Flexible Shortcuts: Linguistic Distributional Information Affects both Shallow and Deep Conceptual Processing

Louise Connell (louise.connell@manchester.ac.uk)
School of Psychological Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK

Dermot Lynott (dermot.lynott@manchester.ac.uk)
Decision and Cognitive Sciences Research Centre, Manchester Business School, University of Manchester
Booth Street West, Manchester M15 6PB, UK

Abstract

Previous research has shown that people use both embodied perceptual simulations and linguistic distributional knowledge during conceptual processing, with linguistic information especially useful for shallow tasks and rapid responding. Using two conceptual combination tasks, we show that this linguistic shortcut is evident in both shallow and deep conceptual processing of novel stimuli. Specifically, in both shallow sensibility judgement and deep interpretation generation tasks, people use the linguistic shortcut as a "quick and dirty" guide to whether the concepts are likely to combine in a coherent situated simulation. Linguistic distributional frequency predicts both the likelihood and timecourse of rejecting a novel word compound as nonsensical or uninterpretable. However, it only predicts the timecourse of successful processing in shallow sensibility judgement because deeper interpretation generation requires conceptual processing in the simulation system.

Keywords: conceptual combination; linguistic distributional information; embodied cognition; simulation.

Introduction

The embodied simulation view of conceptual representation holds that the same neural systems that are responsible for representing information during perception, action, and introspection are also responsible for representing (or *simulating*) the same information during conceptual thought (e.g., Barsalou, 1999; Glenberg, 1997). Furthermore, concepts do not exist in a representational vacuum, but rather are situated within a broader situational context that includes perceptual, motor, affective and social information on how that concept has been experienced in the past (e.g., Barsalou & Wiemer-Hastings, 2005; Lynott & Connell, 2010a). A *cactus*, for example, can potentially include visual information (e.g., its green colour and prickly surface), tactile information (e.g., the sharpness of its spines), and affective information (e.g., negative valence for anyone who has spent days picking spines from skin), all situated relative to other concepts (e.g., in a desert location or as a pot plant on a kitchen windowsill).

However, the simulation system does not act alone. People are sensitive to distributional, statistical patterns in language and the wider environment, and this sensitivity provides a powerful generalised learning mechanism from early infancy (Aslin et al., 1998; Kirkham et al., 2002). Even in adults, the linguistic system contains statistical distributional information in a dynamic web of word-to-word (and phrase-to-phrase) associations that is powerful enough to support superficial strategies in a broad range of linguistic and conceptual tasks (e.g., Barsalou, Santos, Simmons & Wilson, 2008; Louwerse & Jeuniaux, 2008; Lynott & Connell, 2010a). The linguistic and simulation systems are closely interconnected and mutually supportive; linguistic information can activate simulation information, which may in turn activate further linguistic information, and so on. For example, when the word "cactus" is encountered, closely related linguistic tokens such as "prickly" and "sharp" will be activated, which will in turn begin to activate their relevant grounded representations in the simulation system, thus drawing attention to the visual and haptic modalities. Because their structures are both based on experience, the linguistic and simulation systems mirror each other to a certain extent, which suggests that information from language alone can approximate the perceptual, motor, affective, etc. content of concepts. Supporting this view, Louwerse and Connell (2011) have shown that linguistic distributional information is capable of distinguishing words on the basis of their perceptual modality. Words like *rustling*, *glistening*, and *freezing* refer to object properties in particular perceptual modalities (i.e., auditory, visual, and haptic) and occur in language with particular usage patterns. Louwerse and Connell showed that statistical analysis of these distributional patterns (based on 5-gram co-occurrence frequencies from a large corpus) produced three clusters that corresponded to auditory, visuohaptic and olfactogustatory modality groups. In other words, although auditory words were distinct, distributional information could not distinguish vision from touch, nor smell from taste. These three "linguistic modalities" (i.e., modality-specific clusters within the linguistic system) of auditory, visuohaptic and olfactogustatory words are therefore a coarse-grained approximation of the perceptual reality of five modalities. Linguistic distributional information is, at best, a blurred mirror of the simulation system.

