BRIEF REPORT

Flexible and fast: Linguistic shortcut affects both shallow and deep conceptual processing

Louise Connell · Dermot Lynott

Abstract Previous research has shown that people use linguistic distributional information during conceptual processing, and that it is especially useful for shallow tasks and rapid responding. Using two conceptual combination tasks, we showed that this linguistic shortcut extends to the processing of novel stimuli, is used in both successful and unsuccessful conceptual processing, and is evident in both shallow and deep conceptual tasks. Specifically, as predicted by the ECCo theory of conceptual combination, people use the linguistic shortcut as a "quick-and-dirty" guide to whether the concepts are likely to combine into a coherent conceptual representation, in both shallow sensibility judgment and deep interpretation generation tasks. Linguistic distributional frequency predicts both the likelihood and the time course of rejecting a novel word compound as nonsensical or uninterpretable. However, it predicts the time course of successful processing only in shallow sensibility judgment, because the deeper conceptual process of interpretation generation does not allow the linguistic shortcut to suffice. Furthermore, the effects of linguistic distributional frequency are independent of any effects of conventional word frequency. We discuss the utility of the linguistic shortcut as a cognitive triage mechanism that can optimize processing in a limited-resource conceptual system.

L. Connell (✉)
School of Psychological Sciences, University of Manchester,
Oxford Road,
Manchester M13 9PL, UK
e-mail: louise.connell@manchester.ac.uk

D. Lynott (✉)
Decision and Cognitive Sciences Research Centre,
Manchester Business School, University of Manchester,
MBS East, Booth Street West,
Manchester M15 6PB, UK
e-mail: dermot.lynott@manchester.ac.uk

Conceptual processing uses linguistic distributional information. That is, people are sensitive to the distributional, statistical patterns in language and the wider environment, and this sensitivity provides a powerful generalized learning mechanism from early infancy (Aslin, Saffran, & Newport, 1998; Kirkham, Slemmer, & Johnson, 2002). Even in adults, such statistical distributional information forms a linguistic system of dynamic 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 (see, e.g., Barsalou, Santos, Simmons, & Wilson, 2008; Connell & Lynott, 2013; Louwerse & Jeuniaux, 2008; Lynott & Connell, 2010b).

Information from language alone is powerful enough to capture many aspects of real-world experience, even to the extent of approximating the perceptual, motor, affective, and so forth, content of concepts. For example, Louwerse and Connell (2011) have shown that the linguistic system is, to a certain extent, capable of distinguishing between words that relate to different perceptual modalities. 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 alone 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 coarse-grained approximations of the perceptual reality of five modalities. Linguistic distributional information is, at best, a blurred mirror of a deeper, more precise conceptual representation.

The linguistic system can be characterized as being best for "quick-and-dirty" judgments, because it has the potential to provide a response before deeper conceptual processing is fully engaged (Lynott & Connell, 2010b). For example, reading the word "cactus" activates other relevant words— such as "prickly," "sharp," and so on—before the detailed conceptual representation of a *cactus* as a big green desert plant is fully activated. In other words, people have a *linguistic shortcut* available during conceptual processing, in that computationally cheaper information from the linguistic system can usefully inform a response in a particular task before relatively more expensive (but precise) representations are fully available. Support for this idea has come from Louwerse and Connell (2011; see also Jones & Golonka, 2012; Louwerse & Hutchinson, 2012; Solomon & Barsalou, 2004), who examined the ability of the linguistic shortcut to predict modality-switching costs in a property verification task. *Switching costs* refer to the finding that people are slower to confirm that a perceptual object property is true (e.g., the auditory property *leaves can be rustling*) when it follows a property from a different modality (e.g., the visual property *dew can be glistening*), and this processing cost is assumed to arise from the reallocation of attention between modality-specific areas during representation 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 by 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, with its effect size diminishing the longer that participants took to respond. In short, the linguistic system offers a fuzzy heuristic in certain tasks in which a shallow, "good-enough representation" (Ferreira, Bailey, & Ferraro, 2002) can suffice.

Although Louwerse and Connell's (2011) study offers important evidence of the role of the linguistic shortcut in conceptual processing, it was based on the retrieval of familiar information, which 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, 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—that is, an automatic means to determine

whether it is worth expending precious representational capacity and executive-processing resources on a particular conceptual task, or whether such processing 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 be to examine whether the use of the shortcut is evident (1) in the processing of novel stimuli, (2) for successful responses in relatively shallow conceptual tasks, and (3) for apparent failures in which a process is halted as not being worth the effort, regardless of the depth of processing ostensibly involved in the task.

