Title
The Role of Knowledge Support in Creating Noun-Noun Compounds

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 25(25)

ISSN
1069-7977

Authors
Lynott, Dermot
Keane, Mark

Publication Date
2003

Peer reviewed

The Role of Knowledge Support in Creating Noun-Noun Compounds

Dermot Lynott (dermot.lynott@ucd.ie)
Mark Keane (mark.keane@ucd.ie)

Department of Computer Science, University College Dublin, Dublin 4, Ireland

Abstract

Conceptual combination research has largely concerned itself with the comprehension of novel nominal compounds, while the production of novel compounds has long been neglected in both the empirical and computational literature. In this paper, we advance a new paradigm for examining the creation of novel noun-noun compounds. Two experiments are reported, showing that the level of knowledge support (i.e. the familiarity) of the object descriptions affect compound production in two ways. First, people exhibit greater agreement on what compounds to produce when the object description has high knowledge support. Second, people also have increased confidence in the goodness of their compounds with high knowledge support. To conclude, we discuss some of the issues that arise out of this new work and outline a model of the compound creation process.

Introduction

Holiday drug mule, *soccer mom*, *laptop computer*, *trash cookies* and many more nominal compounds illustrate the creativity and pervasiveness of nominal compounds in everyday language use. These compounds are a microcosm of the generative nature of natural language, in which we see new meanings being created from the re-combination of words in syntactically well-formed phrases. It is the need to understand this generativity of language that has motivated several decades of research into nominal compounds (e.g., Clark & Hecht, 1983; Costello & Keane, 2000; Gagné & Shoben, 1997; Gerrig & Murphy, 1992; Hampton, 1987; Levi; 1978; Wisniewski, 1996). However, this research effort has been somewhat unbalanced in that it has mainly concentrated on the comprehension of novel compounds, rather than on their production. This oversight, focusing on how people understand compounds rather than on how they coin new ones, applies to both the empirical testing and actual modelling of nominal compound use.

There is a large empirical literature on the comprehension of lexicalised (e.g., Levi, 1978; Marsh, 1984; Quirk, Greenbaum & Svartik, 1985) and novel nominal compounds (so-called conceptual combination; see e.g., Coolen, Van Jaarsveld & Schreuder, 1991; Costello & Keane, 2000, 2001; Gagne, 2000, 2001, 2002; Gerrig & Bortfeld, 1999; Lynott & Keane, 2002; Murphy, 1990; Smith, Osherson, Rips & Keane, 1988, Wisniewski, 1997). In contrast, work examining the production of nominal compounds has been quite patchy and disjointed (Clark, Hecht & Mulford, 1986; Levi, 1978, Downing, 1977, Windsor, 1993). This may be because the production literature to date tends to be mainly concerned with issues other than compound creation *per se*. For example, Clark et al. and Windsor were primarily concerned with compound production as an aspect of language development, while Downing and Levi focused on debating the existence of a taxonomy of relations within noun-noun compounds. It is also very hard to generalise over these studies because of their widely differing methodologies, from context-free production to picture naming to linguistic analysis.

The ample modelling literature on conceptual combination has also tended to focus on comprehension rather than on production. Three models deserve mention. CARIN (Gagné & Shoben, 1997) proposes that the distribution of relations usually found with the compound's constituents can determine the correct relation / interpretation. The Dual-Process model (Wisniewski, 1996) proposes that the level of similarity between the constituent nouns activates one of two possible processes of comprehension: scenario creation (which produces relation-based interpretations by slot filling) and alignment (which produces property-based interpretations by a type of analogical mapping). Finally, Costello and Keane's (1997, 2000) C^3 model proposes that, at a computational level, there is a space of possible meanings for a novel compound from which one meaning is selected during comprehension by the application of three constraints (informativeness, plausibility and diagnosticity).