The essential difference between the two systems is that the linguistic system is best for "quick and dirty" judgements, while the simulation system is best for deeper conceptual processing. When a word such as "cactus" is heard or read, both systems are kickstarted but the linguistic system peaks in activation (e.g., spreads activation to other tokens “prickly”, “sharp”, and so on) before the simulation system peaks (e.g., forms a visual, haptic, affective situated simulation of a *cactus*). The linguistic system thus has the potential to act as a shortcut and provide a response before the relatively more expensive simulation system is fully engaged. Support for this idea comes from Louwerse and Connell (2011), who compared the abilities of the linguistic and simulation systems to predict modality switching costs in property verification tasks. Switching costs refer to the finding that people are slower to confirm that a perceptual property is true of an object (e.g., auditory *leaves can be rustling*) when it follows a property from a different modality (e.g., visual *dew can be glistening*), and this processing cost is assumed to arise from the re-allocation of attention between modality-specific areas during perceptual simulation of the object property in question (Pecher, Zeelenberg & Barsalou, 2003). When Louwerse and Connell examined whether switching costs were best predicted by “linguistic modalities” (i.e., auditory, visuohaptic, and olfactogustatory word clusters) or actual perceptual modalities (i.e., auditory, gustatory, haptic, olfactory, and visual categories, based on human ratings), they found that the linguistic shortcut was the best predictor of fast responses, whereas perceptual simulation of five modalities was the best predictor of slow responses. In short, the linguistic system offers a fuzzy approximation that can provide an adequate heuristic in certain tasks, whereas the simulation system provides representational precision for more complex and precise conceptual processing.

The Current Study

Although Louwerse and Connell's (2011) study offers important evidence for the role of the linguistic shortcut conceptual processing, it is based on the retrieval of familiar information that is always expected to be successful. Most of human cognition is not like that, however. In order to function in a normal environment, we must be able to represent new concepts and process unfamiliar information, and work within the constraint that our conceptual processing is not always successful. Indeed, one of the key issues of a cognitive system with limited resource capacity is that not everything *should* be processed; a cognitive triage mechanism – an automatic means to determine whether it is worth expending precious representational and executive resources on a particular conceptual task, or whether it should be abandoned pending further clarification / information – would offer an invaluable aid to efficient functioning. A strong test of the linguistic shortcut hypothesis would therefore predict that use of the shortcut should be evident in the processing of (1) novel stimuli, (2) for successful responses in relatively shallow conceptual tasks, and (3) for apparent failures where a process is halted as not worth the effort, regardless of the depth of processing ostensibly involved in the task.

In the present experiments, we examined the role of the linguistic shortcut in conceptual combination using both shallow and deep processing tasks. Conceptual combination is the process of understanding novel word compounds such as *cactus beetle* or *elephant complaint*, and is predicated upon the inherently constructive nature of cognition that allows us to represent new concepts by mentally manipulating old ones For example, a *cactus beetle* may be represented as a beetle that feeds on cacti, or as a green and prickly beetle; both are equally valid end products of a successful combination process. Recently, Lynott and Connell (2010a) proposed the Embodied Conceptual Combination (ECCo) theory, which argues for a distinct role for the linguistic system during conceptual combination that complements that of the simulation system. Specifically, if the two nouns in a compound have little shared statistical, distributional history from language use, then the linguistic system offers people a reasonable heuristic for rejecting the compound as incomprehensible without expending much cognitive effort in attempting to combine the concepts. In contrast, if the nouns have frequently been encountered in close proximity to one another, then the linguistic system offers people a reasonable heuristic for accepting that the concepts can probably be combined in a shared, situated simulation.

Both sensibility judgement and interpretation generation tasks are commonly used in conceptual combination studies, but they differ in the required depth of processing (Lynott & Connell, 2010a). Sensibility judgement (Experiment 1) is relatively shallow because it simply asks people whether or not a particular compound makes sense. Interpretation generation (Experiment 2) is relatively deep because it asks people whether or not they can think of a meaning for a particular compound, and, if so, to specify the meaning. We therefore expected the linguistic system to play a differential role in conceptual combination according to task requirements: as a shortcut for both accepting and rejecting compounds in sensibility judgements, but only for rejecting compounds in interpretation generation because successful processing requires detailed representation in the simulation system.