Present study

In the present experiments, we examined the role of the linguistic shortcut in conceptual combination that involved 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, which 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 interpretations would be equally valid end products of a successful combination process. Alternatively, a *cactus beetle* might be rejected as an incomprehensible phrase by someone who cannot think of one of the above meanings, and this lack of a representation would be a valid end product of the combination process marking it as unsuccessful. Unlike standard psycholinguistic tasks such as lexical decision or word naming, in which a response is necessarily right or wrong, conceptual combination allows us to legitimately examine the time courses of both successful and unsuccessful processing, as well as the relative likelihoods of success and failure.

Recently, Lynott and Connell (2010b) proposed the embodied conceptual combination (ECCo) theory, which is currently the only theory that argues for a distinct role for the linguistic system during conceptual combination. Specifically, if two nouns in a compound have little shared distributional history from language use, then ECCo predicts that the linguistic shortcut would offer 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 been frequently encountered in close proximity to one another, then the linguistic shortcut offers people a reasonable heuristic for accepting that the concepts *could* probably be combined in a deeper, detailed representation.

Both sensibility judgment and interpretation generation tasks are commonly used in conceptual combination studies,

but they differ in the depths of processing they require (Lynott & Connell, 2010b). Sensibility judgment (Exp. 1) is relatively shallow, because it simply asks people whether or not a particular compound makes sense, and as such, does not necessarily require full representation of the combined concepts. Interpretation generation (Exp. 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, which necessarily requires the full representation of the new combination. We therefore expected the linguistic system to play differential roles in conceptual combination according to the task requirements: as a short-cut for both accepting and rejecting compounds in sensibility judgments, but only for rejecting compounds in interpretation generation, because successful processing would require detailed representation for which a linguistic heuristic would not suffice (see Connell & Lynott, 2011). In other words, we expected linguistic distributional frequency to predict the likelihood and latency of both accepting and rejecting a novel compound in sensibility judgment, but the likelihood and latency only of rejecting a novel compound in interpretation generation.

Experiment 1: Sensibility judgment

In this experiment, we presented people with novel noun–noun compounds in a forced choice sensibility judgment task, in which they pressed "yes" if they thought that the compound phrase made sense and pressed "no" if they thought that it was nonsense (see, e.g., Estes, 2003; Gagné & Shoben, 1997; Tagalakis & Keane, 2006). We measured decisions to press either the "yes" (i.e., to accept as sensible because of successful combination) or the "no" (to reject as nonsense because of failed combination) key and the corresponding response latencies. Following ECCo's proposal that the linguistic shortcut leads people to accept a compound as sensible if the constituent nouns have frequently been juxtaposed, we expected distributional frequency (i.e., how frequently the two nouns have shared a context) to be positively related to acceptance rates. In contrast, we did not expect conventional word frequency (i.e., how frequently the individual words occur in language) to play such a role. Regarding response times, conventional word frequency is traditionally negatively related to latency (i.e., higher word frequency leads to faster word recognition), and we expected to find a similar effect here for both acceptance and rejection times. Critically, we expected distributional frequency to behave quite differently from word frequency, in that the linguistic shortcut would predict inverse effects for acceptance and rejection times. Distributional frequency should be negatively related to acceptance times, because compounds with high scores

would quickly appear sensible; the linguistic shortcut allows people to assume that the concepts in question can combine, merely because their two nouns have often appeared in close proximity. In contrast, distributional frequency should be positively related to rejection times, because compounds with low scores would quickly appear nonsensical; the linguistic shortcut allows them to be dismissed out of hand, rather than requiring a costly, and potentially pointless, combination effort.

Method

Materials A total of 41 noun–noun compounds were used in this study: 27 novel test items and 14 lexicalized filler items. The test items comprised novel noun–noun compounds (e.g., *elephant complaint*: see the Appendix), with a British National Corpus (BNC; BNC Consortium, 2001) phrase frequency of zero, and featured a range of concept types (i.e., artifacts, natural kinds, and abstract concepts). The filler items were lexicalized noun–noun compounds (e.g., *hospital wing*, *guerrilla warfare:* BNC frequency > 20) and were included to provide a baseline of highly sensible combinations to ensure that participants attended to the task.

