Journal of Quantitative Linguistics
2001, Vol. 8, No. 1, pp. 33–50
0929-6174/01/0801-0033$16.00


# A Multifactorial Analysis of Syntactic Variation: Particle Movement Revisited*

Stefan T. Gries
Southern Denmark University, Denmark

## ABSTRACT

The present paper investigates the word order alternation of English transitive phrasal verbs such as, e.g., *to pick up the book* versus *to pick the book up*. It builds on traditional monofactorial analyses, but argues that previously used methods of analysis are grossly inadequate to describe, explain and predict the word order choice by native speakers. A hypothesis integrating virtually all relevant variables ever postulated is proposed and investigated from a multifactorial perspective (using GLM, linear discriminant analysis and CART). As a result, more than 84% of native speakers' choices can be predicted. Further implications (linguistic and methodological) are discussed.

## INTRODUCTION

A notoriously difficult problem for syntactic research is the existence of syntactic variation, i.e. closely related syntactic variants with truth-conditionally equivalent meanings. Examples in English include the well-known word order alternations Dative Movement, Preposition Stranding and Particle Movement in (1), (2) and (3) respectively.[1]

(1) (a) John [$_{VP}$ gave [$_{NP}$ the book] [$_{PP}$ to [$_{NP}$ Bill]]].
(b) John [$_{VP}$ gave [$_{NP}$ Bill] [$_{NP}$ the book]].

*Address correspondence to: Stefan T. Gries, Southern Denmark University, Institut for Erhvervssproglig Informatik og Kommunikation, Grundtvigs Allé 150, DK-6400 Sønderborg. E-mail: stgries@sb.sdu.dk

[1]The grammatical notation is not committed to any particular grammatical framework and serves expository reasons only. Likewise, the choice of terminology in terms of movement processes is not meant to truly imply any such processes—it merely reflects that these phenomena have most frequently been dealt with within the transformational-generative paradigm.

(2) (a) Who did you [$_{VP}$ see Bill [$_{PP}$ with t$_i$]]?
(b) [$_{PP}$ With whom]$_i$ did you [see Bill t$_i$]?

(3) (a) John picked up [$_{NP}$ the book].
(b) John picked [$_{NP}$ the book] up.

Several interrelated questions arise with respect to these constructional alternations:

- how do the two constructional variants of each pair differ from each other?
- why and to what extent do different variables influence the subconscious choice of construction by native speakers in a natural setting?
- how can native speaker choices of constructions in a particular discourse situation be predicted? More specifically, which techniques are most suitable for the prediction of native speakers' choices given the complexity of natural discourse settings?

This study investigates these questions for the last of the above-mentioned word order alternations of transitive phrasal verbs; the construction where no element intervenes between verb and particle will be referred to as construction—the construction where a direct object NP intervenes between verb and particle is referred to as construction$_1$:[2]

(4) (a) John picked up the book. construction$_0$
(b) John picked the book up. construction$_1$

Section 2 is concerned with a brief summary of previous analyses of Particle Movement. At the same time, several methodologically motivated points of critique are raised. Section 3 outlines the methods by means of which the above three objectives are pursued. Section 4 deals with the results of the study: monofactorial as well as multifactorial results will be presented in some detail. Finally, Section 5 concludes by situating the study within a broader psycholinguistic framework and by briefly discussing further implications going beyond the immediate scope of Particle Movement.

---

[2]The reason for this seemingly arbitrary and counter-intuitive labelling will be addressed later; as a mnemonic help, consider the index to name the number of contituents intervening between verb and particle.

## PREVIOUS ANALYSES

### Variables that purportedly govern the alternation

The position of particles in transitive phrasal verbs has been investigated time and again within the last 100 years. The approaches come from widely diverging linguistic schools of thought such as Chomskyan transformational-generative grammar (cf., e.g., Fraser, 1974, 1976; Den Dikken, 1992, 1995; Rohrbacher, 1994, to name but a few), traditional grammarians (Sweet, 1892; Jespersen, 1928; Kruisinga & Erades, 1953), cognitive grammar (Yeagle, 1983), discourse-functional oriented approaches (Chen, 1986), psycholinguistically-oriented approaches (Hawkins, 1994) etc. Over the time, various variables have been proposed in order to account for both the optimal structural configuration of the two constructions and the question of which construction is chosen by native speakers. Table 1 provides an overview over

Table 1. Variables that allegedly govern the alternation.

