

Experimental Methodologies. A Corpus-Based Approach

Anca PĂUNESCU*
Ana Maria DEMETRIAN**

<https://doi.org/10.52744/AUCSFLSA.2024.01.43>

Abstract

On the basis of data, we analyse two different pairs of syntactic patterns, the so-called dative alternation and particle placement of transitive phrasal verbs. Although it has sometimes been argued that only experimental data can contribute to studies of priming, the analysis shows that (a) the corpus-based results for datives are very similar to the experimental ones; (b) priming is also obtained for the verb-particle construction, a construction hitherto not explored in the priming literature and (c), most importantly, in line with much previous psycholinguistic and corpus-linguistic work, priming effects turn out to be strongly verb-specific such that some verbs are much more resistant or responsive to priming than others. As a variety of studies has shown, speakers tend to repeat syntactic structures they have just encountered (produced or comprehended) before.

Keywords: methodology; study; linguistics; priming; syntax; structure

The range of experimental methodologies has also been broadened considerably and now includes a wide variety of offline experimental paradigms such as sentence completion tasks and picture descriptions in dialogs. In addition, researchers have carried out online studies where priming effects were also measured in terms of production latencies.

While many studies have investigated the dative alternation and the active-passive alternation in English, more recent work has also looked at the equivalent constructions in Dutch as well as Dutch locative PP alternations, the order of auxiliary verb and past participle in Dutch as well as dative-accusative verb alternation in German (Scheepers & Corley, 2000), and the order of syntactic functions in Japanese (Yamashita *et al.*, 2002).

The currently most pressing issues concerning syntactic priming (many of which will also be addressed in the present approach) are the following:

* PhD, University of Craiova, Craiova, Romania

** PhD, University of Craiova, Craiova, Romania



- the duration of syntactic priming: on the one hand, Levelt and Kelter (1982) and Branigan *et al.* (1999) report that priming (in spoken and written production respectively) is fairly short-lived. On the other hand, other studies report priming effects across longer time interval or more intervening material (Bock & Griffin, 2000; Pickering *et al.*, 2000; Chang *et al.*, 2000).
- the directionality of syntactic priming: Branigan *et al.* (1995) discuss a variety of different studies which, taken all together, support the assumption that syntactic priming can operate from production to production (Bock, 1986; Bock & Loebell, 1990), from comprehension to comprehension (Branigan *et al.*, 1995 for an overview) and from comprehension to production (Branigan *et al.*, 2000; Bock, 2002).
- the grammatical characteristics of the priming verb: Pickering and Branigan (1998) found that (a) syntactic priming is stronger if the priming verb lemma and the target verb lemma are identical (compared to different lemmas in prime and target) and that (b) morphological differences between the priming verb and the target verb (in terms of tense, aspect and number) do not result in strongly varying priming strengths.
- the degree to which syntactic priming is asymmetric and construction-(pair)-specific: From a between-alternations perspective, Bock (1986) found stronger priming for the two syntactic frames involved in the dative alternation than for those involved in the active-passive alternation in English; a similar prominence of datives over transitives was found for English by Potter and Lombardi (1998) and for Dutch by Hartsuiker and Kolk (1998). In addition, from a within-alternation perspective, further asymmetries were sometimes obtained: Bock (1986) found there was stronger priming for ditransitives than for prepositional datives while Potter and Lombardi (1998) report the opposite (and Pickering *et al.*, 2002: 587 mention evidence for symmetric/ balanced priming).
- the degree to which syntactic priming is language-specific: Hartsuiker *et al.*, (2002) demonstrate syntactic priming from comprehending Spanish to producing English, Salamoura (2002: Exp. 2) demonstrates priming from Greek (L1) structures to English (L2) structures, and Gries and Wulff (in press) show that German learners of English as a foreign language exhibit priming in an English sentence completion task.

The present study is concerned with the issues raised above. However, it is different from most others in two respects. First, its main



point is that it goes beyond previous works by paying closer attention to the role individual verbs play for priming.

Recent studies demonstrated that different verbs exhibit differentially strong associations to particular syntactic patterns or, put differently, constructions. Although the experimental priming studies mentioned above did control for the frequencies of prime constructions and for item-specific effects (in terms of F1 and F2 statistics) and, thus, allowed for a clear confirmation of syntactic priming, there appear to be no studies at all which investigated to what degree, if any, the strength of priming effects is conditioned by the prime and target verbs.