To approximate the linguistic distributional information available for the novel compounds, we employed 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 bidirectional 5-gram frequency of occurrence between the modifier and head nouns (e.g., the summed counts of the combinations *octopus . . . apartment* and *apartment . . . octopus* with zero, one, two and three intervening words; for a similar approach, see Louwerse & Connell, 2011). Distributional frequencies were then log-transformed as $\ln (f + c)$, where f is the raw frequency and c is a constant (minimum nonzero frequency) added to enable log-calculation of zero counts. Finally, we also calculated the individual word frequency of each compound by summing

---

[1] A broader co-occurrence measure like LSA (Landauer & Dumais, 1997) is not the same as 5-gram frequency counts. LSA measures co-occurrence over a broad paragraph-length window before reducing the total matrix to approximately 300 dimensions so that the distance between words can be calculated as the cosine of the angle between two points in this high-dimensional space. LSA scores between words therefore reflect 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., 0–3 intervening words for 5-grams). *N*-gram frequencies between words therefore reflect whether words are used in close proximity to one another. They do not reflect similarity of meaning, because synonyms, which occur within 0–3 words of each other only rarely, should receive low scores.

the word counts of the constituent nouns (Brysbaert & New, 2009; SUBTLEX$_{US}$) and log-transforming them. All of the log distributional frequencies and log word frequencies were within 2.5 standard deviations of their means.

All of the novel compounds were potentially sensible because they had been successfully interpreted by a majority of participants in previous studies (Lynott & Connell, 2010a). Critical to our present purposes, since we hypothesized that the linguistic shortcut would more likely be used when the task requirements favored a "quick-and-dirty" heuristic (Lynott & Connell, 2010b), the data from an offline pretest (i.e., an open-response task under no time constraints, $N = 20$) showed no reliable relationship between the items' distributional frequencies and the success rates of interpretation, $r(25) = .176$, one-tailed $p = .190$.

Participants A group of 24 native speakers of English completed the experiment for a nominal sum. One participant was excluded for judging a majority of the lexicalized fillers as being nonsensical.

Procedure The participants were told that they would be presented with two-word phrases on screen; some of these phrases would be familiar to them, while others would not. They were instructed to press, using their index fingers, a key labeled "yes" ("K" on a QWERTY keyboard) to indicate that the phrase made sense or a key labeled "no" ("D") to indicate that the phrase was nonsense. Participants were asked to respond as quickly and accurately as possible.

Each trial began with the word "Ready" appearing onscreen for 2,000 ms, followed by the compound, which remained visible until the participant made a decision. All text appeared in lowercase black Arial font on a white screen. Response times were recorded in seconds from the onset of the compound until the participant's keypress (yes/no button). An interstimulus interval (ISI) of 1,000 ms separated each trial from the next. Each participant saw all compounds presented in a different random order, and the experiment took approximately 10 min.

Design and analysis Response decision data (i.e., whether a compound was accepted or rejected) were analyzed in a mixed-effects logistic regression model (logit link function) with crossed random factors for participants and items. The inclusion of items was empirically validated, because it improved model fit over participants alone, $\chi^2(1) = 38.31$, $p < .0001$. Distributional frequency (i.e., the log 5-gram frequency per compound) and word frequency (i.e., the log summed frequency of the individual nouns per compound) acted as fixed predictor variables. Response time data were analyzed in a mixed-effects linear regression model with participants as a random factors. Items was not included as a crossed random factors 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 distributional frequency (i.e., log 5-gram frequency per compound) acted as fixed interacting predictors, with word frequency (i.e., log summed frequency of the individual nouns per compound) as an additional fixed predictor.[2] Mixed-effects analysis has the dual advantage of determining the effects of item-level predictors while simultaneously taking participant variability into account, and it offers greater power than aggregated analyses over participants or items (Baayen et al., 2008; Locker, Hoffman, & Bovaird, 2007). Regression coefficients are reported as unstandardized β values. The effect size r for each predictor was calculated from t (Cohen, 1988).