Experiment 1: Sensibility Judgement

In this experiment, we presented people with novel noun-noun compounds in a forced-choice sensibility judgement task, where they pressed “yes” if they thought the compound phrase made sense, and pressed “no” if they thought it was nonsense. Similar methods have been used in a number of previous conceptual combination studies (e.g., Gagné & Shoben, 1997; Estes, 2003; Tagalakis & Keane, 2006). We measured response times to press both “yes” (i.e., accept as sensible because of successful combination) and “no” (reject as nonsense because of failed combination) keys. Following ECCo's proposal regarding the nature of the linguistic shortcut in sensibility judgements, we predicted inverse effects for acceptance and rejection of compounds. Linguistic distributional frequency (i.e., how frequently the two nouns have shared a context) should be *negatively* related to acceptance rates and times because high-frequency compounds will quickly appear sensible: the linguistic shortcut allows people to assume the concepts in question can combine merely because their two nouns have been frequently juxtaposed. In contrast, linguistic distributional frequency should be *positively* related to rejection times because low-frequency compounds will quickly appear nonsensical: the linguistic shortcut allows them to be dismissed out of hand rather than attempting a costly and potentially pointless combination effort in the simulation system.

Method

Materials Forty one noun-noun compounds were used in this study: 27 novel test items and 14 lexicalised filler items. Test items comprised novel noun-noun compounds (e.g., *octopus apartment*, *elephant complaint, whale knife*) with a British National Corpus phrase frequency greater than 20 (BNC, 2001), and featured a range of concept types (i.e., artifacts, natural kinds, abstract concepts). Filler items were lexicalised noun-noun compounds (e.g., *hospital wing*, *guerrilla warfare*) with a BNC frequency greater than 20, and were included to provide a baseline of highly sensible combinations to ensure that participants attended to the task.

In order to approximate the linguistic distributional information available to novel compounds, we carried out a corpus analysis using the Web 1T 5-gram corpus (Brants & Franz, 2006), which contains over a trillion tokens culled from Google indices and thus allows extensive analysis of linguistic distributional patterns[1]. For each compound, we calculated the cumulative 5-gram frequency of occurrence between the modifier and head nouns (e.g., the summed count of *octopus … apartment* with zero, one, two and three intervening words: for a similar approach, see Louwerse & Connell, 2011). Finally, frequencies were log-transformed as ln $(f + c)$, where f is the raw frequency and c is a constant (minimum non-zero frequency) added to all values to enable log calculations of zero counts.

All novel compounds were potentially sensible because they had been successfully interpreted by a majority of participants in previous studies (Lynott & Connell, 2010b). Critical to our present purposes, data from an offline pretest (i.e., an open-response task under no time constraints: N = 20) showed no reliable relationship between items' linguistic distributional frequency and success rate of interpretation, $r(25) = .170$, one-tailed $p = .198$.

Participants Twenty-four native speakers of English completed the experiment for a nominal sum. One participant was excluded for judging a majority of lexicalised filler items as nonsensical.

Procedure Participants were told that they would be presented with two-word phrases onscreen; some of these phrases would be familiar to them, while others would not. They were instructed to press the key labelled "Yes" to indicate that the phrase made sense or to press the key labelled "No" to indicate the phrase was nonsense. All responses were made with the participant's dominant hand.

Each trial began with the word "Ready" appearing on the screen for 2000 ms, followed by the compound which remained onscreen until the participant made a decision. Response times were recorded in seconds from the onset of the compound until the participant's keypress ("Yes" or "No" button). There was a blank screen interval of 1000 ms until the start of the next trial. Each participant saw all compounds presented in a different random order. The experiment took less than 10 minutes to complete.