| Value for $construction_0$ | Variable | Value for $construction_1$ |
|---|---|---|
| Long DO | Length of the DO in words (Length W) | |
| Long DO | Length of the DO in syllables (LengthS) | |
| Complex | Complexity of the DO (Complex) | |
| | NP-Type of the DO: semi-pronominal (Type) | pronominal |
| Indefinite | Determiner of the DO (Det) | definite |
| No | Previous mention of the DO (Lm) | yes |
| Low | ◀——Times of preceding mention of the DO (Topm)——▶ | high |
| High | ◀——Distance to last mention of the DO (Dtlm/ActPC)——▶ | low |
| High | ◀——News Value of the DO——▶ | low |
| Yes | (Contrastive) Stress of the DO | |
| Yes | Subsequent mention of the DO (NM) | no |
| High | ◀——Times of subsequent mention of the DO (Tosm)——▶ | low |
| How | ◀——Distance to next mention of the DO (Dtnm/ClusSC)——▶ | high |
| | Overall frequency of the DO (OM) | |
| | following directional adverbial (PP) | yes |
| Yes | Prep of the following PP is identical to the particle (Part = Prep) | |
| | Register | |
| Idiomatic | ◀——Meaning of the VP (Idiomaticity)——▶ | literal |
| Low | ◀——Cognitive Entrenchment of the DO——▶ | high |
| Inanimate | Animacy of the DO (Animacy) | animate |
| Abstract | Concreteness of the DO (Concreteness) | concrete |

the multitude of variables proposed so far. The central column names the variable proposed whereas the left and right column name the values/levels of each variable associated with a particular preference for a construction.

**Comments and points of critique**

This list of variables may seem quite impressive at first. It is especially interesting to note that at a first superficial glance quite simple word order alternation is in fact influenced by variables from many different sub-branches of linguistics: phonology, syntax, semantics, pragmatics and other variables. Unfortunately, however, there are also several shortcomings that have hindered progress considerably.

First of all, most variables are based on introspective analysis (i.e., acceptability judgements) and non-authentic example sentences. While there are some linguistic frameworks which consider this to be a rewarding way of gathering data, I would contend that (i) acceptability judgements very often do not necessarily constitute objective, reliable and valid data (cf., e.g., Labov, 1975; Schütze, 1996); (ii) it is questionable that an analysis solely based on dreamt-up sentences can in fact obtain representative results; and (iii) 'no one has ever presented even a hint of evidence that any part of the human's linguistic competence is the ability to evaluate sentences produced artificially, out of context' (Prince, 1991, p. 80).

Second, most analyses only consider monofactorial results (i.e., the effect one variable has on the alternation in isolation) although for the speaker all variables are given simultaneously. For instance, Fraser (1974) argued that verbs without initial stress prefer construction$_1$, offering the following sentences as what he claims to be supporting evidence.

(5) (a) ?I will insult back the man.
(b) I will insult the man back.

(6) (a) ?We converted over the heating to steam.
(b) We converted the heating over to steam.

(7) (a) ?They attached up the tag on the wall.
(b) They attached the tag up on the wall.

But what is problematic about this approach? Is this not an example of one of the most traditional and well-established methods in linguistics, namely the minimal-pair test? The problem lies in the fact that the examples do not warrant this claim at all: the preference for construction$_1$ in these examples (if there is one at all, recall the scepticism expressed above concerning such isolated acceptability judgements) need not be related to Fraser's claim at all and might as well derive from the fact that short and simple direct objects already favour construction$_1$, as do definite determiners and literal VP meanings (cf. Table 1). If we generalize from this phenomenon to other analyses (which we can do: nearly all previous analyses are monofactorial) we find that, given the complexity of 20 or so interacting variables, we cannot rely on monofactorial analyses to describe Particle Movement adequately.

Finally, it is generally accepted that, normally, science tries to describe, explain and predict phenomena. With Particle Movement (and many other cases of syntactic variation), however, the most rigorous test of one's theory, namely the actual prediction of speakers' behaviour, has never been attempted. Every analysis has aimed at describing particle placement at least to some extent; some analyses have aimed at explaining particle placement, but there are only few analyses aiming at subsuming the variables under a common (set of) factor(s); no analysis has aimed at predicting particle placement in natural discourse situations. This is of course a consequence of the previously mentioned shortcomings: If, for instance, one is not able to quantify the importance of the individually proposed variables, then traditional accounts would already fail to predict constructional choices when only two variables have conflicting preferences. Consider *John picked up a book*. The short direct object prefers construction$_1$ whereas the indefinite determiner prefers construction$_0$. Evidently, without a way to weigh individual variables' preferences, traditional accounts cannot even predict speakers' preferences in the simple cases where only two variables are concerned, which is why so far nobody has managed to predict speakers' choices simultaneously accounting for more than a dozen variables.

As is evident from the three above-mentioned research questions (cf. section 1) I intend to overcome these shortcomings. The following section is concerned with the methods I use to that end.