Results and discussion

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

Acceptance/rejection rates Overall, 31.6 % of the novel compounds were judged as being sensible, and 68.4 % as nonsense. As predicted, the likelihood of accepting a noun–noun compound as sensible increased with distributional frequency, $t(578) = 6.24$, $p < .0001$, $\beta = 0.388$, $r = .251$. Word frequency had a marginal effect, but in the opposite direction, $t(578) = -1.79$, $p = .074$, $\beta = -0.439$, $r = .074$ (see Fig. 1 for odds ratios). Thus, even though all of the compounds were novel, with no prespecified 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 ($M = 2.364$, $SE = 0.144$), $t(556.2) = 3.05$, $p = .002$, $\beta = 1.247$, $r = .128$. Distributional frequency had a positive effect on overall response times, $t(555.5) = 2.23$, $p = .026$, $\beta = 0.097$, $r = .094$, but critically interacted with response decision to produce a negative effect on acceptance times, $t(555.8) = -2.78$, $p = .006$, $\beta = -0.194$, $r = .117$. Word frequency overall was negatively related to response times, $t(555.1) = -2.65$, $p = .008$, $\beta = -0.346$, $r = .112$. Separate analysis of "yes" and "no" responses showed the predicted inverse effects (Fig. 2): The time taken to accept a novel compound as sensible decreased with greater distributional

---

[2] Because we expected conventional word frequency to exert the same effect on both acceptance and rejection times (i.e., a negative relationship), we did not initially enter an interaction with response decision into the regression model. Later checks showed that entering this interaction as a fixed variable was nonsignificant in all analyses and did not alter the pattern of findings. We therefore report the more parsimonious model that omits this interaction.
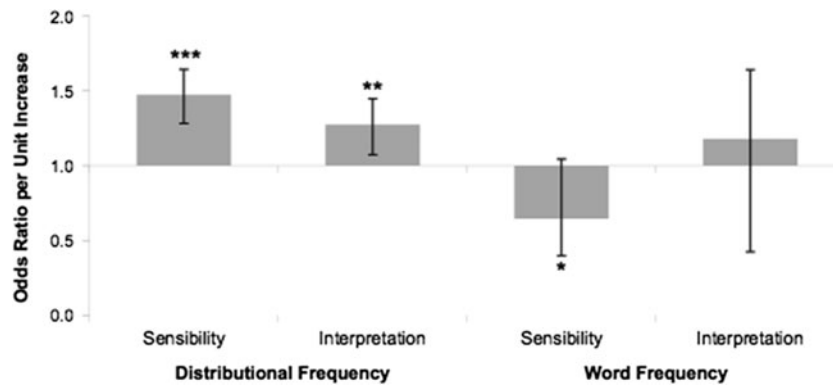
Fig. 1 Effects of distributional frequency and word frequency, expressed as odds ratios of acceptance over rejection responses, for sensibility judgment (Exp. 1) and interpretation generation (Exp. 2). Values over 1 represent increased likelihood of accepting a compound, and values below 1 represent decreased likelihood. The error bars represent 95 % confidence intervals, and asterisks represent the significance of directional predictions where specified (*** $p < .01$, ** $p < .05$, * $p < .1$)

frequency, $t(162.4) = -1.92$, $p = .028$, $\beta = -0.126$, $r = .149$, whereas the time to judge a compound as nonsense increased with distributional frequency (i.e., low-frequency compounds were rejected quickly, high-frequency compounds were not), $t(375.9) = 1.98$, $p = .024$, $\beta = 0.086$, $r = .102$. Word frequency had the same negative effect for both acceptance, $t(163.0) = -1.50$, $p = .067$, $\beta = -0.402$, $r = .117$, and rejection, $t(375.6) = -1.83$, $p = .034$, $\beta = -0.276$, $r = .094$.

In other words, the linguistic shortcut acts to facilitate shallow conceptual combination by providing a heuristic of sensibility. Higher 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. This effect is in addition to that of conventional word frequency, which also facilitates acceptance times. Lower distributional frequency, however, facilitates rejection: Words that rarely share a context are quickly and frequently judged to form a non-sensical 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. In this case, faster rejection times emerge from both lower distributional frequency *and* higher conventional word frequency. Of course, participants do not have to rely solely on this linguistic shortcut just because it exists, and they are free to base their sensibility judgments on full conceptual combination. 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 partialed out.