The main issue of this study is, therefore, the question of whether particular verbs are more responsive or resistant to priming as target verbs such that, across many different prime verbs, they either have a tendency for a particular construction that overrides the prime structure or not.

More generally speaking, the present study takes into account the degree to which syntactic priming may be verb-specific. Second, contrary to most previous works, the present study is based on naturalistic corpus data rather than psycholinguistic experimentation. Given that the first studies reporting syntactic repetition were based on naturalistic data, it may appear somewhat surprising that so little corpus based work on priming has been conducted, especially since larger corpora and the software necessary for their analysis is so widely available.

This absence can probably partly be attributed to the fact that, although Bock's (1990) first example for what she later refers to as syntactic priming is drawn from naturalistic conversation, priming researchers such as Branigan and Pickering have argued against corpus-based approaches to priming by stating that:

there are several nonsyntactic factors which could lead to repetition. [...] Corpora have proved useful as a means of hypothesis generation, but unequivocal demonstrations of syntactic priming effects can only come from controlled experiments (Branigan et al., 1995: 492; Pickering & Branigan, 1999: 136).

In order to investigate syntactic priming corpus-linguistically, first all ditransitive constructions and all prepositional datives with *to* and *for* in the British component of the International Corpus of English (ICE-GB) were identified. Out of the 3793 cases, 790 had to be discarded for the priming analysis because they were the first or last construction either in one of the 500 corpus files or in a subtext of a corpus file, leaving 3003



prime-target pairs (i.e. subsequent constructions of either type) for the analysis. Each of these was then coded for a variety of variables:

- Medium: the medium in which prime and target occurred: spoken vs. written (automatically retrieved from the corpus files).
- CPrime and CTarget: the constructions of the first and the second of the two constructions constituting a prime-target pair: ditransitive vs. prepositional dativ (automatically retrieved from the annotated parse trees within the corpus files).
- CID: the fact whether the constructions in prime and target are identical: yes or no (this coding task was performed semi-automatically with a spreadsheet software applied to the output of the concordancing software).
- Distance: the distance in parsing units between the occurrence of prime and target within each subtext of each file as determined from the annotation of the corpus: 0, ≤ 1 , ≤ 2 , ≤ 3 , ≤ 4 , ≤ 5 , ≤ 6 , ≤ 7 , ≤ 8 , ≤ 9 , ≤ 10 , ≤ 15 , ≤ 20 , ≤ 25 and > 25 (a parsing unit is the basic structural unit of each corpus file; in the majority of cases it corresponds to a clause or sentence).
- VFormPrime and VFormTarget as well as VLemmaPrime and VLemmaTarget: the exact verb form and the verb lemma of each prime and target (the verb forms in both constructions were retrieved automatically from the corpus files, the lemmatization was done manually by myself).
- VFormID and VLemmaID: whether both constructions involved the same verb form and verb lemma: yes or no (this coding task was performed semi-automatically with a spreadsheet software).
- SpeakerID: whether in the spoken data both constructions were produced by the same speaker or not: yes or no (this coding task was performed semi-automatically with a spreadsheet software applied to the output of the concordancing software).

Not only has the corpus-based analysis of syntactic priming revealed significant priming effects for ditransitives and prepositional datives, the results are also strikingly similar to those of previous experimental studies in terms of strength of effects, the influence of morphological characteristics of the verbs, construction-specificity, directionality and distance effects (i.e. the time course of priming).

On the one hand, the results from the corpus-based investigation of priming have documented a clear priming effect for both the ditransitive and the prepositional dative, and many characteristics of the corpus-based priming effects clearly resemble those of previous experiments.



However, it has also provided evidence that offers a much more detailed perspective on this global effect by showing that priming effects are verb-specific: Individual verb associations to particular constructions as measured by distinctive collocation strength result in some verbs allowing priming in one particular direction more readily than others.

This tendency is masked by the overall priming effects, but taking into account individual verbs' behavior can provide a more precise picture of the processing of the verbs and the structures in which they are used. It illustrates the potential of this way of analysis for a different constructional alternation that has so far received little attention in the literature on priming, namely verb-particle constructions.

Many studies are concerned with active/passive and datives. As in Hartsuiker *et al.* (1999), however, the results indicate that priming effects can also be obtained for cases where the alternants consist of the same phrases in different orders. The general findings concerning particle placement are somewhat similar to those of the dative alternation. There is a general priming effect so that constructions are likely to be repeated at the next opportunity.

In addition, both the dative alternation and particle placement exhibit a similar strength of the priming effect, and in both cases one construction primes more strongly than the other.

