Research Report

Conceptual Combination During Sentence Comprehension

Evidence for Compositional Processes

David Swinney,[1] Tracy Love,[1,2] Matthew Walenski,[1] and Edward E. Smith[3]

[1]*University of California, San Diego;* [2]*School of Speech, Language and Hearing Sciences, San Diego State University; and* [3]*Columbia University*

ABSTRACT—***This experiment examined the time course of integration of modifier-noun (conceptual) combinations during auditory sentence comprehension using cross-modal lexical priming. The study revealed that during ongoing comprehension, there is initial activation of features of the noun prior to activation of (emergent) features of the entire conceptual combination. These results support compositionality in conceptual combination; that is, they indicate that features of the individual words constituting a conceptual combination are activated prior to combination of the words into a new concept.***

A fundamental issue in language and cognition concerns how individual concepts (such as those denoted by single words) are combined to form larger, complex concepts, a process referred to as *conceptual combination*. Any account of conceptual combination must address the issue of compositionality. In the strong sense of this notion, a combinatorial process is compositional if it begins with the meanings of the components and operates only on these components. Smith, Osherson, Rips, and Keane (1988) proposed such a compositional account of adjective-noun combinations, and their model's reliance on strong compositionality led to numerous problems in accounting for a variety of data (e.g., Hampton & Springer, 1989; Medin & Shoben, 1988; Murphy, 1990; for a review, see Murphy, 2002). However, a weak version of compositionality may still be viable. According to such an account, the combinatorial process begins with meanings of the constituent concepts, but other information may enter the combinatorial process at a later point.

Address correspondence to Tracy Love, San Diego State University, School of Speech, Language and Hearing Sciences, 5500 Campanile Dr., San Diego, CA 92182-1518, e-mail: tlove@mail.sdsu.edu.

Some surprising results by Springer and Murphy (1992) appear to provide decisive evidence against even a weak version of compositionality. These authors focused on emergent features—properties that emerge in a combination and that are not properties of its component parts—and supposedly demonstrated that some emergent features are activated prior to features of the constituents; this result contradicts weak compositionality (henceforth, just "compositionality"). In their first study, Springer and Murphy presented adjective-noun pairs, such as "boiled celery," followed by probe properties that were true or false of the noun in isolation (e.g., "is green," "is blue") or true or false of the combination ("is soft," "is crisp"; "is soft" is an emergent property of the combination). Verification for emergent properties of combinations was as fast as, or faster than, verification for properties of the nouns themselves. This implied that properties of the combinations were available at least as quickly as properties of the nouns, contrary to a compositional account.

Springer and Murphy (1992) conducted a second study in which the same phrases were presented via rapid serial visual presentation. On each trial, the adjective appeared first, for 150 ms, and was then replaced by the noun, also for 150 ms; finally, the probe property (e.g., "is green") was presented until subjects responded. Again, verification reaction times (RTs) for emergent properties of the combinations were at least as fast as verification RTs for the nouns. In a third study, the probe property that subjects made decisions about (e.g., "green") was presented for 1,000 ms prior to presentation of the modifier and noun. Verification of properties for the conceptual combinations was again faster than verification of properties for the nouns.

These surprising findings, along with the conclusion that compositionality must be abandoned, require serious consideration. Our concerns center on methodology. The verification tasks Springer and Murphy (1992) employed may require conscious processing that follows comprehension and results in the meaning of the entire phrase (and hence its properties) being more available than the meaning of the constituents (see, e.g., Foss & Swinney, 1973). The fact that verification RTs in Springer and Murphy's studies were relatively long supports this possibility (RTs averaged 1,411 ms in Study 1 and more than 2,000 ms in Studies 2 and 3). In addition, in a pilot study examining the experimental stimuli that Springer and Murphy used, we found that for half of the adjectives employed in their conceptual combinations, our subjects provided associates that were either identical or highly related to the emergent properties tested for the same combination. For example, "hard," which was the emergent feature for the combination "stale buns," was an associate our pilot subjects commonly supplied for "stale" presented in isolation. Thus, it is possible that in some cases, the emergent property of the entire phrase was inadvertently activated by the initial adjectives in Springer and Murphy's Studies 1 and 2.[1]