Although these three models offer a rich diversity of mechanisms to model the creation of novel nominal compounds, no one model seems to be directly applicable. At its simplest, one might argue that the production process is the reverse of comprehension, taking an interpretation as input and outputting possible nominal compounds. However, this proposal does not stand up to close scrutiny. Intuitively, production seems more complex than this reversion, as one has to select a set of terms that best convey the intended meaning in many future contexts, rather than simply extracting a possible meaning from two terms presented in a single, current context. One would expect there to be common component processes between comprehension and production, but exactly what these might be is unclear.

Of course, the paucity of the empirical literature means that any model is mere speculation until there is some hard evidence to model. To fill this important gap, we propose a new paradigm for the empirical investigation of novel compound creation. Towards the end of the paper, we will return with this evidence to the issue of computational modelling.

A Paradigm for Compound Production

Vendler (1967) describes the process of nominalization as "packing a sentence into a bundle that fits into other sentences". We have developed a novel experimental paradigm for studying compound production that mimics just such a process. In this paradigm, we give people a sentence describing an object, and ask them to produce a compound to convey the information in this description. Computationally, one can cast this problem as finding the minimal subset of terms whose meaning will accurately and unambiguously convey the given description. Consider a number of cases to illustrate this proposal. If the object description is the sentence:

1a. A wine that is made from grapes and contains alcohol.

then the minimal set of terms to capture it is likely to be just one word "wine". Prior knowledge tells us that wine is typically made from grapes and is alcoholic. So, "wine" will accurately and unambiguously convey the description. However, if the description were changed to 1b, a single term will no longer suffice.

1b. A wine that is made from apricots and contains alcohol.

Using "wine" alone will not convey the meaning accurately. Indeed, the minimal set of terms is probably now two; namely, "apricot wine". "Apricot" is needed to overwrite the default shared knowledge that wine is made from grapes and replace it with the knowledge that this wine is made from apricots. However, even "apricot wine" will only work if people readily understand the structure an "X wine" to be "a wine made from X" rather than "a wine in X shaped bottles" or "a dessert wine served with Xs". So, it is not just the case that a term can be ignored if prior knowledge is available. It must also be the case that other competing interpretations from prior knowledge should not be suggested by the chosen terms.

Knowledge or Word Order or Both?

The above analysis indicates what needs to be computed when creating nominal compounds but it says less about what allows this computation to be achieved. At least three factors are likely to play role:

- *pragmatic knowledge* about effective communication
- *available world knowledge* about the described objects
- *syntactic knowledge,* or *word order* in the object description

From the pragmatic perspective, when people produce a novel nominal compound, either in speech or in writing, they are forced to make decisions regarding their choice of words, knowing that the resulting compound must be comprehensible to the listener or reader[1]. The pragmatics literature suggests that shared world knowledge is likely to have a crucial impact on this choice of words (see Grice, 1975; Sperber & Wilson, 1986). Indeed, in their C^3 model of comprehension, Costello and Keane (2000) employ such pragmatic principles (e.g. informativeness) to show how supporting world knowledge is central to arriving at a plausible meaning for a novel compound. In the studies reported here, we attempt to manipulate the shared (or available) knowledge given in the sentence description to look at its effect on the compounds produced. Intuitively, if there is supporting knowledge then people may be more confident that the compound they produce will convey its intended meaning (see Experiments 1 and 2).

Although it is difficult to quantify levels of knowledge support, the support for a particular object description certainly relies on our prior knowledge and experience of the object. In other words, knowledge support is about how normal a description seems with respect to what we know about the world.

Furthermore, when supporting world knowledge is available, it is possible that it will reduce variability in the compounds people produce – i.e. people may produce fewer unique compounds for an object description that is supported by world knowledge (see Experiments 1 and 2).

From a syntactic perspective, it also seems feasible that the selection of appropriate words from the sentence may be informed by its syntactic structure. Consider the following:

2a A note left for a milkman on a pole.
2b A note left on a pole for a milkman.

If people were being solely directed by the syntactic position of the words in the sentence then we would expect different choices of compound (*milkman note* for 2a and *pole note* for 2b). However, if the primary

[1] This is not to suggest that these decisions are conscious, merely that the compound producer is implicitly aware of the need to provide an informative compound.

influence is from the world knowledge brought to bear, and not from the word order in the description, then alternative syntactic forms should have little effect (see mainly Experiment 2).

