Title
Geographical Estimates are Explained by Perceptual Simulation and Language Statistics

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

ISSN
1069-7977

Authors
Tillman, Richard
Hutchinson, Sterling
Louwerse, Max

Publication Date
2013

Peer reviewed

Geographical Estimates are Explained by Perceptual Simulation and Language Statistics

Richard Tillman (rntllman@memphis.edu)
Department of Psychology/ Institute for Intelligent Systems, University of Memphis
365 Innovation Drive, Memphis, TN 38152 USA

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

Several studies have demonstrated that language encodes geographical information. That is, the relative longitude and latitude of city locations can be extracted from language. Whether people actually rely on these linguistic features is less clear. Recent studies have suggested that language statistics plays a role in geographical estimates, but these studies rely on map drawings, a fundamentally perceptual task. The current study investigated the extent to which people rely on map representations and statistical linguistic frequencies by using a linguistic task. Participants saw U.S. city pairs in their iconic positions (a more northern city is presented above a more southern city, or a more western city is presented to the left of a more eastern city), and in their reverse-iconic positions (a more southern city is presented above a more northern city, or a more eastern city is presented to the left of a more western city). For iconic city pairs both in the east – west (*Seattle – Boston*) and north – south (*Memphis – Miami*) configurations, RTs were determined by the iconicity. No effect was obtained for statistical linguistic frequencies. However, when city pairs were presented in a reverse-iconic configuration, for both horizontal (*Boston – Seattle*) and vertical (*Miami – Memphis*) orientations, both perceptual and linguistic factors explained RTs. These findings support the idea that cognition relies on a shallow heuristic, a linguistic system, and a fine-grained and more precise perceptual simulation system.

Keywords: embodied cognition; symbolic cognition; geography; spatial cognition

Introduction

Is San Francisco close to New York? Is Boston close to Miami? Judging the distance between cities can be approached in more than one way. This judgment can be deep and precise, as with perceptual simulation, or quick and shallow, as with symbolic representation. For instance, humans can make geographical estimates on the basis of their perceptual experiences from locomotion and stationary viewing, from static pictorial representations, such as diagrams, paintings and photos, provided on a map, and they can acquire information via dynamic pictorial representations, including animations, and videos (Freundschuh & Mercer, 1995).

The importance of a perceptual simulation system has been strongly advocated by accounts of embodied cognition (Barsalou, 1999; Barsalou, 2008; Glenberg & Kaschak, 2002; Pecher & Zwaan, 2005; Semin & Smith, 2008). According to Barsalou, Solomon, and Wu (1999), perceptual states are transferred into memory and function symbolically, rather than an arbitrary representation such as language. As an example, overwhelming evidence in favor of an embodied cognition account has accumulated, showing that processing within modalities is faster than having to map across modalities, and suggesting that modality switching comes at a price (e.g., Marques, 2006; Pecher, Zeelenberg, & Barsalou, 2003; Spence, Nicholls, & Driver, 2001). Furthermore, language comprehension seems to be influenced by action representations primed in experimental tasks (e.g., Glenberg & Kaschak, 2002; Kaschak et al., 2005; Klatzky, Pellegrino, McCloskey, & Doherty, 1989; Zwaan, Stanfield, & Yaxley, 2002), and visual representations get activated during language comprehension (see also Boroditsky, 2000; Fincher-Kiefer, 2001; Matlock, Ramscar, & Boroditsky, 2005).

One particular study nicely illustrates the embodied cognition account. Zwaan and Yaxley (2003) presented iconic word pairs either as they occur in the real world, such as *attic* over *basement*, or the reverse-iconic orientation, such as *basement* over *attic.* They found significant differences between the iconic and reverse-iconic configurations of these word pairs. They concluded that the explanation for the iconicity effect was that words activate their perceptual representations (attics presented above basements are processed faster than basements above attics, because of their iconic relationship in the real world).