Experiment 2: Interpretation generation

While in the previous experiment we 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

Fig. 2 Effects of distributional frequency and word frequency, expressed as standardized regression coefficients, on acceptance and rejection times for sensibility judgment (Exp. 1) and interpretation generation (Exp. 2). Error bars represent 95 % confidence intervals, and asterisks represent the significance of directional predictions where specified (*** $p < .01$, ** $p < .05$, * $p < .1$)
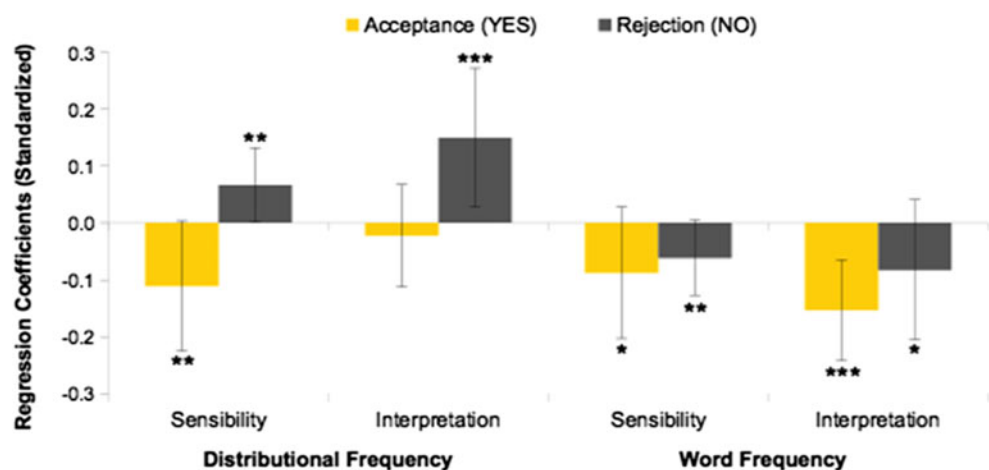
form of processing, in that people were asked to provide an actual interpretation for each compound. As before, we used a forced choice task in which participants pressed "yes" if they could think of a meaning for the compound phrase (and then told us the meaning that they had generated) and "no" if they could not. Because, by asking people to think of a meaning, the interpretation generation task invites deeper processing than does sensibility judgment, previous research has found that it leads to more liberal use of "yes" decisions to novel compounds (Tagalakis & Keane, 2006). Since pretests confirmed that our compounds were all interpretable, we therefore expected a larger proportion of items to be accepted than in Experiment 1, but as before, we expected this acceptance rate to be positively related to distributional frequency. Words that seldom appear in the same contexts should be quickly and frequently rejected as uninterpretable because the linguistic shortcut suggests that their concepts may not combine. Rejection times, as such, should show the same positive relationship with distributional frequency that we saw for sensibility judgment. However, since acceptance times reflect the latency of full conceptual combination, in which participants must generate a meaning for the compound, in this circumstance the linguistic shortcut cannot suffice (Connell & Lynott, 2011), so acceptance times should not be predicted by mere distributional frequency.

Method

The method and materials were the same as in Experiment 1, unless otherwise specified.

Participants A group of 18 native speakers of English completed the experiment for a nominal sum.

Procedure The participants were asked to press a key labeled "yes" to indicate "Yes, I can think of a meaning" (whereupon a screen appeared for them to type in the interpretation they had just generated) or to press a key labeled "no" to indicate "No, I cannot think of a meaning." They were instructed to be sure that they had an interpretation in mind when pressing the "Yes" button. The experiment took approximately 20 min to complete and had a short, self-paced break halfway through.

Design and analysis The data were analyzed with crossed random factors, because model fit improved with the inclusion of items for both the logistic regression of response decision data, $\chi^2(1) = 69.95$, $p < .0001$, and the linear regression of response time data, $\chi^2(1) = 6.17$, $p = .013$.

Results and discussion

In all, 2.2 % of the "yes" responses to novel compounds resulted in a 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 ("yes" responses, 1.3 %; "no" responses, 2.8 %).

Acceptance/rejection rates Overall, 68.5 % of the 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 distributional frequency, $t(438) = 3.26$, $p = .001$, $\beta = 0.243$, $r = .154$. Conventional word frequency, however, had no effect, $t(438) = 0.66$, $p = .511$, $\beta = 0.166$, $r = .031$ (see Fig. 1).