## METHODS

### The Processing Hypothesis (PH)

In order to explain why speakers choose the construction they do, I propose the following hypothesis: the multitude of variables (most of which are concerned with the direct object NP) that seems to be related to Particle Movement can all be related to the processing effort of the utterance.[3] My idea of the notion of processing effort is a fairly broad one: it encompasses not only purely syntactic determinants, but also factors from other linguistic levels. More specifically, I assume that virtually all levels of linguistic description mentioned above can contribute to processing effort:

- phonologically indicated processing cost: contrastive stress on a linguistic expression increases the amount of processing effort because the speaker focuses (which is itself not an effortless task) the hearer's attention on the referent of the contrastively-stressed expression;
- morphosyntactically determined processing cost: the longer and the more complex the direct object NP is, the more effort (and working memory) is needed to process the utterance correctly;
- semantically conditioned processing cost: if the meaning of the VP is idiomatic, then the meaning of the whole of the transitive phrasal verb is not computable from the meaning of the individual parts so that the parts of the idiomatic phrasal verbs rely more on one another than with totally literal phrasal verbs (which mostly designate caused motion). Following, say in Hawkins (2000), we may assume that there is a tendency to minimize what he refers to as lexical dependency domains (LPDs), i.e. (slightly simplified) the distance between expressions (at times mutually) dependent on one another for their interpretation. With idiomatic phrasal verbs (e.g., *to eke out*), the semantic dependency between verb and particle is higher than the dependency for literal verbs (e.g., *to bring back*), so we would accordingly expect construction$_0$, minimizing the distance between the component parts of the phrasal verb, whereas with literal phrasal verbs, no preference for a construction is to be expected because the low degree of interdependence does not require a particularly small distance between the component parts and, thus, licenses both word orders. In connection with that, concrete DOs are more likely to correlate with a literal interpretation of the verb-particle construction since these referents can undergo the caused motion that verb-particle constructions commonly denote—abstract DOs, on the other hand,

[3]The question arises as to whether I refer to the processing effort of the speaker or the hearer. My own focus is on the speaker's processing effort—with Particle Movement, however, we find that what makes processing efficient for speakers is also beneficial to hearers. Thus, no strict differentiation between the two interlocutors is necessary here.

give rise to less literal interpretations (*to bring back peace* is not a case of literal caused motion). Thus, concreteness/abstractness of the DO's referents correlates with the literalness/idiomaticity of the verb-particle constructions and yields identical predictions.

- discourse-functionally determined processing cost: if the referent of the direct object NP is discourse-given or can be inferred from the preceding context, then its activation and production incurs less processing cost than the activation and production of some discourse-new or even completely unknown referent (cf., e.g., Givón, 1992). A morphosyntactic phenomenon strongly correlating with degree of givenness of NP referents is the choice of determiner. It is widely acknowledged that indefinite determiners are typically used for new referents while definite determiners are more often found with given referents.

Since (i) speakers strive to communicate whatever they intend to communicate with as little effort as possible and (ii) construction$_0$ is inherently easier to process (cf. Hawkins, 1994; Rohdenburg, 1996), they will tend to use construction$_0$ in situations where the processing effort associated with the utterance is already high. In other words, if the VP does not require a lot of processing effort (due to its limited length, the degree of activation of the DO's referent, etc.) then construction$_1$ is chosen—if the VP requires a lot of processing effort (due to the processing cost for the DO's referent) then construction$_0$ is chosen in order to facilitate communication by minimizing the structurally determined processing effort. Note, however, that this hypothesis implies that some of the variables mentioned above will not be relevant for the choice of construction since there is no reason why variables concerned with the discourse *following* the verb-particle construction should play a role just as there is no reason why the animacy of the direct object's referent should be important. Finally, contrary to a previous analysis (cf. Gries, 1999), the variable of entrenchment is not considered to be relevant since the variables that, taken together, constitute the entrenchment hierarchy used previously are investigated here separately and thus much more precisely (cf. Gries, 2000 for a more detailed statistical analysis of these variables).

## The data

In order not to rely on made-up sentences and their (at times) doubtful acceptability judgements, I advocate the use of naturally-occurring data. I have, therefore, compiled a sample of 403 utterances with verb-particle

Table 2. Data from the British National Corpus.

| | $Construction_0$ | $Construction_1$ | Row totals |
|---|---|---|---|
| Spoken | 67 | 133 | 200 |
| Written | 127 | 76 | 203 |
| Column totals | 194 | 209 | 403 |

constructions from the British National Corpus. The verb-particle constructions chosen mainly consist of combinations of the most frequent verbs and particles entering into transitive phrasal verbs.[4] Table 2 shows the distribution of my data.

To each of these sentences the 10 preceding and subsequent clauses (without false starts or discourse markers such as *You know* or *I mean*) were added. Then, each sentence was investigated with respect to the variables listed above in Table 1.[5] The table resulting from these processes was the basis for the analysis to follow.