Louwerse (2008) questioned whether the Zwaan and Yaxley (2003) finding should be solely attributed to perceptual simulation. Statistical linguistic frequencies, the co-occurrence of words in a given frame, showed that items that are normally high in space preceded items that are normally low in space more frequently than vice versa, suggesting that language encodes spatial information (e.g., we say *up and down*, *top and bottom*, *knees and toes*, rather than *down and up*, *bottom and top* and *toes and knees*). Moreover, statistical linguistic frequencies explained RTs better than the perceptual factor. These findings demonstrate that there is a complementary linguistic explanation to a perceptual simulation explanation.

Louwerse and Jeuniaux (2010) showed that the extent to which cognitive processes can be explained by perceptual simulation or language statistics (frequency of word co-occurrence) depends on a variety of factors, including the nature of the stimulus (e.g., words versus pictures) and the cognitive task (e.g., shallow or deep cognitive task). In Louwerse and Jeuniaux (2010), participants saw either pictures or words in their natural orientation (e.g., *ceiling* above *floor*), or in their reverse orientation (e.g., *floor* above *ceiling*). Statistical linguistic frequencies were better able to explain RTs than perceptual ratings when the word pairs were used, with the reverse result when picture pairs were used. Similarly, when participants were asked to make a real-world judgment task, the effect for perceptual ratings on RTs was larger than that for statistical linguistic frequencies, with the opposite result for a semantic judgment task. Importantly, effects for both language statistics and perceptual simulation were found for both stimulus types and both cognitive tasks, however, their relative dominance was modified by task and stimulus.

These findings have been captured through the Symbol Interdependency Hypothesis, which proposed that conceptual processing can be explained by both symbol and embodied mechanisms (Louwerse, 2007; 2008; 2011). When we encounter a word, a rough meaning is elicited by using the linguistic, that is symbolic, neighbors. This is accomplished by using language statistics, where words that often appear together are related in important ways that can facilitate initial cognitive processing. In order to fully ground the word, we can mentally simulate the features of the word in order to process the word in a deeper way. Human beings can use the fuzzy sense of words by a linguistic (symbolic) short-cut when processing language as it occurs. In addition, language is encoded with sensori-motor and spatial information. The Symbol Interdependency Hypothesis is composed of three components. First, language encodes perceptual information. Second, during cognitive processes users of language rely on language statistics and perceptual simulation. Finally, the dominance of either language statistics or perceptual simulation is dependent on the type of task and stimulus.

Do these three claims also hold for spatial cognition within geographical representation? Using newspapers such as the *New York Times* and the *Wall Street Journal* Louwerse and Zwaan (2009) were able to estimate the longitude and latitude of the largest cities in the US computationally, based on the idea that "cities that are located together are debated together." That is, by computing the *n* x *n* frequencies of the co-occurrence of city names in the newspapers, a two-dimensional multidimensional scaling analysis yielded correlations with the longitude and latitude of the cities. The Louwerse and Zwaan (2009) findings are not limited to the English language. Louwerse, Hutchinson, and Cai (2012) found similar results using Arabic for predicting cities in the Middle East, and Chinese for predicting cities in China. It is interesting to note the presence of this effect was found for three languages each with different writing directions (English- left to right, Arabic- right to left, and Chinese, at least historically- top to bottom). This shows, at the least, that it is possible to map out cities in different locations, within different writing systems, by using the frequency of co-occurrences of city names within a large corpus.