Acceptance/rejection times Interpretation acceptance times ($M = 3.348$, $SE = 0.110$) were marginally faster than rejection times ($M = 3.713$, $SE = 0.194$), $t(412.9) = 1.66$, $p = .097$, $\beta = 0.925$, $r = .082$. Distributional frequency had an overall positive relationship with response times, $t(121.3) = 2.04$, $p = .043$, $\beta = 0.192$, $r = .183$, but critically, it interacted with response decision to negatively affect response times, $t(412.3) = -2.30$, $p = .022$, $\beta = -0.240$, $r = .112$. As before, word frequency was negatively related to response times, $t(22.0) = -2.72$, $p = .012$, $\beta = -0.569$, $r = .502$. Results for the separate analyses of "yes" and "no" responses were as predicted (Fig. 2): The time taken to accept and interpret a novel compound was unaffected by distributional frequency, $t(284.4) = -0.50$, $p = .620$, $\beta = -0.026$, $r = .029$. As with sensibility judgments, however, the time to reject a compound as uninterpretable increased with distributional frequency, $t(120.5) = 2.43$, $p = .008$, $\beta = -0.235$, $r = .216$. Word frequency negatively affected response times for both acceptance, $t(283.6) = -3.40$, $p < .001$, $\beta = -0.707$, $r = .198$, and rejection, $t(121.3) = -1.33$, $p = .093$, $\beta = -0.414$, $r = .120$.

People use the linguistic shortcut as a "quick-and-dirty" guide to whether concepts are likely to combine successfully in both shallow sensibility judgment and deep interpretation generation tasks. Interpretation generation requires deeper conceptual processing than does accepting a compound as sensible (i.e., some 700 ms longer), which meant that successful interpretation times were no longer reliably predicted by information from the linguistic system. Rejection times were also slower for interpretation generation than for sensibility judgment (i.e., 1,300 ms longer), which suggests that at least some "no" responses resulted from tried-and-failed conceptual combination. However, the fact that rejection times were still strongly predicted by distributional

frequency shows that the linguistic shortcut offered an important heuristic for avoiding this resource-wasting event.

General discussion

We have demonstrated three novel findings in the present article. First, we have shown that linguistic distributional frequency can predict not only the time course of successful conceptual processing (i.e., "yes" responses in sensibility judgment), but also the time course and likelihood of failure (i.e., "no" responses). Second, the use of this linguistic shortcut extends beyond simple retrieval into the processing of novel stimuli in conceptual combination: The more often that 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 have shown 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 that two words have appeared in close proximity, the faster people are to reject their compound as uninterpretable, which allows people to avoid the risk of costly failure in later processing. These findings support theories that argue for the important role of the linguistic system in conceptual combination (i.e., ECCo; Lynott & Connell, 2010b) and in conceptual processing more generally, such as the language-and-situated-simulation theory (Barsalou et al., 2008) and the symbol interdependency theory (Louwerse, 2011; Louwerse & Jeuniaux, 2008).

Moreover, the effects of distributional frequency are independent from those of conventional word frequency, and in the case of rejection times, run in the opposite direction. Words that are highly frequent in language are processed more rapidly than are words that are infrequent, which can be seen in standard psycholinguistic tasks such as lexical decision or word naming (e.g., Brysbaert & New, 2009). We found the same effects in both of our experiments: Reading a compound was faster when the constituent nouns had higher individual word frequencies, regardless of whether the task involved sensibility judgment or interpretation generation, and regardless of whether the compound was accepted or rejected. However, accepting a compound as sensible was additionally facilitated when the words had *frequently* appeared in close proximity, and rejecting a compound as nonsensical or uninterpretable was additionally facilitated when the words had *seldom* appeared in close proximity. This differential pattern of results suggests that the effects of conventional word frequency operate at an earlier stage of processing than do those of distributional frequency: Word frequency facilitates

recognizing the individual word forms, and distributional frequency facilitates the conceptual combination process by acting as a linguistic shortcut that can provide a useful basis for a response before deeper conceptual processing has been fully engaged. Indeed, the only situation in which this linguistic shortcut does not provide a useful basis is when successfully generating an interpretation for a novel compound (Exp. 2). While there may be other reasons why distributional frequency does not affect the latency of successful interpretations, we believe that the theoretical account of the roles of the linguistic and simulation systems in conceptual combination (Lynott & Connell, 2010b), and in conceptual processing in general (Barsalou et al., 2008; Connell & Lynott, 2013; Louwerse, 2011), provides the most likely explanation, and future work should explore the interaction of these systems. Regardless, the linguistic shortcut is a "quick-and-dirty" heuristic that simply does not work when a complex and/or detailed conceptual representation is required.

