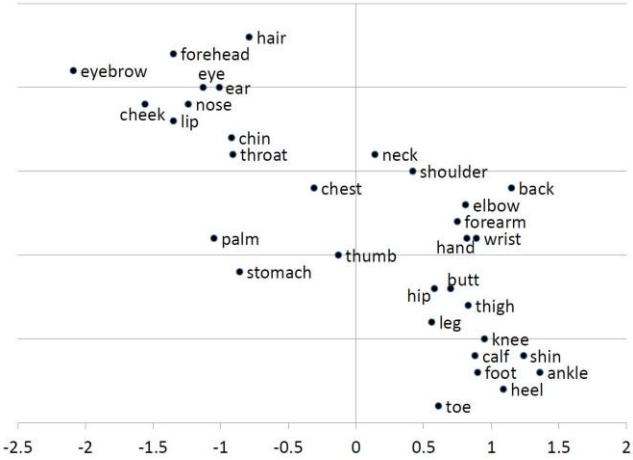


Figure 3. Multidimensional Scaling of 32 body parts (Van Elk & Blanke, 2011).

Discussion

Recent literature has shown that perceptual information, such as geographical locations, modalities, and iconicity, is encoded in language. The current paper extended these findings by addressing the question whether language encodes (literally) embodied information: whether statistical linguistic frequencies can explain the relative location of different parts of the body. Results from two computational studies showed that such frequencies indeed can estimate the relative location of body parts. In study 1, we demonstrated that computationally derived values can explain human similarity estimates of body-parts. Study 2 similarly found that word frequencies can estimate physical distances between body parts. Both of these studies support the claim that language encodes body information.

We conclude that language inherently contains body part information, such that experimental results can be approximated computationally. This is in line with previous research that has demonstrated that language encodes geographical information (Louwerse & Zwaan, 2009), that it encodes modality specific information (Louwerse & Connell, 2011), spatial information (Louwerse, 2008), and social relations (Hutchinson, Datla, & Louwerse, in press). The current study adds to this literature and suggests that cognition is indeed both embodied and symbolic.

References

Barsalou, L. (1999). Perceptual symbol systems. *Behavior and Brain Sciences*, 22, 557-660.

Blake, B., Schulze, S., & Hughes, J. (2003). *Perceptual mapping by multidimensional scaling: A step by step primer*. Research Reports in Consumer Behavior. Cleveland, OH: Cleveland State University.

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

Fodor, J. (1975). *The Language of Thought*. New York, NY: Crowell.

Friedman, A. & Kohler, B. (2003). Bidimensional regression: A method for assessing the configural similarity of cognitive maps and other two-dimensional data. *Psychological Methods*, 8, 468-491.

Glenberg, A. (1997). What memory is for. *Behavior and Brain Sciences*, 20, 1-55.

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

Goldstone, R., & Barsalou, L. (1998). Reuniting perception and conception. *Cognition*, 65, 231-262.

Hutchinson, S., Datla, V., & Louwerse, M. M. (in press). Social networks are encoded in language. *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Jacobowitz, D. (1973). *Development of semantic structures*. Unpublished dissertation. University of North Carolina-Chapel Hill, NC.

Kaschak, M., Madden, C., Therriault, D., Yaxley, R., Aveyard, M., Blanchard, A., & Zwaan, R. (2005). Perception of motion affects language processing. *Cognition*, 94, B79-89.

McCloskey, B., Klatzky, R. L., & Pellegrino, J. (1992). On rubbing your stomach while tapping your fingers: Interference between motor planning and semantic judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 948-961.

Kruskal, J. & Wish, M. (1977). *Multidimensional Scaling*. Beverly Hills, CA: Sage Publications.

Lakoff, G. & Johnson, M. (1999) *Philosophy in the Flesh*. New York: Basic Books.

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

Louwerse, M. (2008). Embodied representations are encoded in language. *Psychonomic Bulletin and Review*, 15, 838-844.