**Statistical techniques**

First of all, for each variable a monofactorial correlation was computed. Depending on the measurement scale of the independent variables, the coefficients given in Table 3 were determined.[6]

It shall be noted, however, that the monofactorial correlations are only determined in order to test previous monofactorial analyses empirically, most of which have not been empirically tested before. The primary purpose of this paper, i.e., the prediction of speakers' choices, can only be achieved with more elaborate techniques. The multifactorial techniques that will be used are the

[4]I have determined the most frequent verbs and particles in transitive phrasal verbs on the basis of my own collection of 1,357 transitive phrasal verbs from several dictionaries.

[5]The degrees of complexity and idiomaticity were measured on ordinal scales: simple NPs, intermediate NPs (NPs with modification by adjectives and/or genitives) and complex NPs (containing embedded clauses) for complexity and simple/literal, metaphorical/figurative and idiomatic/opaque VPs.

[6]The ranking is roughly according to the size of the correlation (but cf. below). While the two values of $\phi$ and $\lambda$ do not always coincide, the sizes of all correlation coefficients (once with the $\phi$ coefficient, once with $\lambda$) correlate highly significantly ($\gamma = 0.85$; $z = 6.923$; $p < 0.001^{***}$), which is why these minor ranking differences will not be dealt with.

Table 3. Monofactorial correlations computed for the corpus data.

| Measurement scale of the independent variable | Correlation coefficient |
|---|---|
| Nominal/categorical | Phi/Cramer's I and Lambda |
| Ordinal | $\gamma$ (equalling Kendall's $\tau$ with correction for ties) |
| Interval | Pearson product–moment correation |

general linear model (GLM), linear discriminant analysis (LDA) and classification and regression trees (CART).

## RESULTS

### Monofactorial results

Table 4 summarizes the monofactorial correlation coefficients of every independent variable; for nominal and ordinal variables, the correlations between individual levels and the choice of construction are also provided.[7] The PH is strongly supported by

- those variables/values that indicate a low degree of processing cost (e.g., literal VPs with pronominal DOs or short lexical DOs with definite determiners where the DO has been mentioned before frequently ) favour construction$_1$;
- those variables /values that indicate a high degree of processing cost (e.g., idiomatic VPs with discourse-new lexical referents of long DO NPs with indefinite determiners).

On the whole, the morphosyntactic variables are most influential, semantic and discourse-functional variables are less powerful in determining the choice of construction. However, three comments are necessary. First, given the

[7]The ranking is roughly according to the size of the correlation (but cf. below). While the two values of $\phi$ and $\lambda$ do not always coincide, the sizes of all correlation coefficients (once with the $\phi$ coefficient, once with $\lambda$) correlate highly significantly ($\gamma = 0.85$; $z = 6.923$; $p < 0.001^{***}$), which is why these minor ranking differences will not be dealt with.

Table 4. Monofactorial correlations between variables and the choice of construction.

| Variable/Variable: *Value* | Correlation coefficient |
|---|---|
| Complexity of the DO | $\gamma = -0.85$*** |
| Idiomaticity of the VP | $\gamma = -0.6$*** |
| Complex: *simple* NP | $\phi = 0.522$*** ($\lambda = 0.49$) |
| NP Type of the DO | $\phi = 0.492$*** ($\lambda = 0.366$) |
| Length of the direct object in syllables | $r_{\text{pbis}} = -0.481$*** |
| Type: *lexical* NP | $\phi = 0.47$*** ($\lambda = 0.366$) |
| Type: *pronominal* NP | $\phi = 0.468$*** ($\lambda = 0.32$) |
| Complex: *intermediate* NP | $\phi = 0.455$*** ($\lambda = 0.412$) |
| Distance to last mention of the DO | $r_{\text{pbis}} = 0.452$*** |
| Cohesiveness of the DO to the preceding discourse | $r_{\text{pbis}} = 0.429$*** |
| Length of the DO in words | $r_{\text{pbis}} = -0.423$*** |
| Times of preceding mention of the DO | $r_{\text{pbis}} = 0.414$*** |
| Last mention of the DO | $\phi = 0.411$*** ($\lambda = 0.387$) |
| Overall mention of the DO | $r_{\text{pbis}} = 0.357$*** |
| Concreteness of the DO | $\phi = 0.339$*** ($\lambda = 0.314$) |
| Idiomaticity: *idiomatic* VP | $\phi = -0.328$*** ($\lambda = 0.253$) |
| Determiner of the DO | $\phi = 0.319$*** ($\lambda = 0.206$) |
| Idiomaticity: *literal* VP | $\phi = 0.314$*** ($\lambda = 0.268$) |
| Register | $\phi = 0.291$*** ($\lambda = 0.263$) |
| Det. *indefinite determiner* | $\phi = -0.288$*** ($\lambda = 0.206$) |
| Directional adverbial following the DO | $\phi = 0.284$*** ($\lambda = 0.16$) |
| Det. *no determiner* | $\phi = 0.232$*** ($\lambda = 0.191$) |
| Complex: *complex NP* | $\phi = -0.193$*** ($\lambda = 0.077$) |
| Times of subsequent mention of the DO | $r_{\text{pbis}} = 0.191$*** |
| Animacy of the DO | $\phi = 0.166$*** ($\lambda = 0.057$) |
| Cohesiveness of the DO to the subsequent discourse | $r_{\text{pbis}} = 0.142$** |
| Next mention of the DO | $\phi = 0.104$* ($\lambda = 0.072$) |
| Distance to next mention of the DO | $r_{\text{pbis}} = 0.1$* |
| Type: *semi-pronominal NP* | $\phi = 0.092$*** ($\lambda = 0$) |
| Idiomaticity: *metaphorical NP* | $\phi = -0.047$ ns ($\lambda = 0.016$) |
| Type: *proper name* | $\phi = 0.023$ ns ($\lambda = 0$) |
| Det: *definite determiner* | $\phi = -0.\ 018$ ns ($\lambda = 0$) |
| Particle equals the preposition of the following PP | $\phi = 0.003$ ns ($\lambda = 0$) |