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 by actually attempting to combine the concepts. Indeed, it is possible that some particularly cautious individuals may even base every sensibility judgment on whether the concepts can combine into a new and coherent concept. However, an easy shortcut is hard to refuse. Because the linguistic shortcut is faster and computationally cheaper than more precise conceptual representations (e.g., Louwerse & Connell, 2011), and because on-the-fly conceptual processing does not have to be perfect (only "good enough": Ferreira et al., 2002), participants can safely exploit it across a wide variety of tasks and goals. The present findings highlight people's sensitivity to distributional regularities in the linguistic environment and demonstrate how such patterns can impact on higher-level cognitive processing. In a limited-resource cognitive system, not everything should be processed in depth, and the present findings provide one example of a cognitive triage mechanism that can determine relatively cheaply whether it will be worth expending precious resources on a particular conceptual task. Inevitably, there will be times when such a mechanism will cause an individual to abandon a process that with persistence would have succeeded, but, so long as the benefits outweigh the costs, cognitive triage mechanisms like the linguistic shortcut remain a valuable aid to efficient cognitive function.

Appendix

Table 1 Novel noun–noun compounds, with their log distributional (5-gram) frequencies and log summed individual word frequencies, employed in Experiments 1 (sensibility judgment) and 2 (interpretation generation), alongside the acceptance rates and mean acceptance and rejection times in each experiment

Compound	Log 5-Gram Frequency	Log Word Frequency	Sensibility Judgment			Interpretation Generation		
			Acceptance Rate (%)	Acceptance Time (ms)	Rejection Time (ms)	Acceptance Rate (%)	Acceptance Time (ms)	Rejection Time (ms)
airplane chemical	5.017	3.056	40.9 (10.7)	2,073 (330)	3,200 (714)	75.0 (11.2)	2,819 (264)	4,735 (2,016)
antelope coconut	3.807	2.456	18.2 (8.4)	5,518 (1,770)	2,219 (407)	35.3 (11.9)	4,145 (374)	3,636 (529)
army decision	9.006	3.857	77.3 (9.1)	1,870 (259)	4,634 (2,166)	94.1 (5.9)	3,071 (472)	1,719 (0)
bed helicopter	3.807	40.15	4.8 (4.8)	2,194 (0)	2,219 (325)	47.1 (12.5)	3,866 (791)	3,071 (589)
cabbage cup	8.755	3.445	36.4 (10.5)	2,061 (293)	2,801 (642)	70.6 (11.4)	3,232 (389)	5,705 (1,458)
celebrity spear	6.853	2.783	14.3 (7.8)	4,788 (1,382)	2,728 (526)	50.0 (12.9)	3,957 (386)	3,332 (477)
chocolate clay	7.812	3.325	54.5 (10.9)	2,489 (629)	2,366 (396)	82.4 (9.5)	3,451 (520)	4,369 (1,847)
drill tile	5.866	2.909	36.4 (10.5)	1,967 (316)	3,128 (666)	47.1 (12.5)	4,683 (759)	5,340 (1,106)
elephant complaint	3.807	3.041	23.8 (9.5)	3,597 (994)	2,831 (696)	81.2 (10.1)	3,863 (759)	1,474 (157)
lobster college	5.268	3.674	9.5 (6.6)	1,384 (125)	2,547 (502)	68.8 (12.0)	3,019 (592)	3,324 (925)
monkey chisel	3.807	3.245	19.0 (8.8)	2,069 (555)	2,221 (456)	50.0 (12.9)	2,529 (455)	3,784 (469)
octopus apartment	3.807	3.637	31.8 (10.2)	2,567 (526)	2,087 (381)	76.5 (10.6)	3,239 (628)	4,342 (1,065)
onion bus	3.807	3.602	9.1 (6.3)	1,441 (49)	1,863 (277)	50.5 (12.9)	3,461 (814)	3,158 (1,041)
piano cow	5.193	3.411	0.0 (0.0)	–	2,228 (322)	60.0 (13.1)	3,998 (775)	3,097 (968)
pig restaurant	7.383	3.641	28.6 (10.1)	2,121 (437)	1,965 (373)	80.0 (10.7)	2,686 (382)	2,371 (756)
rhinoceros cactus	4.500	2.274	45.0 (11.4)	4,060 (1,137)	1,982 (668)	70.6 (11.4)	4,268 (520)	3,944 (666)
scarf coconut	3.807	2.677	0.0 (0.0)	–	2,222 (314)	40.0 (13.1)	5,233 (1354)	4,438 (1,123)
seal viper	3.807	2.946	28.6 (10.1)	2,061 (432)	2,989 (667)	58.8 (12.3)	3,973 (702)	4,162 (1,078)
skunk tent	4.718	3.025	31.8 (10.2)	4,158 (1,118)	2,163 (464)	88.2 (8.1)	3,117 (335)	3,054 (1,517)
slug tree	7.092	3.553	38.1 (10.9)	3,674 (870)	2,009 (406)	68.8 (12.0)	3,602 (735)	2,927 (474)
snail shark	5.606	2.933	40.9 (10.7)	2,957 (679)	2,627 (490)	87.5 (8.5)	3,508 (295)	3,490 (1,817)
spider moth	6.344	2.802	81.0 (8.8)	2,066 (321)	1,510 (333)	100.0 (0.0)	3,428 (546)	–
steam saxophone	3.807	2.889	4.5 (4.5)	2,210 (0)	2,311 (315)	53.3 (13.3)	3,380 (807)	3,030 (633)
tent shirt	5.814	3.513	61.9 (10.9)	2,864 (493)	2,015 (581)	94.1 (5.9)	3,137 (285)	3,608 (0)
termite frog	3.807	2.808	50.0 (10.9)	2,555 (577)	2,677 (643)	75.0 (11.2)	2,955 (419)	2,791 (261)
whale knife	5.347	3.472	61.9 (10.9)	2,107 (470)	1,386 (249)	81.2 (10.1)	2,621 (399)	4,695 (1,461)
whiskey giraffe	3.807	2.954	0.0 (0.0)	–	2,112 (359)	58.8 (12.3)	3,866 (919)	3,617 (648)