Design & Analysis Response decision data (i.e., whether a compound was accepted or rejected) were analysed in a mixed-effects logistic regression model (logit link function) with crossed random factors of participants and items. The inclusion of items was empirically validated because it improved model fit over participants alone, $\chi^2(1) = 38.31$, $p < .0001$. Linguistic frequency (i.e., log 5-gram frequency per compound) acted as a fixed predictors variable. Response time data were analysed in a mixed-effects linear regression model with participants as a random factor. Items were not included as a crossed random factor because it did not further improve model fit, $\chi^2(1) = 2.55$, $p = .111$ (Baayen, Davidson & Bates, 2008). Response decision (i.e., yes or no) and linguistic frequency (i.e., log 5-gram frequency per compound) acted as fixed interacting predictors variables. The primary advantages of mixed effects analysis as regards the present experiment is that it can determine the effect of item-level predictors while simultaneously taking participant variability into account, and that it offers greater power than analysing aggregated responses over participants or items (Baayen et al., 2008; Locker, Hoffman & Bovaird, 2007). Regression coefficients are reported as unstandardized β values. Effect size r for each predictor was calculated from t (Cohen, 1988).

Results & Discussion

Data points more than 2.5 standard deviations from each participant's mean time per response decision were removed as outliers: 1.6% for "yes" responses and 2.4% for "no".

Acceptance / Rejection Rates Overall, 31.6% of novel compounds were judged as sensible and 68.4% as nonsense. As predicted, the likelihood of accepting a noun-noun

---

[1] Note that a broader co-occurrence measure like LSA (Landauer & Dumais, 1997) is not the same as the 5-gram frequency count we use here. LSA measures co-occurrence over a broad paragraph-length window before reducing the total matrix to approximately 300 dimensions, so distance between words can be calculated as the cosine of the angle between two points in this high-d space. LSA scores between words therefore reflect a broad linguistic similarity, such that synonyms, which often occur in the same general contexts, should receive a high score. In contrast, n-gram frequencies measure co-occurrence within a narrow window of local context (i.e., with 0-3 intervening words for 5-grams). N-gram frequencies between words therefore reflect whether words are used in close proximity with one another. They do not reflect similarity of meaning because synonyms, which occur within 0-3 words of each other only rarely, should receive a low score.

compound as sensible increased with linguistic distributional frequency, *t*(606) = 4.63, $p < .0001$, β = 0.251, *r* = .185. Even though all the compounds were novel stimuli with no pre-specified definition, the fact that two nouns had been relatively frequently juxtaposed was enough to allow their combination to seem sensible.

Acceptance / Rejection Times Sensibility acceptance times (*M* = 2.625, *SE* = 0.096) were generally slower than rejection times (*M* = 2.364, *SE* = 0.144), *t*(557.1) = 3.03, *p* = .003, β = 1.151, *r* = .127. Linguistic frequency had a marginally positive effect on overall response times, *t*(556.5) = 1.89, *p* = .059, β = 0.084, *r* = .080; but critically interacted with response decision to produce a negative effect on acceptance times, *t*(556.8) = –2.70, *p* = .007, β = –0.187, *r* = .114. Separate analysis of "yes" and "no" responses showed the predicted inverse effects (see Figure 1). The time taken to accept a novel compound as sensible decreased with greater linguistic frequency, *t*(162.9) = –2.62, *p* = .005, β = –0.140, *r* = .201, whereas the time to judge a compound as nonsense increased with linguistic frequency (i.e., low frequency compounds were rejected quickly, high frequency compounds were not), *t*(376.8) = 1.77, *p* = .039, β = 0.077, *r* = .091.

In other words, the linguistic shortcut acts to facilitate shallow conceptual combination by providing an heuristic of sensibility. Higher linguistic distributional frequency facilitates acceptance of a novel stimulus: words that often share a local context are quickly and frequently judged to be a sensible phrase, which constitutes successful (albeit "quick and dirty") processing of the combination. Lower linguistic frequency, however, facilitates rejection: words that rarely share a context are quickly and frequently judged to be a nonsensical phrase, which may appear to constitute a failed conceptual combination process, but is perhaps better regarded as successful avoidance of a potentially costly but fruitless cognitive effort. Of course, participants do not have to rely solely on this linguistic shortcut just because it exists, and are free to base their sensibility judgements on the simulation system. Nevertheless, the results of this experiment demonstrate a statistical tendency to use linguistic distributional information as a sensibility heuristic, even when individual differences between participants and items are partialled out. We return to this issue in the general discussion.