mathematically different ways of calculating these coefficients, it is not possible to simply compare the variables' power by simply comparing the absolute values of the correlation coefficients. Second, some variables that were predicted not to correlate significantly with choice of construction do in fact display a significant correlation so further investigation is called for.

Lastly, as was argued above, a monofactorial perspective (i) does not do justice to the complexity of the phenomenon and (ii) defies any cognitively realistic account of the alternation.

**Multifactorial results: GLM and Linear Discriminant Analysis (LDA)**

The multiple correlation between all variables included by the PH and the choice of construction as determined by the GLM is highly significant: $r = 0.786$; $F_{71,331} = 7.512$; $p < 0.001^{***}$. Given the large number of inter-correlations between the predictor variables, however, this correlation coefficient needs to be adjusted for shrinkage using Wherry's formula; $r_{\text{adjusted}} = 0.732$. Still, this value is still quite high and strongly supports the PH. It does so especially when we consider the following two points:

- $R$ obtained on the basis of the PH is even larger than $R$ obtained when we include all variables mentioned in Table 1 (namely, $R_{\text{adjusted}} = 0.718$; $F_{126,276} = 4.4$; $p < 0.001^{***}$), which shows that the variables I have chosen to eliminate only add random noise to the analysis anyway;
- multiple correlations that are obtained in other behavioural sciences are often much smaller so the account of variance accounted for in the present study is comparatively large.

But what about the predictive power of my hypothesis? An LDA shows that the variables included in the PH make it possible to predict which construction a speaker will choose in a particular discourse situation. The discriminant function is highly significant (canonical $r = 0.73$; $\chi^2 = 297.37$; d.f. $= 20$; $p = 0^{***}$). Moreover, the discriminant function can classify 86.1% of the constructional choices within the sample. However, it is more important to also cross-validate this result in order to avoid circularity of reasoning by using cases for their own 'prediction'. Two measures were therefore taken to improve the analysis:

(a) the leave-one-out method for cross-validation, yielding a prediction accuracy of 84.1%; a result that is virtually impossible to obtain by pure chance, according to the exact binomial test, the chance for 339 correct hits in 403 trials approaches zero;

(b) the split-sample technique, where I divided the corpus data into a learning sample and a prediction sample in order to derive a discriminant function from the learning sample which was subsequently applied to the prediction sample. In order not to be accused of a possibly biased choice

Table 5. Cross-validated prediction accuracy of LDA for split samples.

| Learning sample | Prediction sample | Correct predictions for prediction sample |
|---|---|---|
| 200 spoken sentences and 150 written sentences | 53 written sentences | 84.9%; ($p_{\text{binomial test}} \approx 1.184 \times 10^{-7***}$) |
| 150 spoken sentences and 200 written sentences | 53 spoken sentences | 64.2%; ($p_{\text{binomial test}} \approx 0.027^{*}$) |
| 174 spoken sentences and 176 written sentences | 26 spoken sentences and 27 written sentences | 75.5%; ($p_{\text{binomial test}} \approx 1.343 \times 10^{-4***}$) |
| Average | | 74.8% |

of samples, this was done three times with randomly chosen sentences from the different registers. Table 5 gives an overview of the results.

Obviously, the results are quite robust: the prediction accuracies are all significant, as determined by testing the correct hit rate against the one expected by pure chance using the exact binomial test.[8] The best prediction results are achieved for written data, the worst for oral data, which is to be expected, given the more spontaneous and interactive nature of natural discourse as opposed to planned writing.