For these reasons, we thought it important to examine the question of compositionality via an on-line task that is sensitive to the time course of information activation during sentence comprehension. We used cross-modal lexical priming (CMLP; Nicol, Swinney, Love, & Hald, 2006; Swinney, Onifer, Prather, & Hirshkowitz, 1979), which has been demonstrated to be sensitive to activation of word meanings in a sentence within roughly 150 to 300 ms. In a CMLP study, each sentence is presented auditorily. At some point while comprehension of the sentence is ongoing, a visual letter string appears on a computer screen, and the subject makes a lexical decision (word/nonword) about that string. The lexical decision is speeded (primed) when the letter string is semantically related to a word that has just been processed in the sentence. Priming is determined by comparing RT to a semantically related probe with RT to an unrelated probe matched for a priori RT. When priming is found during the time course of sentence understanding, it is evidence that a related word has been processed in the sentence. If noun meanings are activated prior to activation of modifier-noun combinations, as compositionality demands, this CMLP task would have the temporal sensitivity to detect this sequence of activations.

PRELIMINARY STUDIES

We designed two preliminary studies to create materials for the on-line study. One purpose of these preliminary studies was to identify properties that were independently associated with potential modifiers, with potential nouns, and with their combinations, and that could be used as CMLP target probes. In addition, we needed to identify control probes that were matched to the target probes on a priori RT.

[1]Springer and Murphy kindly shared their experimental materials with us, and we greatly appreciate their openness.

Preliminary Study 1

For this study, we created a pool of 96 modifier-noun pairs, each with associates (properties) for the noun and the combined phrase. These 96 conceptual combinations included the 20 items used by Springer and Murphy (1992). Fifteen undergraduates at the University of California, San Diego (UCSD), rated the degree to which each noun associate and combination associate was related to each of the three elements of interest (using a 5-point scale: 1 = *not related*, 5 = *highly related*): the modifier, the noun, and the modifier-noun combination. This procedure resulted in six ratings for each potential combination. Our goal was to choose items for which the noun associate and the combination associate were each highly related to the appropriate element (3 or greater on the 5-point scale) and unrelated to the other two elements (less than 3 on the scale). Thirty-two items met these criteria and were chosen as stimuli for the on-line experiment.

For these final items, the noun associates' average rating of relatedness to the corresponding nouns was 3.6 ($SD = 0.37$); ratings of relatedness to both the combinations ($M = 2.1$, $SD = 0.62$) and the modifiers ($M = 1.8$, $SD = 0.61$) were lower. Noun associates' relatedness to the nouns was significantly greater than their relatedness to the combinations, $t(31) = 10.42$, $p < .0001$, $d = 2.9$, or to the modifiers, $t(31) = 12.56$, $p < .0001$, $d = 3.57$. The combination associates' average rating of relatedness to the corresponding combinations was 3.71 ($SD = 0.48$); ratings of relatedness to the nouns ($M = 2.1$, $SD = 0.58$) and to the modifiers ($M = 2.23$, $SD = 0.63$) were lower. The combination associates' relatedness to the combinations was significantly greater than their relatedness to the nouns, $t(31) = 10.66$, $p < .0001$, $d = 3.02$, or to the modifiers, $t(31) = 10.71$, $p < .0001$, $d = 2.7$.

Preliminary Study 2

A separate group of 39 UCSD undergraduates participated in a lexical decision task intended to find matched control items for the 32 noun associates and 32 combination associates chosen in Preliminary Study 1. For each of the 64 already-chosen associates, we selected six words that were similar in frequency, length, form class, and animacy (but unrelated semantically or associatively to any component of the corresponding combination). These candidate control words were presented in random order (along with the selected associates) in an isolated lexical decision task. Using these data, we chose an RT-matched control word for each associate. The mean RTs were 595 ms ($SD = 60$) for the associate probes and 589 ms ($SD = 52$) for the matched control probes, $t(104) = -0.54$, $p = .593$, $d = 0.1$.

EXPERIMENT: CROSS-MODAL LEXICAL PRIMING

The main experiment tested for activation of both noun and combination (emergent) meanings during ongoing sentence comprehension. The test points were at the onset of the noun in the combination and the onset of the word following that noun.

Method

Subjects and Stimuli

Ninety-seven UCSD undergraduates participated. All subjects were right-handed, native English speakers.

The 32 combinations chosen from the preliminary tests were placed in sentence formats, as in the following example: "The executive delivered a speech proposing the *peeled banana* as the company's logo." Sentences ranged between 13 and 22 words in length (varied to prevent anticipation of target probes). The combination always appeared in the verb phrase of the sentence, roughly between the middle of the sentence and two thirds of the way through it. Each experimental sentence was paired with a noun associate (e.g., "yellow" in the case of the example just given) and a combination associate (e.g., "white"), which were employed as related target probes, as well as with two RT-matched control probes.[2] These four target probes were used in the concurrent lexical decision task.