Experiment 2: Interpretation Generation

While the previous experiment examined a relatively shallow form of conceptual combination (i.e., judging whether a noun-noun compound made sense, but without having to specify why), this experiment focuses on a deeper form of processing by asking people to provide an actual interpretation for each compound. As before, we used a forced-choice task, where participants pressed "yes" if they could think of a meaning for the compound phrase (and then told us the meaning they had generated), and pressed "no" if they could not. Because the interpretation generation task invites deeper processing than sensibility judgement by asking people to think of a meaning, previous research has found it leads to more liberal use of "yes" decisions to novel compounds (Tagalakis & Keane, 2006). We therefore expected a larger proportion of items to be accepted than in

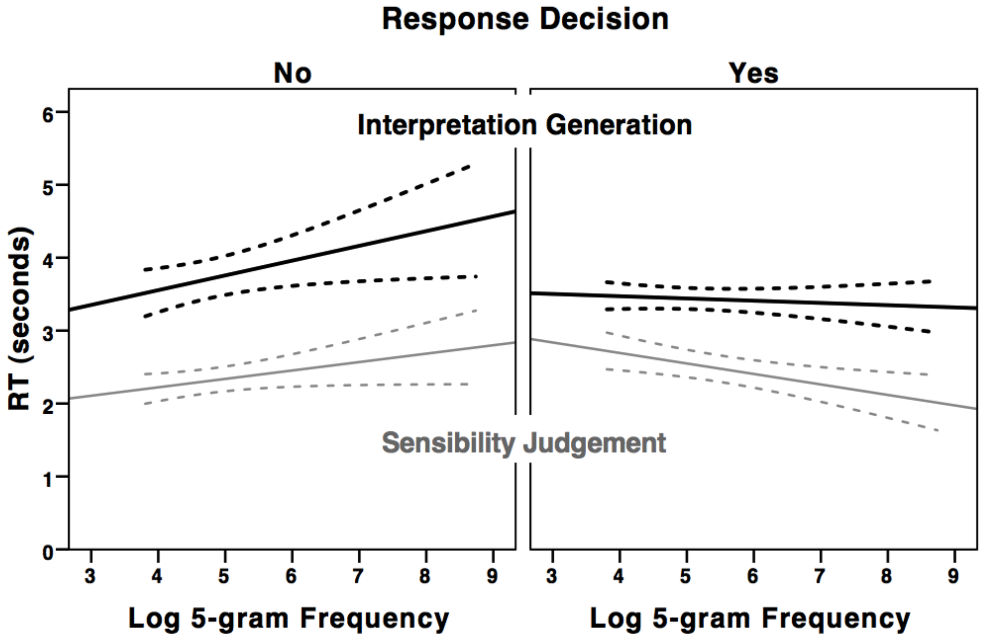

Figure 1: Regression plots of linguistic distributional frequency against model predicted response times for rejection ("no" decision") and acceptance ("yes" decisions) of novel noun-noun compounds in Experiment 1's sensibility judgement and Experiment 2's interpretation generation tasks. Dotted lines represent 95% confidence intervals around the mean. All fits except "yes" responses in interpretation generation are significant at $p < .05$.

Experiment 1, but, as for sensibility judgement, we expected this acceptance rate to be positively related to linguistic distributional frequency. The linguistic shortcut should quickly make high-frequency compounds appear interpretable, and – because most people can generate meanings for these items when they try – their subsequent combination in the simulation system is likely to succeed. Acceptance times thus reflect the latency of full conceptual combination, and as such should not be predicted by mere linguistic frequency. Rejection times, on the other hand, should show the same positive relationship with linguistic frequency that we saw for sensibility judgement: words that seldom appear in the same contexts will be quickly and frequently rejected as uninterpretable because the linguistic shortcut suggests their concepts may not combine.