Let us now try to find out which variables are responsible for the good discrimination between the two constructions. The previous results were concerned with an LDA where, for theoretical reasons, only the variables of the PH were included. But we also need to find out whether it is empirically plausible to exclude some variables from further consideration, be it only to support the results obtained by the GLM. Thus, a second LDA was computed where all variables where included; Table 6 summarizes its results: the higher the absolute value of a factor loading for a variable the more important it is for the choice of construction in the only cognitively realistic analysis of the situation, namely when all of the variables are considered simultaneously.

[8]One might wonder why the split-sample technique yields slightly worse results than the leave-one-out method. This is due to the fact that the learning samples for the leave-one-out method contain 52 sentences more than those of the split-sample technique.

Table 6. Factor loadings of the discriminant analysis.

| Variable | Factor loading | Kind of variable | Choice of construction |
|---|---|---|---|
| Length S | 0.522 | Morphosyntactic | High variable values ⇒ |
| Type: *lexical* | 0.498 | | $\text{construction}_0$ |
| Complex: *intermediate* | 0.479 | | |
| LengthW | 0.447 | | Low variable values ⇒ |
| Idiom: *idiomatic* | 0.325 | Semantic | $\text{construction}_1$ |
| DET: *indefinite* | 0.281 | Morphosyntactic | |
| Complex : *complex* | 0.184 | Morphosyntactic | Due to the low factor |
| Idiom: *metaphorical* | 0.044 | Semantic | loadings ($< 0.22$) |
| DET: *definite* | 0.016 | Morphosyntactic | these variables do |
| Disfluency | 0.006 | Other | not discriminate |
| Part = Prep | 0.002 | Other | well between the |
| Type: *proper name* | −0.021 | Morphosyntactic | two constructions |
| Type: *semipronominal* | −0.086 | | |
| ClusSC | −0.094 | Discourse-functional (subsequent context) | |
| NM | −0.098 | | |
| CohSC | −0.135 | | |
| Animacy | −0.157 | Semantic | |
| TosM | −0.183 | Discourse-functional (s. c.) | |
| DET: *no determiner* | −0.223 | Morphosyntactic | High variable values ⇒ |
| PP | −0.278 | Other | $\text{construction}_1$ |
| Idiom: *literal* | −0.309 | Semantic | |
| Concrete | −0.337 | | Low variable values ⇒ |
| OM | −0.358 | Discourse-functional | $\text{construction}_0$ |
| LM | −0.422 | Discourse-functional | |
| TopM | −0.427 | (preceding context) | |
| CohPC | −0.445 | | |
| ActPC | −0.474 | | |
| Type : *pronominal* | −0.496 | Morphosyntactic | |
| Complex : *simple* | −0.573 | | |

Obviously, the PH is again strongly supported. Not only do we find that all variables included in the PH correlate with the choice of construction as predicted—in the multifactorial analysis, we see that the variables concerned with the subsequent context are indeed irrelevant, as was predicted by the PH. On the whole, we find the following ranking of strength of variable groups:

discourse-functional variables (preceding context), syntactic variables, semantic variables and other variables.[9]

**Multifactorial results: Classification and Regression Trees (CART)**

While the results of the LDA are quite convincing, there is one objection concerning the application of an LDA that might be raised. It is concerned with the standard assumption that an LDA requires a multivariate normal distribution of the data, and one might (correctly) claim that it is doubtful that my data do indeed meet this demand and, thus, that the above results are to be taken with a grain of salt. However, there are several arguments supporting the above analysis, results and interpretation even though the distributional assumptions of LDAs are not met.

First, while many researchers tend to emphasize the importance of distributional assumptions (such as normality, homogeneity of variances and the like), a number of scholars argue that, in practice, these assumptions are not as essential as they might seem on a purely mathematical basis (cf. Winer et al., 1991, p. 5). Second, it has even been claimed that there is no test that reliably identifies multivariate normal distributions (cf. Bortz, 1999, p. 435). Third, the difference between LDA and CART is of course not just a statistical/mathematical one—rather, there is also a conceptual difference: while an LDA includes all variables simultaneously in the calculation to compute a prediction for one of the two constructional choices, the tree resulting from CART analyses includes variables sequentially. For a native speaker however, I believe that the model underlying LDA is more realistic. It is intuitively more plausible to assume that all the variables' values/levels I have discussed are somehow set at the point of time the speaker chooses the word order rather than that the values/levels are included one by one in a sequential fashion. Moreover, while there is still considerable debate whether psycholinguistic theories of speech production should incorporate parallel or serial models of processing, I, following Berg (1998), consider parallel processing theories more rewarding from both a theoretical and a practical point of view. I have decided, for these reasons, to predict native speakers' choices with an LDA which, as opposed to CART, comes closest to predicting choices on the basis of a simultaneous/parallel inclusion of the relevant data.