Outline of Experiments

In our first experiment, we used different object descriptions to manipulate the amount of supporting world knowledge available and asked people to produce labels for the objects being described. Experiment 1 is mainly focused on the world knowledge issues, while Experiment 2 looks at whether altering the word order of the object descriptions affects peoples' production of novel compounds.

Pre-Test

A pre-test was required before the main experiment in order to confirm a reliable difference in levels of knowledge support between High Support and Low Support items.

21 native English speakers were given booklets containing 44 object descriptions. Participants were asked to rate how normal each of the descriptions were (1 being not normal at all and 7 being completely normal), and they were free to decide what normal meant to them.

From participant responses, 44 pairs of object descriptions were selected with one description being of High Support and one being of Low Support. The mean rating for High Support items was 5.451, while the mean for Low Support items was 3.563. This difference was reliable both by subjects ($F1(1, 20) = 312.640$, $MSe = 809.994$, $p < 0.0001$) and by items ($F2(1, 43) = 126.4$, $MSe = 791.217$, $p < 0.0001$).

Experiment 1

Having confirmed differing levels of knowledge support, we then used this set of object descriptions to examine the effects of knowledge support on how people produce novel noun-noun compounds. Two main types of object descriptions were used: High Support (HS) and Low Support (LS):

3a A disease in a swamp caused by flies (HS)
3b A disease in a school caused by flies (LS)

World knowledge suggests that *swamps* are more associated with *flies* and *disease* than *schools* are and so 3a is categorised as being a High Support item while 3b is categorised as Low Support.

In this experiment, we looked at three measures that might be influenced by this variable:

1. *Variability of compound production* – This refers to the number of unique compounds produced for a given object description. We expected increased knowledge support to reduce the variability of compounds produced – i.e., increased knowledge support would give rise to greater agreement between participants.
2. *Frequency of compound production* – This refers to the number of times a compound is produced for a given description. We might expect to see a more unbalanced distribution of frequencies for High Support descriptions, with particular compounds being produced with great frequency and the remainder with only very low frequencies. Low Support descriptions, on the other hand, might provide a more even distribution, with several compounds being produced with relatively high frequency.
3. *Confidence ratings* – This refers to how well participants feel their compounds adequately convey the information contained in the sentence description. We expected High Support items to lead people to give higher confidence ratings. If people produce compounds to describe something that closely matches what they know of the world, then this should make them more confident about the goodness of their compound.

Method

Materials Eighty-eight sentences were used in the experiment: 44 of which were High Support and 44 Low Support versions of the sentence. Each sentence had a subject and two objects that were linked by various relations e.g. Subject *made* of Object 1 *located in* Object 2 (e.g., see sentences 2a, 2b, 3a, and 3b). The High Support and Low Support versions of a sentence were differentiated by the change of one word (either Object 1 or Object 2).

A set of 15 pairs of filler items was also created. Filler sentences had the same syntactic form but were tautological in nature; for example, *A wine that is made from grapes and contains alcohol*. The filler items could be adequately described using only one word (e.g., wine) whereas test items would require two or more words.

Procedure Participants were 24 Computer Science undergraduates from University College Dublin. All were volunteers and were native speakers of English. Two participants were removed prior to analysis for not completing the task. Participants were given booklets of sentence descriptions and were asked to produce a description of one or two words that would best describe the object given in each description. They were also asked to rate, on a seven-point scale, how well their new phrase conveyed same information as the sentence description (1 being very good and 7 being very poor). Participants were not asked explicitly to produce noun-noun compounds, only to produce responses of one or two words. However, the examples that were given in the instructions were all noun-noun responses.

Two sets of booklets were prepared with a random selection of object descriptions but with equal numbers of high- and low-support items. The order of descriptions was randomised for each participant. Support was a within-subject factor with no participant seeing both the high support and low support versions of the same sentence.