Method

Materials As per Experiment 1.

Participants Eighteen native speakers of English completed the experiment for a nominal sum.

Procedure Instructions were identical to Experiment 1 except that participants were asked to press the key labelled "Yes" to indicate that "Yes, I can think of a meaning" (whereupon a screen appeared for them to type in the interpretation just generated), or to press the key labelled "No" to indicate that "No, I cannot think of a meaning". The experiment took approximately 20 minutes to complete and had a short, self-paced, break halfway through.

Design & Analysis Data were analysed with crossed random factors because model fit improved with the inclusion of items for both logistic regression of response decision data, $\chi^2(1) = 69.95$, $p < .0001$, and linear regression of response time data, $\chi^2(1) = 6.17$, $p = .013$. All other details were the same as Experiment 1.

Results & Discussion

2.2% of "yes" responses to novel compounds resulted in blank or invalid interpretation (e.g., "a", "I don't know") and were excluded from analysis as they did not represent successful combination. Data points more than 2.5 standard deviations from each participant's mean per response decision were removed as outliers: 1.3% for "yes" responses and 2.8% for "no".

Acceptance / Rejection Rates Overall, 68.5% of compounds were accepted and successfully interpreted and 31.5% were rejected as uninterpretable. Each compound had a variety of different, coherent interpretations, such as a *whale knife* as "A knife that has a picture of a whale on it" or "knife used by whalers", or an *elephant complaint* as "a large complaint" or "a complaint about elephants in the area". As predicted, the likelihood of successfully interpreting a noun-noun compound increased with linguistic distributional frequency, $t(439) = 2.10$, $p = .019$. $\beta = 0.145$, $r = .100$.

Acceptance / Rejection Times Interpretation times ($M = 3.348$, $SE = 0.110$) were marginally faster than rejection times ($M = 3.713$, $SE = 0.194$), $t(413.7) = 1.77$, $p = .077$, $\beta = 0.974$, $r = .087$. Linguistic frequency had an overall positive relationship with response times, $t(117.0) = 2.14$, $p = .034$, $\beta = 0.209$, $r = .194$; but, critically, it negatively interacted with response decision, $t(413.8) = -2.44$, $p = .015$, $\beta = -0.259$, $r = .119$. Results for separate analysis of "yes" and "no" responses were as predicted (Figure 1). The time taken to accept and interpret a novel compound was unaffected by linguistic distributional frequency, $t<1$. Like sensibility judgements, however, the time to reject a compound as uninterpretable increased with linguistic frequency, $t(120.7) = 2.74$, $p = .007$, $\beta = -0.256$, $r = .242$.

In both shallow sensibility judgement and deep interpretation generation tasks, people use the linguistic shortcut as a "quick and dirty" guide to whether the concepts are likely to combine in a coherent situated simulation. Building a representation that is detailed enough to provide an interpretation is a function of deep conceptual processing in the simulation system, and took some 700 ms longer than accepting a compound as sensible. This extra depth of processing meant that successful interpretation times were no longer predicted by information from the linguistic system. Rejection times were also slower for interpretation generation than for sensibility judgement, and the 1300 ms difference suggests that at least some "no" responses resulted from tried-and-failed conceptual combination in the simulation system. However, the fact that rejection times were still strongly predicted by linguistic distributional frequency shows that the linguistic shortcut offered an important heuristic for avoiding this resource-wasting event.

General Discussion

There are three novel findings in the present paper. First, we show that linguistic distributional frequency can predict not only the timecourse of successful conceptual processing (i.e., "yes" responses in sensibility judgement), but also the timecourse and likelihood of failure (i.e., "no" responses). Second, use of this linguistic shortcut extends beyond simple retrieval into the processing of novel stimuli in conceptual combination. The more often two words have appeared in close proximity to one another, the faster people are to accept the compound as sensible and the slower they are to reject it as uninterpretable nonsense. Third, we show that the influence of such linguistic shortcuts is not restricted to shallow conceptual tasks, but is also useful in deeper conceptual processing as a form of cognitive triage. The less often two words have appeared in close proximity, the faster people reject their compound as uninterpretable rather than risk costly failure in the simulation system. These findings support theories that argue for complementary roles of the linguistic and simulation systems in conceptual combination (Lynott & Connell, 2010a) and conceptual processing more generally (Barsalou et al., 2008; Louwerse & Jeuniaux, 2008).