Language encodes geographical information. The question is whether this also means that humans use these encodings. Louwerse and Zwaan (2009) stated that between 16% and 35% of the latitude and longitude variance in human location estimates can be attributed to linguistic coding. These percentages were found by using a bidimensional regression analysis correlating human and computational longitude and latitude estimates (by a large newspaper corpus). However, it is unclear whether 84% and 65% and of the latitude and longitude variance in human location estimates can be attributed to spatial information. Moreover, given that language encodes spatial information, it is difficult to disentangle linguistic and perceptual processes. It could be argued that proximity can explain estimation bias when determining distance between two locations (Tobler, 1970). However, Friedman, Kirkman, & Brown (2002) tested this hypothesis by comparing latitude estimates by participants in Canada and Texas. Their findings did not support the proximity hypothesis, whereas participants in Texas exhibited greater bias in their estimates of Mexican locations than the participants from Canada. The explanations proposed by Friedman et al. included cognitively based beliefs, geopolitically based beliefs, and socio-culturally based beliefs. It was also argued by Brown (2002) that seeding effects can affect real-world judgments, such as proximity and size estimation of two cities. However, many of the experiments contained in Brown (2002) were designed for numerical estimates such as population, or how many square kilometers is for a given country. While they were robust and interesting effects, they do not necessarily apply here, because the tasks in the present study utilize the distance between two cities, not estimations of numbers about those locations.

Louwerse and Benesh (2012) investigated to what extent geographical estimates could come from language statistics and from perceptual simulations by comparing readers who had read Tolkien's *Lord of the Rings* trilogy and *The Hobbit* with participants who studied a map and had never seen the text. As in Louwerse and Zwaan (2009), computational estimates of co-occurrence of the location of the cities in Middle Earth were determined. Participants were asked to draw the location of the cities on a piece of paper. Again, computational estimates of co-occurrence for cities mentioned in the text correlated with the longitude and latitude of cities in Middle Earth. Interestingly, estimates from those who studied a map correlated with the actual geographical location in Middle Earth more than the estimates from those who had read the text did. On the other hand, estimates from those who had read the text correlated more with the computational estimates of co-occurrence than the estimates from those who studied a map did. These results support the claims made by the Symbol Interdependency Hypothesis: 1) Language (*Lord of the Rings*) encodes geographical (Middle Earth) information; 2) Those who read *Lord of the Rings* and those who studied the map relied both on language statistics and perceptual simulation in their estimates; 3) the relative dominance of language statistics and perceptual simulation factors is modified by whether participants read the text or studied the map.

Importantly, human estimates in Louwerse and Zwaan (2009) and Louwerse and Benesh (2012) were derived from an experimental setting in which participants were asked to draw the location of cities on a piece of paper, which is a perceptual task. Given that the cognitive task determines the effect of language statistics and perceptual simulations (Louwerse & Jeuniaux, 2010), the estimates how much of human geographical estimates come from language statistics and come from perceptual simulations is likely to be biased.

We therefore conducted an experiment in which participants were not asked to draw a map (a perceptual task) but to estimate geographical distances from words (a task that better justifies linguistic processing).

Experiment

In a between subjects design, participants viewed United States city pairs in either a horizontal or vertical orientation. These city pairs randomly appeared in either their natural orientation (i.e., a more northern city was presented above a second city, or a more western city was presented to the left of a second city), or the opposite of their natural orientation. In this iconic orientation, we predicted that participants would rely on perceptual information. Conversely, when the location of the city pairs was reversed (i.e., reverse-iconic), we predicted that participants would rely on language statistics.

Methods

Participants Ninety-three undergraduate native English speakers at the University of Memphis (67 females) participated for extra credit in a Psychology course. Forty-five participants were randomly assigned to the vertical presentation condition and forty-eight participants were randomly assigned to the horizontal presentation condition.

Materials The experiment consisted of the largest 50 cities in the United States using the U.S. Census data from 2000 and were presented in 2,450 name pairs.

Procedure In two presentation conditions (horizontal or vertical), we presented subjects with city pairs in their iconic configuration and their reverse iconic configuration. Participants were randomly assigned to view either the vertical or horizontal configuration. To reduce order effects, participants were counterbalanced across four groups per condition.