Results & Discussion Approximately 75% of all responses were noun-noun compounds and the results relate to these compounds.

Using a one-tailed Mann-Whitney test we observed a main effect of knowledge support on compound variability, with High-Support descriptions giving rise to fewer unique compounds than Low-Support descriptions; $U = 610.5$, $n1 = 44$, $n2 = 44$, $p < 0.001$. The mean number of unique compounds produced for High Support descriptions was 3.48 while the mean for Low Support descriptions was 4.34.

People were reliably more confident about compounds produced for highly supported object descriptions than those produced for low-supported descriptions. This finding was reliable both by-participants; $F1(1, 21) = 37.788$, $p < 0.001$, $MSe = 82.039$, and by-items; $F2(1, 43) = 30.421$, $p < 0.001$, $MSe = 80.620$. On the 7-point scale, the mean rating for high-support items was 2.945, while the mean for low-support items was 3.598. There was also a reliable correlation between the frequency of production of a compound and the confidence ratings it received ($p < 0.005$, $r = 0.163$, $N = 350$). In other words, compounds that were produced the most often tended to have the highest confidence ratings. Finally, we found no difference between the frequency of compound production for High Support descriptions and Low Support descriptions ($p > 0.1$).

Experiment 2

There is the possibility that the results of Experiment 1 were in some way confounded by the word order of the object descriptions i.e. the order in which the concepts appeared in the description. Thus, the purpose of Experiment 2 is twofold. First, we can investigate whether the findings from Experiment 1 were being driven by the word order of the object descriptions or by the knowledge of the concepts themselves. Second, it provides a replication of Experiment 1 as it follows the same structure and task demands.

The question of whether nominal compounding is knowledge-driven or syntax-driven is an important one. If it is the case that the process is knowledge-driven, then we can make predictions regarding the production of novel compounds by focussing only on the concepts involved and not on how they have been described syntactically. On the other hand, if syntax or word order is the driving force behind compound production then it is factors such as sentential structure and word order that would need to be attended to first and foremost.

Experiment 2 uses the same base set of materials as Experiment 1, except that the word order of the descriptions has been altered. This allowed us to compare the response from both experiments to ascertain the effect of changing the word order on the production process. To construct this new set of materials Experiment 1 descriptions were changed by switching the positions of the first and second objects. The subject remained at the start of the description. If an object description in Experiment 1 were of the form 4a then it was re-written in the form 4b for Experiment 2 (see also examples 2a and 2b).

4a An X caused by a Y in a Z
4b An X in a Z caused by a Y

So although the constituents of the description remain the same, the word order has changed. In altering the word order in this way, six of the object descriptions became ambiguous or nonsensical. These descriptions were excluded from analyses.

Method

Materials 38 of our 44 original object description pairs were used along with 15 filler items. As with Experiment 1, the object descriptions were categorised as either High Support or Low Support. Two sets of booklets were prepared with a random selection of descriptions, with Support as a within-subject factor with no participant seeing both the High Support and Low Support versions of the same sentence.

Procedure Participants were 20 undergraduate students in computer science at University College Dublin. All were volunteers and were native speakers of English. Participants received booklets of descriptions whose order was randomised for each person. As with Experiment 1, participants were asked to rate on a scale of 1-7 how good their compound was in conveying the same information as the object description (1 being very good and 7 being very poor).

Results As with Experiment 1, approximately 75% of responses were noun-noun compounds and these were used in the analyses that follow. We found a replication of all of the main findings reported in Experiment 1.

The level of knowledge support had a reliable effect on the variability of compounds produced ($U = 557.5$, $n1 = 38$, $n2 = 38$, $p < 0.05$), with High Support descriptions giving rise to 2.947 compounds on average, compared to 3.421 for Low Support descriptions.

Again we found that the level of knowledge support affected participants' confidence ratings, with compounds produced for High Support descriptions receiving better scores. This was reliable both by participants ($F1(1, 19) = 4.67$, $p < 0.05$) and by items ($F2(1, 37) = 5.151$, $p < 0.05$). The frequency of compound production was not affected by the level of knowledge support ($p > 0.1$).