Louwerse, M. & Benesh, N. (in press). Representing spatial structure through maps and language: Lord of the Rings encodes the spatial structure of Middle Earth. *Cognitive Science*.

Louwerse, M., Cai, Z., Hu, X., Ventura, M., & Jeuniaux, P. (2006). Cognitively inspired natural language based knowledge representations: Further explorations of Latent Semantic Analysis. *International Journal of Artificial Intelligence Tools*, 15, 1021-1039.

Louwerse, M., Cai, Z., & Hutchinson, S. (in press). The Chinese route argument: Predicting the longitude and latitude of cities in China and the Middle East using statistical linguistic frequencies. *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Louwerse, M., & Connell, L. (2011). A taste of words: Linguistic context and perceptual simulation predict the

modality of words. *Cognitive Science*, 35, 381-398.

Louwerve, M. & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. de Vega, A. Glenberg, & C. Graesser (Eds.). *Symbols, embodiment, and meaning* (pp. 309-326). Oxford University Press: New York, NY.

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

Louwerve, M. & Zwaan, R. (2009). Language encodes geographical information. *Cognitive Science*, 33, 51-73.

Marques, J. (2006). Specialization and semantic organization: Evidence for multiple semantics linked to sensory modalities. *Memory & Cognition*, 34, 60-67.

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 Cognition*, 1, 1-8.

Pecher, D., van Dantzig, S., Zwaan, R., & Zeelenberg, R. (2009). Language comprehenders retain implied shape and orientation of objects. *The Quarterly Journal of Experimental Psychology*, 62(6), 1108-14.

Pecher, D., & Zwaan, R. (Eds.). (2005). *Grounding cognition: The role of perception and action in memory, language, and thinking*. New York: Cambridge University Press.

Pylyshyn, Z. (1984). *Computation and cognition: Towards a foundation for cognitive science*. Cambridge, MA: MIT Press.

Horton, W. & Rapp, D. (2003). Out of sight, out of mind: Occlusion and the accessibility of information in narrative comprehension. *Psychonomic Bulletin & Review*, 10, 104-110.

Semin, G., & Smith, E. (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.

Smeets, M., Klugkist, I., van Rooden, S., Anema, H., & Postma, A. (2009). Mental body distance comparison: a tool for assessing clinical disturbances in visual body image. *Acta Psychologica*, 32, 157-165.

Spence, C., Nicholls, M., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, 63, 330-336.

Stanfield, R., & Zwaan, R. A. (2001). The effect of implied orientation derived from verbal context on picture recognition. *Psychological Science*, 12, 153-156.

Struiksma, M., Noordzij, M., & Postma, A. (2011). Reference frame acceptability in haptics differs for the blind and sighted in the horizontal but not in the vertical plane. *Perception*, 40, 725-738.

Tobler, W. (1994). Bidimensional regression. *Geographical Analysis*, 26, 186-212.

Van Dantzig, S., Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2008). Perceptual processing affects conceptual processing. *Cognitive Science*, 32, 579-590.

Van Elk, M. & Blanke, O. (2011). The relation between body semantics and spatial body representations. *Acta Psychologica*, 138, 347-358.

Young, F., Takane, Y., & Lewyckyj, R. (1978). ALSCAL: A nonmetric multidimensional scaling program with several different options. *Behavioral Research Methods and Instrumentation*, 10, 451-453.

Zwaan R. A. (2004). The immersed experience: toward an embodied theory of language comprehension. In B. H. Ross, Ed. *The Psychology of Language and Motivation*, 44. New York, NY: Academic Press.

Zwaan R., Madden C., & Whitten S. (2000) The presence of an event in the narrated situation affects its activation. *Memory and Cognition*, 28, 1022-28

Zwaan, R., Stanfield, R., & Yaxley, R. (2002). Language Comprehenders Mentally Represent the Shapes of Objects. *Psychological Science*, 13(2), 168-171.

Zwaan, R., & Yaxley, R. (2003). Spatial Iconicity Affects Semantic Relatedness Judgments. *Psychonomic Bulletin and Review*, 10(4), 954-8.