---

[9]This was determined by calculating (i) the AMs of the absolute values of the factor loadings for each variable group and (ii) the medians of the ranks of all variables in a group when the vaiables are ordered according to their factor loadings. Both results were identical.

Table 7. Parameters and settings of the CART analysis.

| Parameter | Setting |
|---|---|
| Method | CART-style exhaustive search for univariate splits |
| Stop rule | FACT-style direct stopping:fraction of objects = 0.5 |
| Prior probabilities | identical: construction$_0$: $p \approx 0.5$<br>construction$_1$: $p \approx 0.5$ |
| Goodness-of-fit index | *Gini* measure |

Nevertheless, it might very well be the case that these reasons do not satisfy truly mathematically-oriented researchers. I have, therefore, also analyzed my data using the CART module of Statistica 5.5; the algorithms used therein are based on CART by Breiman et al. (1984), where CART and QUEST algorithms are used to classify and predict data in the absence of distributional assumptions. My CART analysis of the data was based on the parameters and settings given in Table 7.

The result of the analysis can be summarized as follows: out of all 403 sentences, 349 (86.6%) were classified correctly while 54 (13.4%) were classified wrongly, again a result that is extremely unlikely to be obtained randomly. Again, however, we must also determine the prediction accuracy by a cross-validation technique. Since the leave-one out method for CART is not available in Statistica 5.5, I used only the split-sample technique analogous to the LDA; the samples and the results are listed in Table 8.

Admittedly, the cross-validated prediction accuracy of CART is not as high as the LDA results, but, apart from the prediction sample for oral data alone,

Table 8. Cross-validated prediction accuracy of CART for split samples.

| Learning sample | Prediction sample | Correct predictions for prediction sample |
|---|---|---|
| 200 spoken sentences and 150 written sentences | 53 written sentences | 81.1%; ($p_{\text{binomial test}} \approx 2.775 \times 10^{-6***}$) |
| 150 spoken sentences and 200 written sentences | 53 spoken sentences | 56.6%; ($p_{\text{binomial test}} \approx 0.205$ ns) |
| 174 spoken sentences and 176 written sentences | 26 spoken sentences and 27 written sentences | 77.4%; ($p_{\text{binomial test}} \approx 4.086 \times 10^{-5***}$) |
| Average | | 71.7% |

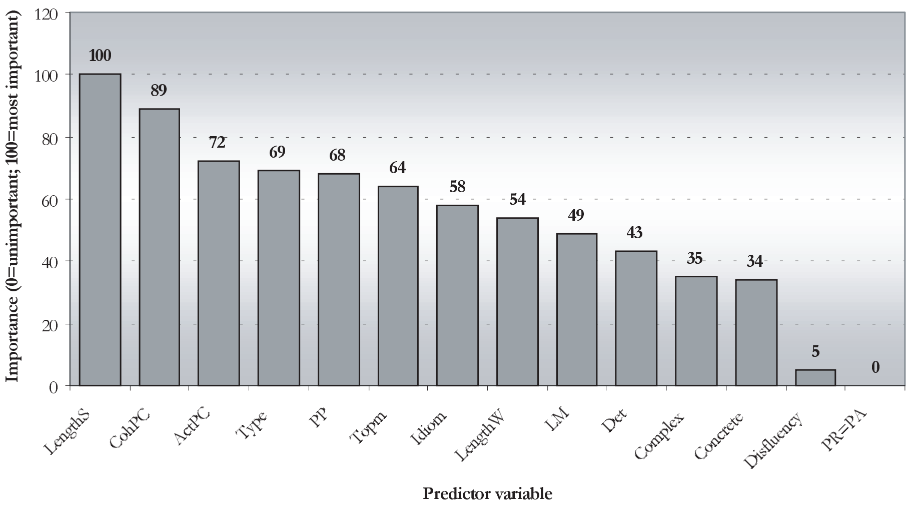


Fig. 1. Importance of predictor variables for CART.

they are still way better than what might be expected by pure chance. Moreover, there is a reason for these minor differences. Given the above parameter settings, the CART technique does not utilize all variables for the prediction of a choice of construction but only the most important ones as determined by the analysis. Thus, for cases where variables of an overall minor importance are decisive, false predictions are more likely.

As far as the importance of the individual variables is concerned, the overall picture does not differ strongly from the results of the LDA. The overall ranking of the variable groups from CART is identical to that of the LDA; for the sake of completeness, Figure 1 shows the results for the individual variables.

## SUMMARY

For each relevant variable ever investigated, it was shown how it contributes to particle placement in isolation. Moreover, it was shown how all of these variables together yield a preference for a construction in particular discourse situations. It is now possible to predict with quite a high prediction accuracy what a speaker will say if the discourse situation he/she is engaging in is known to the analyst.