Finally, there are slight effects of the verb form and the verb lemma which, although insignificant, are at least in the same direction as the analogous (experimental and corpus-based) effects for the datives; the same holds for the directionality of priming. While the results are in need of additional evidence, they provide *prima facie* evidence of structural priming for a construction hardly related to syntactic priming in previous work. But let us now see whether particle placement is subject to the same kind of verb-specificity effects as the dative alternation.

Branigan and Pickering (1999) argue against corpus-based approaches to syntactic priming, claiming that the priming effects obtained from naturalistic data may have to be attributed to nonsyntactic reasons. The non-syntactic factors they mention include „temporary switches to more formal registers at certain points in the interview [... or ...] the well-known facilitatory effects of repeating particular words” (Pickering & Branigan, 1999: 136). Also, they briefly refer to discourse-motivated syntactic repetitions. In the light of these potential points of critique, it is necessary to address how corpus data compare to experimental approaches in the analysis of priming.

It is true that, in general, experimental studies are in a better position to single out particular aspects of priming more easily than corpus-based studies, and the possibility to hold experimental conditions



constant across a variety of trials and (combinations of) conditions should not be underestimated. However, the exploratory benefits of corpus data are mentioned by Pickering and Branigan themselves, and from a different perspective, the controlled nature of experimental conditions also has some drawbacks.

First, the priming data are usually collected in a very narrowly defined and artificial setting. While this is desirable from the point of view of delimiting error variance, it does not allow generalizations of the role of register effects on syntactic priming – the corpus data, by contrast, allow for a multifactorial analysis of syntactic repetition in natural settings.

In addition, in their discussion of previous experimental approaches to priming, Hartsuiker and Kolk (1998: 148) criticize much previous work for not taking into consideration the overall frequencies of syntactic constructions, which – if not considered properly—may introduce frequency effects into the priming results. In the present approach, the corpus data allow for a natural computation of construction baseline frequencies.

Second, not all experimental studies managed to account for all potential explanatory factors. For example, Bock and Loebell's (1990) findings may be interpreted as evidence for the irrelevance of thematic utterance characteristics and that function words are irrelevant to priming until Hare and Goldberg's (1999) and Bencini *et al.*'s (2002) replications showed that this was not necessarily the case. Of course, this does not invalidate the experimental approach as such, but it points out that the number of factors to be taken into consideration is so high that it is not always possible to hold them all constant. Thus, including such confounding factors into a corpus-based evaluation may sometimes be a useful alternative.

Finally, by investigating syntactic priming from a corpus-based perspective, one can determine to what degree it plays a role for grammatical variation, i.e. the phenomenon that in a given discourse situation the speaker may have the choice between two truth-conditionally equivalent, nearly synonymous constructions (e.g. between the two dative constructions, active vs. passive, or the of-genitive vs. the s-genitive etc.). Including the priming effects into the research design may make it possible to increase the accuracy of predicting the construction the speaker will choose (subconsciously).

Apart from these general methodological arguments, some other more specific comments on Pickering and Branigan's non-syntactic factors are due because not all of these lend themselves to an explanation of the present results (and Pickering and Branigan do not provide



empirical evidence for their claims). For example, the fact that one of the two constructions may be predominant in a particular register is taken into account here since (a) the corpus data cover a wide variety of registers and (b) the medium (speaking vs. writing) was included into the analysis. Note also that neither alternation investigated is inherently related to a particular level of formality, nor explaining the frequent cases of syntactic priming by hundreds of sudden register/formality changes does not seem very plausible.

Similarly, the effects cannot be straightforwardly reduced to, say, the givenness or semantic characteristics of the direct object's referent: First, both datives have information structure properties (Thompson, 1990), so why should only one result in priming in corpora? Second, one might suggest that the slight priming prominence of the verb-particle construction with a VP-final particle is due to the fact that this construction is associated with a given referent of the direct object (Gries, 2003): Once the referent of the direct object has been introduced, the verb particle constructions in the subsequent discourse will place it before the particle.

However, Gries (2003:120ff) finds priming effects for VPCs regardless of whether the referent of the direct object NP in the second construction is co-referential with that of the first. Third, the kind of animacy/argument effects that might in principle affect datives (such that animacy affects constructional choices) cannot explain the results on VPCs where animacy plays no role (cf. Gries, 2003a, 88–89]) and the particle is often aspectual or idiomatically used and can, thus, not be attributed argument status.