Standard errors are in parentheses

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. doi:10.1016/j.jml.2007.12.005

Barsalou, L. W., Santos, A., Simmons, W. K., & Wilson, C. D. (2008). Language and simulation in conceptual processing. In M. de Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 245–283). Oxford: Oxford University Press.

BNC Consortium. (2001). British National Corpus, Version 2 (BNC World). Retrieved from www.natcorp.ox.ac.uk

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

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977–990. doi:10.3758/BRM.41.4.977

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

Connell, L., & Lynott, D. (2011). Modality switching costs emerge in concept creation as well as retrieval. *Cognitive Science, 35,* 763–778. doi:10.1111/j.1551-6709.2010.01168.x

Connell, L., & Lynott, D. (2013). *Principles of representation: Why you can't represent the same concept twice*. Manuscript submitted for publication.

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

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11,* 11–15. doi:10.1111/1467-8721.00158

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.

Jones, L. L., & Golonka, S. (2012). Different influences on lexical priming for integrative, thematic, and taxonomic relations. *Frontiers in Human Neuroscience, 6,* 205. doi:10.3389/fnhum.2012.00205

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. doi:10.1037/0033-295X.104.2.211

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. (2011). Symbol interdependency in symbolic and embodied cognition. *Topics in Cognitive Science, 3,* 273–302.

Louwerse, M., & Connell, L. (2011). A taste of words: Linguistic context and perceptual simulation predict the modality of words. *Cognitive Science, 35,* 381–398. doi:10.1111/j.1551-6709.2010.01157.x

Louwerse, M., & Hutchinson, S. (2012). Neurological evidence linguistic processes precede perceptual simulation in conceptual processing. *Frontiers in Cognitive Science, 3,* 385. doi:10.3389/fpsyg.2012.00385

Louwerse, M., & Jeuniaux, P. (2008). Language comprehension is both embodied and symbolic. In M. de Vega, A. M. Glenberg, & A. C. Graesser (Eds.), *Symbols and embodiment: Debates on meaning and cognition* (pp. 309–326). Oxford: Oxford University Press.

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

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

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14,* 119–124. doi:10.1111/1467-9280.t01-1-01429

Solomon, K. O., & Barsalou, L. W. (2004). Perceptual simulation in property verification. *Memory & Cognition, 32,* 244–259. doi:10.3758/BF03196856

Tagalakis, G., & Keane, M. T. (2006). Familiarity and relational preference in the understanding of noun–noun compounds. *Memory & Cognition, 34,* 1285–1297. doi:10.3758/BF03193272