The city pairs were presented on a 1280x1024 computer screen. Participants were asked whether the named United States cities were closely located. The vagueness of the question intentionally left open the question of closeness for the participant to decide. A more specific question would have added a number of constraints that would influence the judgment in unintended ways. The center of the screen was positioned at eye level. Each trial began with the presentation of a fixation cross for 3000ms. The participants would select their choice (yes or no) by designated buttons on a keyboard then a fixation cross would appear on the screen for the next trial.

Results

Outliers were defined as response times (RTs) that were 2.5 SD above the mean per subject per condition and were removed from the analysis. This affected less than 5% of the data.

The perceptual factor was operationalized as the differences in latitude or longitude of the cities. Language statistics was operationalized as the log frequency of $a - b$ (e.g., for North – South: *New York – Miami*; for East – West: *Los Angeles – Boston*), or $b - a$ (e.g., for North – South: *Miami – New York*; for West – East: *Boston – Los Angeles*) order of word pairs using the large Web 1T 5-gram corpus (Brants & Franz, 2006). This corpus consists of 1 trillion word tokens (13,588,391 word types) from 95,119,665,584 sentences. Using the log frequency of the co-occurrence of word pairs enables linear regressions to be performed comparing frequencies with other types of data, because raw frequencies of those co-occurrences are extremely skewed (Gries, 2010).

A mixed-effect regression analysis was conducted on RTs with linguistic frequency and the perceptual factor as fixed factors 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). Participants and items were treated as random factors in the analysis.

Note that the strength of a model association is represented as a weighted ratio of the F statistic. R^2 and F used in ordinary regression analysis are closely related, since where k is the number of model parameters and N is the number of cases, such that *F* has (*k*, *N* - *k* - 1) df. See also Pedhazur (1997, p. 105) and Louwerse and Jeuniaux (2010). See Figure 1.

$$F = \frac{\frac{R^2}{k}}{\frac{1 - R^2}{N - k - 1}}$$

Figure 1. Weighted ratio of the *F* statistic.

Vertical Configuration The perceptual factor explained RTs in the iconic pairs, $F(1,964.821) = 17.7$, $p < .001$, with larger distances yielding lower RTs. The linguistic factor, however, did not explain RTs for the iconic word pairs, $F(1,960.549) = 0.45$, $p = .50$.

For the reverse iconic configuration the perceptual factor also explained RTs, $F(1,984.502) = 8.382$, $p = .004$, except that the effect was considerably smaller. Importantly, for these reverse-iconic word pairs a significant effect on RTs was obtained for the linguistic factor, $F(1,970.543) = 6.18$, $p = .013$, with higher frequencies yielding lower RTs. Figure 2 gives an estimate of effect sizes, which are calculated by differences between groups as opposed to within the two original groups.

Horizontal Configuration For the horizontal configuration, a similar pattern emerged as for the vertical configuration. That is, the perceptual factor explained RTs for city pairs in their iconic order, $F(1,962.735) = 9.645$, $p < .002$, but no significance found for language statistics when the position of the city pair was in the iconic order, $F(1,995.626) = 1.254$, $p = .263$.

For the reverse-iconic order the perceptual factor again explained RT, $F(1,987.520) = 9.565$, $p = .002$. Importantly, an effect for language statistics was obtained when city pairs were presented in their reverse-iconic order, $F(1,1012.479) = 4.068$, $p = .044$ (Figure 3).

Discussion

The goal of the present study was to determine to what extent humans rely on language statistics and on perceptual simulation in spatial cognition. Previous work has found that language encodes geographical information, so much so that by computing the rates of co-occurrence of city names in the text, multidimensional scaling techniques allow for estimating the relative longitude and latitude of cities. Experiments have shown that humans rely on perceptual simulation, for instance, a perceptually grounded memory of the text. However, there is also evidence humans rely on language statistics, similar to those obtained from computational estimates. Because the existing literature used human estimates from map drawings, the current paper investigated to what extent linguistic and perceptual factors would affect cognitive processes in a more linguistic task.