For the lexical decision task, probes were presented at two points during each sentence—at the onset of the noun in the combination and at the onset of the next word (usually a preposition). The occurrence of the first target probe was timed to be coincident with the onset of the key word in the sentence whose activation was being tested (the noun), given the evidence that by the time a response to the lexical decision probe is made (typically after 600–700 ms), the auditory word with which it co-occurred has typically been processed (e.g., Swinney, 1990; Zwitserlood, 1989).

The four target probes and two test points were counterbalanced across eight presentation conditions. Forty-two additional sentences were constructed as filler materials. These sentences were similar in structure to the test sentences, but the target probes were presented at a range of positions throughout the fillers, to prevent subjects from predicting when the target probe would appear. The 32 test sentences and 32 of the filler sentences were pseudorandomized into a script, such that no more than 3 test or 3 filler sentences occurred in a row; all of these filler sentences were paired with nonword probes. The remaining 10 filler sentences constituted practice items, for which half the target probes were words and half were nonwords. Thus, the overall design for the main variables was 2 (type of target probe: noun property vs. emergent property of the combination) × 2 (test point: onset of the noun, or Test Point 1, vs. onset of the next word, or Test Point 2) × 2 (relatedness of target probe: related vs. unrelated). Counterbalancing across eight between-subjects materials conditions ensured that every subject heard or saw equal numbers of items from all the experimental conditions, but never heard or saw the same item (sentence or target probe) more than once.

[2]All materials can be found on the Web at http://psy.ucsd.edu/~tlove/concombo_stims.htm.

Finally, six comprehension questions were placed into the script at semirandomly spaced intervals. The requirement to respond to these auditorily presented questions ensured that subjects were attending to and understanding the experimental stimuli.

Procedure

Subjects were tested in individual, darkened, semisoundproof booths. They were seated in front of a computer screen and wore headphones, over which they heard the 74 sentences in the study. Target probes appeared centered on the computer display and remained on the screen for 300 ms. Subjects were instructed to press a button to indicate whether each target letter string was a real word (right index finger) or a nonword (left index finger). They were instructed to make their decision as quickly and accurately as possible. The interval between sentences was 2 s except when a sentence was followed by a comprehension question. Participants wrote untimed responses to the comprehension questions on an answer sheet, which had numbers for 17 questions even though only 6 actually occurred (so that subjects could not predict either the end of the experiment or the last comprehension question).

Results

RTs for correct responses were averaged for each subject for each experimental condition and submitted to subjects (F_1) and items (F_2) analyses of variance. Data from any subject who made more than 3 incorrect responses on the critical probes or more than 15 errors overall were excluded from analyses, as were data from any subject who made more than 3 errors on the comprehension questions. Data from 13 subjects were excluded on these bases, leaving 84 subjects in the analysis. Data from two items were deleted because of technical problems in probe presentation throughout the study. Including all the sources of error, the overall error rate was 4.8%. Mean RTs for all conditions are displayed in Table 1.[3]

The RTs demonstrated significant priming for target probes related to the noun of the combination (compared with unrelated control probes) at a probe position coincident with the onset of that head noun (Test Point 1 in Table 1), $F_1(1, 83) = 6.5, p = .01, d = 3.7$; $F_2(1, 31) = 4.1, p = .04$. These data support the view that noun properties have been activated by the time a noun is heard. However, at this same test point, there was no priming for emergent properties of the combination, $F_1(1, 83) = 1.9, p = .17, d = 1.97$; $F_2(1, 31) = 1.1, p = .29$. At a test point coincident with the onset of the word following the noun (Test Point 2 in Table 1), there was significant priming for target probes related to the emergent meaning of the combination (compared with unrelated control probes), $F_1(1, 83) = 6.1, p = .01, d = 3.57$; $F_2(1, 31) = 5.1, p = .02$. The interaction between test point (1 vs. 2) and relatedness of the target (related vs. unrelated) was significant for the combination probes, $F_1(1, 83) = 7.7,$

[3]To achieve sufficient power in these types of mixed factorial designs, it is important to have a minimum of 10 subjects per condition. For this study, this requirement meant we needed at least 80 subjects across the eight conditions.