In order to compare participants' responses from both experiments we extracted the overlap of compounds that were produced in both experiments. We then examined correlations (Pearson's r) between the two sets for compound variability, frequency of compound production and mean confidence ratings for each compound. If word order in the object descriptions were driving the compound creation process, then we would expect non-significant correlations between the results of Experiments 1 and 2. However, this was not the case. The results obtained are more in line with a knowledge-based account of the production process. Significant correlations were obtained for compound variability ($p < 0.001$, $r = 0.375$, $N = 117$), confidence ratings ($p < 0.02$, $r = 0.19$, $N = 117$) and compounds' frequencies of production ($p < 0.001$, $r = 0.624$, $N = 117$).

General Discussion

Experiments 1 and 2 examined the effects of knowledge support and word order on the process of novel noun-noun compound production. We found that differing levels of knowledge support impact on the level of variability between participants' novel compounds and also on people's ratings of how well a novel compound adequately conveys information. However, the level of world knowledge support does not affect the frequency of production of specific compounds. Significant correlations between the responses of Experiments 1 and 2 suggest that it is the available knowledge associated with the concepts of an object description that drive the compound production process, and not the order in which the concepts occur in the description. The implication of these results is that the word order of object descriptions does not play a primary role when people create novel compounds. This has important ramifications when we come to consider how we might model the production process.

So why does increased knowledge support increase people's confidence in the compounds they generate and reduce the variability of compounds produced?

If we take a High Support object description "A uniform worn by a guard in a prison" we can see that there are strong associations between the words *uniform*, *guard* and *prison* and also that the notion of a guard in a prison wearing a uniform is not unfamiliar to us. In contrast, the Low Support description "A uniform worn by a guard in a school" does not have the same strength of association between all its constituents and nor is the notion of a guard at a school as familiar to most people as a guard at a prison. It appears that it is this incompatibility between Low Support descriptions and participants' prior knowledge that leads people to be less confident about the compounds they produce.

What's more, this increased familiarity for High Support descriptions also allows participants to converge on particular compounds - as evidenced by the lower number of unique compounds for High Support Descriptions in both Experiments 1 and 2. This suggests that increased knowledge support makes it easier for people to encapsulate information, so leading to less variation in the compounds produced.

Towards a Model of Compound Production

Experiments 1 and 2 give us but a small part of the picture of novel compound production, but it is important for us to consider how these processes might actually be carried out. What stages are involved in the production process and how might a model reflect the findings reported here?

Below we outline the steps involved in producing a novel compound for a given description, incorporating both local and world knowledge to arrive at a suitable compound. We refer to descriptions 1a (*A wine that is made from grapes and contains alcohol*) and 1b (*A wine that is made from apricots and contains alcohol*) in order to highlight how minor differences in object descriptions will impact on the on the production process and on the resultant compounds. The steps are:

1. Input the object description.
2. Isolate subject of the object description e.g. *wine*
3. Compare the knowledge about this term to the meaning of the object description. If the information matches (e.g., *wine* is made-from-grapes and contains-alcohol) then we have found an adequate label (as in 1a) and the production process stops. However, if the meanings do not match (as in 1b) then further processing is required.
4. We retrieve another object from the description (e.g. *apricot*) and ascertain if this object could fill some role in the subject term (w*ine*) to bring it closer to the meaning of the description. In this example, knowledge about *wine* contains the information that it is made-from-grapes. Since *apricot* can be substituted for grapes in this role (i.e. becomes made-from-*apricots*), we now have a new candidate, *apricot wine*, to consider.
5. Compare the meaning of the new candidate compound (e.g. *apricot wine*) to the meaning of the object description. If the information matches, then we have a compound that adequately summarises the object description and the production process can stop. If the information does not match, we must return to step 3.

With this process, we can cycle through different possible compounds until one is found that satisfactorily reflects the information contained in the description.