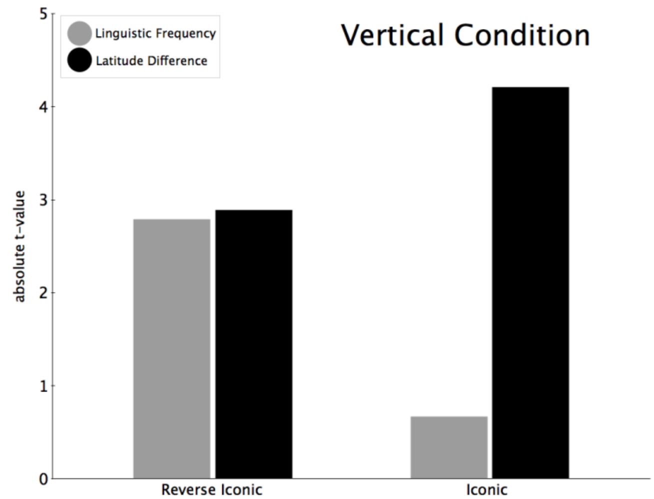

Figure 2. Absolute *t*-values of the linguistic frequency and latitude differences in reverse-iconic and iconic orientation in the vertically positioned city names.

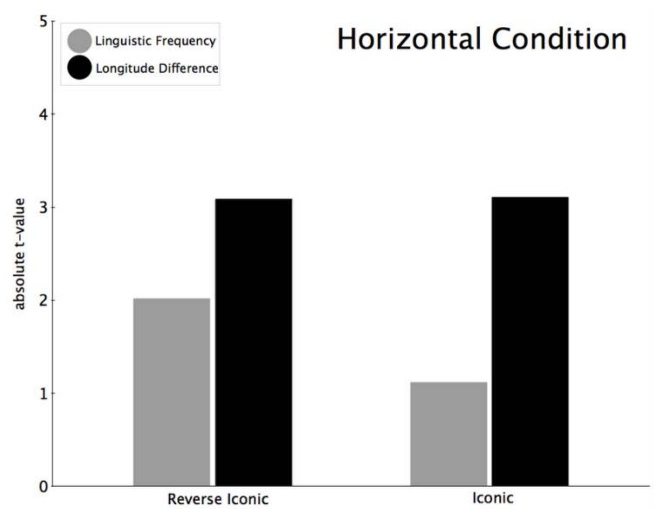

Figure 3. Absolute *t*-values of the linguistic frequency and longitude differences in reverse-iconic and iconic orientation in the horizontally positioned city names.

When city pairs were presented to participants in their iconic order, their distance best explained RTs. The larger the distance, the larger the RTs. No effect was obtained for language statistics in the iconic order. For the reverse-iconic order, the perceptual factor again explained RTs, but language statistics did so as well. This suggests that when the task or the stimulus invites for perceptual simulation, humans rely on perceptual simulation. When perceptual simulation is harder, other heuristics, such as language statistics are used. This finding lies fully in line with the results obtained by Louwerse and Jeuniaux (2010) showing that linguistic and perceptual factors dominate in conceptual processing when they are relevant.

Further research should investigate the weaker effects for the horizontal condition compared to those for the vertical condition. Barsalou (2008) argues that locating objects on a left/right axis is more difficult possibly due to the symmetry of the body and less salient cues to differentiate those objects. Perhaps this weaker effect is due to embodiment factors. However, this difference might also be explained by linguistic factors. When reporting two spatially related words in English, such as up-down or left-right, the top or the left most word is most often reported first. There is the possibility that there are less instances of the left-right phenomenon found in language. Future study of the nature of this phenomenon could illuminate why this weaker effect has been found. In the past, it has been shown that the linguistic system is used more often when quick decisions are made, and the perceptual system is used when slower decisions are made (Louwerse & Connell, 2011). However, more specific investigation is recommended in the future as to the exact mechanisms of these speed differences and to what degree they affect decisions.