A hypothesis was proposed and supported that includes all relevant variables and that correctly predicted some variables not to be relevant. It could be shown that a cognitively realistic approach to language usage made it possible to summarize and extend the previous knowledge on particle placement.

On a methodological level, we have seen that the analysis of syntactic variation can benefit considerably from the use of multifactorial techniques just as the analysis of register variation has profited from Biber's (1988) groundbreaking work. Personally, I would go as far as to say that only by such techniques can we start to really detect hitherto unknown patterns that are not already known from early traditional grammarians' works (as was, unfortunately, the case with many works on Particle Movement). We have also seen that different multifactorial techniques, although quite different from one another with respect to their distributional assumptions, yield comparable results. Both LDA and CART achieve convincing classification and prediction accuracies, which also strongly support the PH. Moreover, the individual variables' importance ratings of the two procedures are strikingly similar, so, at least for the case at hand, the different mathematical requirements of both kinds of analyses do not seem to play a vital role.

Lastly, it was at least briefly hinted at the wealth of information that can be obtained on the linguistic data and the way speakers presumably organize their knowledge in order to arrive at constructional choices. This is not to say that speakers actually perform LDA or CART analyses, but it is meant to imply that we can learn something about the importance of (groups of) variables in the process of online production and any model of language production should be able to accommodate these facts in cognitively/psychologically real ways. One possible model that can accommodate the findings reported above naturally is the Competition model by Bates and MacWhinney (1982, 1989), where different variables' cue strengths compete in order to get their preferences recognized. Moreover, the present findings can be readily integrated into activation models where variable weightings (be it in terms of factor loadings or importance values) correspond to association strengths or similar concepts. In this respect, the investigation of further cases of syntactic variation can probably shed light on the nature of activation networks.

## REFERENCES

Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L.R. Gleitman (Eds.), *Language Acquisition: The State of the Art*. Cambridge: Cambridge University Press.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In E. Bates & B. MacWhinney (Eds.), *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.

Berg, T. (1998). *Linguistic structure and change: An explanation from language processing*. Oxford: Clarendon Press.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Bortz, J. (1999). Statistik für Sozialwissenschaftler. 5. Ausgabe. Berlin, Heidelberg, New York: Springer.

Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1984). *Classification and regression trees.* Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software.

Chen, P. (1986). Discourse and particle movement in English. *Studies in Language*, *10*, 79–95.

Den Kikken, M. (1992). *Particles*. Holland Institute of Linguistics Dissertations. The Hague: Holland Acad. Graphics.

Den Dikken, M. (1995). *Particles: On the syntax of verb-particle, triadic, and causative constructions*. Oxford: Oxford University Press.

Fraser, B. (1974). The phrasal verb in English, by Dwight Bolinger. *Language, 50*, 568–575.

Fraser, B. (1976). *The verb-particle combination in English*. New York: Academic Press.

Givón, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics, 30*, 5–55.

Gries, S.T. (1999). Particle placement: A cognitive and functional account. *Cognitive Linguistics*, *10*, 105–146.

Gries, S.T. (2000). *Towards multifactorial analyses of syntactic variation: The case of particle placement.* Doctoral Dissertation, University of Hamburg, Faculty of Language Sciences.

Hawkins, J.A. (1994). *A performance theory of order and constituency.* Cambridge: Cambridge University Press.

Hawkins, J.A. (2000). The relative ordering of prepositional phrases in english: Going beyond manner–place–time. *Language Variation and Change, 11*, 231–266.

Jespersen, O. (1928). *A modern English grammar on historical principles*. London: George Allen and Unwin Ltd.

Kruisinga, E., & Erades, P.A. (1953). *An English Grammar.* Vol. I. Groningen: P. Noordhoff.

Labov, W. (1975). Empirical foundations of linguistic theory. In R. Austerlitz (Ed.), *The Scope of American Linguistics* (pp. 77–133). Lisse: The Peter de Ridder Press.

Prince, E.F. (1991). On functional explanation in lingistics and the origins of language. *Language and Communication, 11*, 79–82.

Rohdenburg, G. (1996). Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics, 7*, 149–182.

Rohrbacher, B. (1994). English main verbs move never. *The Penn Review of Linguistics, 18*, 145–159.

Schütze, C.T. (1996). *The empirical base of linguistics: Grammaticality judgements and linguistic methodology.* Chicago: University of Chicago Press.

Sweet, H.A. (1892). *A new English grammar.* Oxford: Clarendon Press.

Winer, B.J., Brown, D.R., & Michels, K.M. (1991). *Statistical principles in experimental design*. 3rd ed. New York: McGraw-Hill.

Yeagle, R. (1983). *The syntax and semantics of English verb-particle constructions with off: A space grammar analysis*. Unpublished M.A. Thesis, Southern Illinois University at Carbondale.