This account of the various stages involved in the production process at a computational level may be preliminary in nature, but it incorporates the basic elements required to produce novel compounds. We hope that together with the findings reported here and future empirical work, it will form the basis for a complete, cognitively-motivated model of novel compound production. Such a model should incorporate not only commonsense knowledge regarding objects and relations, and syntactic information of sentence and object descriptions, but should also have the capacity to address pragmatic considerations such as the informativeness of novel compounds and the extent to which knowledge might be shared between speakers. We intend the experimental paradigm we have outlined here, and the evidence shown for the primacy of knowledge support, to provide an important signpost for the direction of future work in this area.

Acknowledgments

This research was partly funded by a grant from the Irish Research Council for Science, Engineering and Technology.

References

Clark, E. V., Hecht, B. F., & Mulford, R. C. (1986). Coining complex compounds in English: affixes and word order in acquisition. *Linguistics, 24,* 7-29.

Coolen, R., Van Jaarsveld, H. J. V. & Schreuder, R. (1991). The interpretation of isolated novel nominal compounds. *Memory and Cognition*, 19, 341-352.

Costello & Keane, (1997). Polysemy in conceptual combination: Testing the Constraint Theory of Combination. *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, M. G. Shafto, P Langley (Eds.). Hillsdale, N.J.: Erlbaum. pp 137-142.

Costello, F. J. & Keane, M. T. (2000). Efficient creativity: Constraints on conceptual combination. *Cognitive Science*, 24, 2, 299-349.

Costello, F. J. & Keane, M. T. (2001). Testing two theories of conceptual combination: Alignment versus diagnosticity in the comprehension and production of combined concepts. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 27(1): 255-271.

Downing, P. (1977). On the creation and use of English Compounds. Language, 53 (4), 810-842.

Gagné, C. L., (2000). Relation-based versus property based combinations: A test of the CARIN theory and dual-process theory of conceptual combination. *Journal of Memory and Language,* 42, 365-389.

Gagné, C. L., (2001). Relation and lexical priming during the interpretation of noun-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 236-254.

Gagné, C. L., (in sub). Conceptual combination and discourse context: Effect of discourse context and modifier relation frequency on conceptual combination.

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.

Gerrig, R. J. & Bortfeld, H. (1999). Sense creation in and out of discourse contexts. *Journal of memory and Language*, 41, 457-468.

Gerrig, R. J. & Murphy, G. L., (1992). Contextual influences on the comprehension of complex concepts. *Language and Cognitive Processes*, 7, 205-230.

Grice, H. P. (1975). Logic and Conversation. In P. Cole and J. L. Morgan (Eds.), Syntax and Semantics (vol. 3): Speech Acts. New York: Academic Press.

Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory and Cognition, 15(1)*, 55-71

Levi, J. N. (1978). The Syntax and Semantics of Complex Nominals. New York: Academic Press.

Lynott, D. & Keane, M. T. (2002) The comprehension of novel noun-noun compounds: the influence of out-of-context interpretations on in-context understanding. *Proceedings of the Twenty-Fourth Annual Conference of the Cognitive Science Society* (p. 1020). Hillsdale, NJ: Erlbaum.

Marsh, E. (1984). A Computational analysis of complex noun phrases in navy messages. In *Proceedings of 10th International Conference on Computational Linguistics*, New York. NY.

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

Quirk, R., Greenbaum, S., Leech, G. & Svartik, J. (1985). Comprehensive Grammar of the English Language. Longman.

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

Sperber, D. & Wilson, D. (1986). Relevance: Communication and Cognition. Oxford: Blackwell.

Vendler, Z. (1967). Linguistics in Philosophy. Ithaca: Cornell University Press.

Windsor, J. (1993). The functions of novel word compounds. *Journal of Child Language*, 20, 119-138.

Wisniewski, E. J., (1996). Construal and Similarity in conceptual combination. *Journal of Memory and Language*, 35, 434-453.

Wisniewski, E. J., (1997). When concepts combine. *Psychonomic Bulletin and Review*, 4, 167-183.