These findings reported in this paper are also in line with the Symbol Interdependency Hypothesis, which claims that cognitive processes rely both on language statistics and perceptual simulation. Because language encodes spatial information, including geographical information, language users can utilize these cues in their comprehension process. Geographical judgments then rely on both a shallow heuristic, called the linguistic system, and a fine-grained and more precise perceptual simulation system.

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*(4), 390–412.

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

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

Barsalou, L. W., Solomon, K. O., & Wu, L. L. (1999). Perceptual simulation in conceptual tasks. In M. K. Hiraga, C. Sinha, & S. Wilcox (Eds.), *Cultural, typological, and psychological perspectives in cognitive linguistics: The proceedings of the 4th conference of the International Cognitive Linguistics Association, Vol. 3*. Amsterdam: John Benjamins.

Boroditsky, L. (2000). Metaphoric structuring: Understanding time through spatial metaphors. *Cognition, 75*(1), 1-28.

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

Brown, N. R. (2002). Real-world estimation: Estimation modes and seeding effects. *Psychology of Learning and Motivation, 41*, 321-359.

Fincher-Kiefer, R., (2001). Perceptual components of situation models. *Memory & Cognition, 29*, 336-343.

Freundschuh, S. M., & Mercer, D. (1995). Spatial cognitive representations of story worlds acquired from maps and narrative. *Geographical Systems, 2*, 217–233.

Friedman, A., Kirkman, D. D., & Brown, N. (2002). Spatial location judgments: A cross-national comparison of estimation bias in subjective North American geography. *Psychonomic Bulletin & Review, 9*, 615–623.

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

Gries, S. Th. (2010). Useful statistics for corpus linguistics. In Aquilino Sánchez & Moisés Almela (Eds.), *A mosaic of corpus linguistics: selected approaches*. Frankfurt am Main: Peter Lang.

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.

Klatzky, R. L., Pellegrino, J. W., McCloskey, B. P., & Doherty, S. (1989). Can you squeeze a tomato? The role of motor representations in semantic sensibility judgments. *Journal of Memory and Language, 28*, 56–77.

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

Louwerse, M. M. (2007). Symbolic or embodied representations: A case for symbol interdependency. In T. Landauer, D. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis*. Mahwah, NJ: Erlbaum.

Louwerse, M. M. (2008). Embodied relations are encoded in language. *Psychonomic Bulletin & Review, 15*(4), 838–844.

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

Louwerse, M. M., & Benesh, N. (2012). Representing spatial structure through maps and language: Lord of the Rings encodes the spatial structure of Middle Earth. *Cognitive Science, 36*(8), 1556–1569.

Louwerse, M. M., & Connell, L. (2011). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science, 35*(2), 381–398.

Louwerse, M. M., & Jeuniaux, P. (2010). The linguistic and embodied nature of conceptual processing. *Cognition, 114*(1), 96–104.

Louwerse, M. M., & Zwaan, R. A. (2009). Language encodes geographical information. *Cognitive Science, 33*(1), 51–73.

Louwerse, M. M., Hutchinson, S., & Cai, Z. (2012). The Chinese Route Argument: Predicting the Longitude and Latitude of Cities in China and the Middle East Using Statistical Linguistic Frequencies. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

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

Matlock, T., Ramscar, M., & Boroditsky, L. (2005). The experiential link between spatial and temporal language. *Cognitive Science, 29*, 655-664.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14*, 119–124.

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

Pedhazur, E. J. (1997). *Multiple regression in behavioral research*. Fort Worth, TX: Harcourt Brace.

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

Spence, C., Nicholls, M. E., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics, 63*(2), 330–336.

Tobler, W. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography Supplement, 46*, 234–240.

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. & Yaxley, R. (2003). Spatial iconicity affects semantic relatedness judgments. *Psychonomic Bulletin and Review, 10*, 954-958.