An additional important point is that other non-syntactic factors can also not be held responsible for the present findings. For example, those who would like to attribute the present results to lexical repetition effects would have to explain why, in the case of dative alternation, it is the ditransitive construction that primes more strongly although (i) it is the prepositional dative which allows for the priming of the function words to and for and (ii) the fact that lexical activation decays too fast makes it unlikely that the long duration of priming effects observed here and in other (experimental studies) is just a lexical memory effect.

Much of the present findings resembles those obtained experimentally so strongly that they cannot be explained away as easily as suggested. While scholars do not rule out discourse-motivated factors of priming at all, it is hard to explain all the similarities between the different kinds of results and still simply uphold the claim that all this is epiphenomenal. Without doubt, further experimental evidence is



necessary, but it seems as if the utility of corpus-based, explorative results should not be underestimated prematurely.

There is also experimental evidence supporting the verb-specificity effect argued for here. Gries and Wulff replicated Pickering and Branigan's (1998) experiments on syntactic priming in English with native speakers of German to determine whether syntactic priming is also obtained with advanced learners of a foreign language. In addition to a general priming effect, they found that, just like in the present study, the strength of the priming effect of the seven dative alternation verbs discussed above is strongly correlated with a general bias of the subjects to use the experimental verbs in particular constructions. Their results are summarized: the y-axis portrays the bias of individual verbs to either the ditransitive or the prepositional dative in the corpus data, basically as measured by collocation strength. The x-axis portrays the preference of individual verbs to be completed using either the ditransitive or the prepositional dative in the sentence-completion task.

Finally, the strong correlation ($r^2 = .8$; $t(5) = -4.47$; $p = .007$) is indicated by the slope of the regression line, which shows that one can predict the outcome in the priming experiment on the basis of the verbs' preferences as measured on the basis of the corpus data.

Even though these results were not obtained with native speakers, they do point to the fact that experimentally primed sentence completion is strongly sensitive to verb bias. Thus, I submit, this issue is clearly in need of further research of which corpus linguistic methods may play an essential role in determining collocation strengths for verbs to be tested (notably, Gries and Stefanowitsch, 2004, on verbs distinctive for actives and passives).

Many previous results are explained within, say, the psycholinguistic model proposed by Pickering and Branigan (1998). In this model, syntactic priming is accounted for in terms of combinatorial nodes which are activated when a verb is used in a particular construction. When a speaker produces a verb in a particular construction, the lemma nodes of all words produced as well as their feature nodes (representing morphological features such as number, tense etc.) and the corresponding combinatorial node are activated. Since the activation level of these nodes and the links relating them decays only gradually, the nodes and links that were just used are more likely to be used again when the next opportunity arises; syntactic priming is the result. Since the combinatorial nodes are directly related to the lemma nodes, priming should be stronger when the same verb is used in both prime and target, but the verb form as such should not influence the priming effect.



Since the findings concerning the dative alternation are so similar, they can of course be equally well integrated into Pickering and Branigan's model. Remember, for example, that not only priming was found, the priming effect was also stronger when the verb lemmas were identical. There is also a tendency for identical verb forms to result in a stronger priming effect, and while that was not hypothesized in the above model, similar tendencies were obtained in Pickering and Branigan's (1998) on singular-vs.-plural form differences. Also, since the model has been argued to involve a shared representation in comprehension and production, the fact that SpeakerID had no strong effect in this study can be explained naturally.

Finally, the fact that the verb-particle constructions exhibit priming supports the idea that order information is encoded within combinatorial nodes. Given this kind of psycholinguistic model, the second kind of finding of this study, the verb-specificity of priming, can be integrated straightforwardly. Recall that each verb lemma is connected to the combinatorial nodes of the construction in which the verb can be used. Since syntactic priming of a construction C involves the repeated activation of C's combinatorial node (so that its resting level is exceeded), it follows naturally that when the link between a verb and C is stronger, priming of (only) that construction should be stronger. This is exactly the point: the verbs which are strongly associated with one construction exhibit priming with this construction much more strongly than with the other construction.

Thus, Pickering and Branigan's model need to be supplemented with the notion that the links between verb lemmas and combinatorial nodes they postulated anyway can also be differentially strong to reflect their degree of attraction/repulsion to a construction as measured by collocation strength. This would allow for the model to accommodate the present findings on verb-specificity, but also allows for an economical representation of many of the findings concerning verb sub-categorization preferences, verb bias etc. Given the strong interest in the issue of whether syntactic priming is best explained as activation patterns or implicit learning (Chang *et al.*, 2000, 2003), it is even conceivable that the network architectures used to test these different conceptions could be somehow enriched with the collocational information.