TABLE 1
Mean Reaction Times and Priming Effects for Noun and Combination Probes (in Milliseconds)

Probe	Test Point 1	Test Point 2
Noun associate		
Related	602	607
Control	634	621
Priming	**32***	**14**
Combination associate		
Related	650	634
Control	632	671
Priming	**−18**	**37***

Note. The following is an example of a stimulus sentence, with test points indicated by superscripts: ''The executive delivered a speech proposing the peeled*[1] banana*[2] as the company's new logo.'' For this sentence, the noun associate is ''yellow,'' and the combination associate is ''white.'' Priming was calculated by subtracting reaction time to related probes from reaction time to control probes; these results are in boldface in the table. Asterisks indicate significant priming ($p < .05$).

$p = .005$; $F_2(1, 31) = 5.4, p = .02$. At Test Point 2, there was no longer significant activation for the noun properties (see Table 1), $F_1(1, 83) = 1.5, p = .22, d = 1.7$; $F_2(1, 31) = 2.6, p = .10$, although the interaction between test point and relatedness of the target was not significant for the noun probes, $F_1(1, 83) = 1.1, p = .30$; $F_2(1, 31) < 1$.

Discussion

These results clearly demonstrate that properties of the head noun are activated before those of the conceptual combination; these findings are consistent with an account of conceptual combination that starts with compositional processes, but not with the data and conclusions of Springer and Murphy (1992). We note in particular that there is clear evidence that activation of the properties related to the combination is later than activation of the properties related to the noun: This conclusion is supported by both the absolute priming effects found at Test Points 1 and 2 (see Table 1) and the interaction between test point and probe relatedness (for the combination probes). The fact that the noun properties were significantly activated (early) at Test Point 1 and not significantly activated at Test Point 2 suggests that decay of such activations was under way by the second test point, but that the activations had not yet fully dissipated (as indicated by the nonsignificant interaction).[4]

[4]In a follow-up experiment similar to the one reported in this article, the test point for the noun property occurred even later, roughly two words after the noun occurred in the sentence; there was no priming effect (−4 ms), which suggests complete decay of the noun-property activation.

Acknowledgments—The research reported in this article was supported by grants from the National Institutes of Health (DC02984, DC00494) and the National Institute on Aging (AG17586). The authors wish to thank Sophia Bagdasaryan, Maura Pilotti, and Lisa Vance Trup for their contribution to this work. David Swinney passed away on April 14, 2006. Our friend and colleague will be greatly missed.

REFERENCES

Foss, D., & Swinney, D. (1973). On the psychological reality of the phoneme: Perception, identification, and consciousness. *Journal of Verbal Learning and Verbal Behavior*, *12*, 246–257.

Hampton, J.A., & Springer, K. (1989, November). *Long speeches are boring: Verifying properties of conjunctive concepts*. Paper presented at the annual meeting of the Psychonomic Society, Atlanta, GA.

Medin, D.L., & Shoben, E.J. (1988). Context and structure in conceptual combination. *Cognitive Psychology*, *20*, 58–190.

Murphy, G.L. (1990). Noun phrase interpretation and conceptual combination. *Journal of Memory and Language*, *29*, 259–288.

Murphy, G.L. (2002). Conceptual combination. In *The big book of concepts* (pp. 443–475). Cambridge, MA: MIT Press.

Nicol, J., Swinney, D., Love, T., & Hald, L. (2006). The on-line study of sentence comprehension: An examination of dual task paradigms. *Journal of Psycholinguistic Research*, *35*, 215–232.

Smith, E.E., Osherson, D.N., Rips, L.J., & Keane, M. (1988). Combining prototypes: A modification model. *Cognitive Science*, *12*, 485–527.

Springer, K., & Murphy, G.L. (1992). Feature availability in conceptual combination. *Psychological Science*, *3*, 111–117.

Swinney, D. (1990). The resolution of indeterminacy during language comprehension: Evidence from lexical and structural processing. In G.B. Simpson (Ed.), *Understanding word and sentence* (pp. 367–385). Amsterdam: North-Holland.

Swinney, D., Onifer, W., Prather, P., & Hirshkowitz, M. (1979). Semantic facilitation across sensory modalities in the processing of individual words and sentences. *Memory & Cognition*, *7*, 159–165.

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*, 25–64.

(RECEIVED 7/1/05; REVISION ACCEPTED 10/6/06; FINAL MATERIALS RECEIVED 10/9/06)