But isn't all this just standard *word frequency effects*? In a word, no. We can't observe the above range of effects in conventional psycholinguistic tasks such as lexical decision or word naming. Firstly, responses in lexical decision and naming tasks are either correct or incorrect (e.g., correctly rejecting a non-word), whereas novel compounds do not necessarily have a "correct" interpretation. Rather, an individual's processing of a compound is either successful or unsuccessful, and even an "unsuccessful" outcome may represent the most efficient use of cognitive resources. Secondly, lexical decision and naming tasks rely solely on the recognition of known concepts, while conceptual combination tasks require the processing of new conceptual entities. Thus, the paradigm in this paper allows us to examine the conceptualisation of novel stimuli at two depths of processing, and demonstrate how the linguistic shortcut offers a useful heuristic in both shallow and deep tasks.

Of course, participants do not have to rely solely on a linguistic shortcut just because it exists. An individual may double-check apparently sensible or apparently uninterpretable compounds within the simulation system by actually attempting to combine the concepts. Indeed, it is possible that some particularly cautious individuals may even base every sensibility judgement on whether the concepts can combine into a coherent simulation. However, an easy shortcut is hard to refuse. Because the linguistic shortcut is faster and computationally cheaper than basing a judgment on the simulation system, and because on-the-fly conceptual processing does not have to be perfect (only "good enough": Ferreira, Bailey & Ferraro, 2002), participants can safely exploit it most of the time.

Acknowledgements

This research was supported in part by the UK Economic and Social Research Council [grant RES-000-22-3248].

References

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*, 321–324.

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. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*, 577-609.

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, A. (Eds.). *Symbols, embodiment, and meaning*. Oxford, UK: Oxford University Press.

Barsalou, L. W., & Wiemer-Hastings, K. (2005). Situating abstract concepts. In D. Pecher & R. A. Zwaan, *Grounding cognition: The role of perception and action in memory, language, and thinking* (pp. 129– 163). Cambridge, UK: Cambridge University Press.

The British National Corpus, Version 2 (BNC World) (2001). Distributed by Oxford University Computing Services on behalf of the BNC Consortium. Available at http://www.natcorp.ox.ac.uk.

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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Connell, L., & Lynott, D. (2011). Modality switching costs emerge in concept creation as well as retrieval. *Cognitive Science, 35*, 763-778.

Estes, Z. (2003). Attributive and relational processes in nominal combination. *Journal of Memory and Language, 48*, 304-319.

Ferreira, F., Ferraro, V., & Bailey, K. G. D. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*, 11–15.

Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 71-87.

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

Kirkham, N.Z., Slemmer, J.A., & Johnson, S.P. (2002). Visual statistical learning in infancy: Evidence of a domain general learning mechanism. *Cognition, 83*, B35–B42.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis Theory of acquisition, induction and representation of knowledge. *Psychological Review, 104*, 211–240.

Locker, L., Hoffman, L., & Bovaird, J. A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavior Research Methods, 39*, 723–730.

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., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. de Vega, A. Glenberg, & A. C. Graesser (Eds.), *Symbols, embodiment, and meaning*. Oxford University Press.

Lynott, D., & Connell, L. (2010a). Embodied conceptual combination. *Frontiers in Psychology, 1*(216), 1-14.

Lynott, D., & Connell, L. (2010b). The effect of prosody on conceptual combination. *Cognitive Science, 34*, 1107-1123.

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

Tagalakis, G., & Keane, M. T. (2006). Familiarity and relational preference in the understanding of noun–noun compounds. *Memory & Cognition, 34*, 1285–1297.