All in all, the present findings demonstrate how useful — in spite of some limitations — corpus-based approaches to priming phenomena can be to support and extend findings obtained with other methodologies, promoting once more the ideal of converging evidence.



Conclusions

Terminological units are subject to linguistic analysis. Since this type of analysis can be carried out in a number of ways, it is necessary to choose the linguistic approach most in consonance with the object of study. Such an approach should be lexically centered and usage-based, as well as focus on meaning and conceptual representation. As we shall see, such is the case of cognitive linguistic approaches. In the past, Terminology and Linguistics have mostly ignored each other. In its initial phase Terminology was interested in asserting its independence from other knowledge areas, and creating a totally independent discipline. This goal led terminologists to go to great lengths to emphasize differences between the two even to the extent of affirming that terms are not words. In a parallel way, linguistic theory has largely ignored terminology, probably because specialized language has been and is often regarded as merely a special case of general language. Thus, it was not considered as worthy of serious study because anything pertaining to general language was also presumed to be true of specialized language.

However, interesting conclusions about specialized language, scientific translation, and language in general can be obtained when terminology is studied in its own right. As such, it is most certainly susceptible to linguistic analysis within the framework of a linguistic model. Oddly enough, some years ago this seemingly innocuous affirmation would have caused quite a hue and cry in terminological circles. The reason for this was that the first approximations to terminology had normalization as a primary objective. Great pains were taken to strive for totally unambiguous communication through standardization. This signified unicity or one-to-one reference between term and concept. The fact that the majority of terms designate concepts that represent objects in a specialized knowledge field meant that such an objective seemed possible to achieve.

References

- Atkins, S.; M. Rundell & S. Hiroaki. (2003). „The contribution of Framenet to practical lexicography”. *International Journal of Lexicography* 16 (3). pp. 334-357.
- Barsalou, L. W. (2003) „Situated simulation in the human conceptual system”. *Language and Cognitive Processes* 18. pp. 513-562.
- Cabr  Castellvi, M. T. (1999). *Terminology. Theory, Methods, Applications*. Translated by Janet Ann DeCesaris. Amsterdam: John Benjamins.
- Cabr  Castellvi, M.T. (2000a). „Elements for a theory of terminology: Towards an alternative paradigm”. *Terminology* 6(1). pp. 35-57.



- Faber, P. (2002). „Terminographic definition and concept representation”. In: Maia, B. (ed.) *Training the Language Services Provider for the New Millennium*. Oporto (Portugal): University of Oporto. pp. 343-354.
- Faber, P. & C. Jiménez Hurtado (eds.) (2002). *Investigar en terminología*. Granada: Comares.
- Fillmore, C. J. & S. Atkins. (1998). „FrameNet and lexicographic relevance”. In: *Proceedings of the ELRA Conference on Linguistic Resources*, Granada. pp. 417-423.
- Fillmore, C. J.; C. R. Johnson & M. Petruck. (2003) „Background to Framenet”. *International Journal of Lexicography* 16(3). pp. 235–250.
- Gaudin, F. (1993) *Pour une socioterminologie: Des problèmes pratiques aux pratiques institutionnelles*. Rouen: Publications de l'Université de Rouen.
- Gaudin, F. (2003). *Socioterminologie: Une approche sociolinguistique de la terminologie*. Bruxelles: Duculot.
- Geeraerts, D. (1995) „Representational models in cognitive semantics”. *Folia Linguistica: Acta Societatis Linguisticae Europaeae* XXIX (1-2). pp. 21-41.
- Kerremans K.; R. Temmerman & J. Tummers. (2004). „Discussion on the Requirements for a Workbench supporting Termontography”. In: Williams, G. & S. Vessier (eds.) *Proceedings of the XIth Euralex International Congress 2004, 6-10 July, Lorient, France*. Lorient: Université de Bretagne-Sud. pp. 559-570.
- Pihkala, T. (2001). „Socioterminology”. *Terminfo 1/2001 – Summaries, Nordterm 2001*. [Available at http://www.tsk.fi/fi/ti/ti101_teksti.html].
- Strehlow, R. (1993). „Terminological standardization in the physical sciences”. In: Sonneveld, H. B. & K. L. Loening (eds.) *Terminology: Applications in Interdisciplinary Communication*. Amsterdam/Philadelphia: John Benjamins. pp. 127–140.