Language users as creatures of habit: A corpus-based analysis of persistence in spoken English[*]

BENEDIKT SZMRECSANYI

Abstract

For different reasons, speakers re-use recently used or heard linguistic options whenever they can, a tendency which is referred to as 'persistence' in the present paper. The phenomenon has been largely neglected in extant corpus-based, variationist research, and no standard methodology for dealing with the phenomenon is available. By analyzing three well-known alternations (analytic vs. synthetic comparatives, particle placement, and future marker choice) in several spoken corpora of English, this paper demonstrates that factoring in persistence increases the researcher's ability to account for linguistic variation. It is also shown that persistence itself is subject to several determinants, such as textual distance between two successive choice contexts in discourse, or turn-taking. In conclusion, I argue that persistence is a factor which deserves empirical attention, and that its existence has consequences for both linguistic theory and practice.

Keywords: alternations, variation, persistence, priming, repetition

1. Introduction

Dwight Bolinger once noted that

> at present we have no way of telling the extent to which a sentence like *I went home* is a result of innovation, and the extent to which it is a result of repetition […]. Is grammar something where speakers 'produce' (i. e., originate) constructions, or where they 'reach for' them, from a pre-established inventory? (Bolinger 1961: 381).

Indeed, as a corpus linguist dealing with naturalistic data, one regularly has the nagging suspicion that language users are creatures of habit who seem to 'reach for' at least as much as they 'produce'. Consider, for

Corpus Linguistics and Linguistic Theory 1–1 (2005), 113–149 1612-5681/05/0001–0113

instance, the variation between the future markers be going to and WILL in English in the conversational snippet in (1):

(1) JOE: Uh is there LCL accounts gonna be maintained here,
JIM: cause cause
JOE: I mean, or … is there gonna be a separate, they're gonna have an account in Chicago, for the funds to pass through? Or is it gonna be passthrough funds here at the bank? Or, is that …
JIM: x… Well, … w- … what we'll do is, those'll probably wire transfer out.
JOE: Through Boltmans or something,
JIM: Well, … through the Fed, what … I think what will happen, … but we … Matt'll find this out, and, I mean, we'll get involved in it.
(Corpus of Spoken American English, text *"Bank Products"*)[1]

To explain the variation in future marker reference observable in (1), one could point out semantic constraints (e. g., *gonna* is used in *is there LCL accounts gonna be maintained here* because *LCL accounts* are assumed to be on the path to being maintained). One could argue, too, that there are phonological factors (cliticized *will* is used in *Well, … w- … what we'll do* because there are many words in this clause that start in /w/), or that there is an idiolect issue (Jim only uses WILL markers, Joe only uses BE GOING TO markers). Alternatively, one could also note that in (1), all the BE GOING TO markers and all the WILL markers are heavily clustered and argue that successive variable sites in discourse possibly influence each other. This paper will deal with the latter type of explanation, and will refer to this phenomenon as *persistence*.

There is much evidence that persistence plays an important role in language use. For one thing, there is a sizable body of psycholinguistic, experimental research demonstrating that language users are hard-wired to go for recently used (or activated) linguistic patterns whenever they can. This phenomenon is known as *production priming*. For an instance of production priming, take what is known as *syntactic priming*: Bock (1986) investigated syntactic priming in the choice of active/passive constructions and in prepositional/double object constructions. In the experiments she set up, subjects had to read out a priming sentence containing one of the relevant constructions. Subsequently, they were presented with an unrelated event in a picture which they had to describe. Bock found that the structural properties of the priming sentence significantly influenced subjects' subsequent description of the pictures – those subjects who were asked to read out loud a passive sentence such as *John was seen by Mary* were substantially more likely to describe an event using a passive sentence such as *the church was struck by lightning* (rather than, e. g., *lightning struck the church*) than were subjects who were presented with an alternative priming sentence (such as *Mary saw John*). Crucially, priming is not restricted to syntax: there is evidence of semantic priming (Meyer and Schvaneveldt 1971), lexical priming (Levelt and Kelter 1982), morphological priming (Kempley and Morton 1982), form priming (Tanenhaus et al. 1980), and word order priming (Hartsuiker and Westenberg 2000). Production priming phenomena are often assumed to be due to spreading activation levels in a network of memory which is presumably organized in terms of lexical, morphological, phonological/formal, or syntactic similarity. When a word, morpheme, phonological form, or syntactic structure is recognized, some site in the network is activated, and this activation may subsequently spread to nodes of related patterns or tokens (for instance, Tanenhaus et al. 1980).

Along different lines, discourse analysts and – in particular – conversation analysts have revealed the important role repetitiveness plays in managing discourse. For example, Tannen has devoted two papers published in *Language* (Tannen 1982, 1987) and part of a book (Tannen 1989) to the question of why repetitiveness is so pervasive in conversation, and what effects conversationalists can achieve by being repetitive. Tannen claims that repetition in conversational interaction maintains involvement, connection, and interaction, and that it can be functionally exploited: repetitiveness is speaker-economical in that it provides for planning time, and hearer-economical in that it can help relax the processing load that comes with otherwise informationally dense discourse.

In summary, repetitiveness and production priming have been shown to be important aspects of language use. In this spirit, Sankoff and Laberge (1978) suggested that while it is corpus-linguistic standard practice to view successive occurrences of a variable as independent binomial trials (like independent, unrelated throws of a dice), there may, in fact, exist interactions between neighboring variables, depending on the syntagmatic proximity between them. Yet, corpus linguists working quantitatively have not really followed up on Sankoff and Laberge (1978) and, more often than not, have chosen to ignore the psycholinguistic and discourse-analytic evidence of the pervasiveness of persistence. To the best of my knowledge, only a handful of published corpus studies have dealt quantitatively with persistence and related phenomena, and in most of these the authors quite accidentally stumbled across the phenomenon as one factor among many: Poplack (1980) studied factors favoring retention or deletion of the plural marker in Puerto Rican Spanish and found that when plural markers were deleted on tokens preceding a variable, the plural is likely to be deleted on the variable as well. Weiner and Labov (1983) were somewhat surprised to find that structural parallelism is one of the most potent predictors influencing the choice between actives and passives (their findings were later elaborated on by Estival 1985). Scherre and Naro (1991) investigated plural marking in Brazilian Portuguese, obtaining much the same results with regard to morphological parallelism as Poplack (1980). Poplack and Tagliamonte (1993, 1996) studied past tense marking in Nigerian Pidgin English and in "early" Black English. In both varieties, past temporal reference is optional but persistent in that previous marking increases the probability for marking on the variable. Finally, Gries (forthcoming) investigated syntactic priming in the English dative alternation and in English particle placement through a corpus-based method.

While especially Gries (forthcoming) has begun to remedy the dearth of systematic, quantitative corpus-linguistic investigations into the phenomenon, the author seeks to contribute to psycholinguistic model building. In an attempt to contribute to variationist model building, by contrast, three overarching objectives will guide the present study:

1. To show that corpus data, to the extent that this is possible at all (see fn. 2), can match psycholinguistic data;
2. To suggest a methodology to integrate persistence into variationist research designs;
3. To demonstrate that consideration of the phenomenon can increase the linguist's ability to account for linguistic variation, and to predict speakers' linguistic choices more accurately.

To this end, I will conduct three case studies where I analyze the determinants of well-known alternations in English, including persistence-related predictors, on the basis of data drawn from several corpora of spoken English.

2. Design and Methods

I will refer to what discourse analysts call 'repetitiveness' and what psycholinguists term 'production priming' as *persistence* in language usage.[2] I operationally define

> *persistence* as referring to the tendency that iff speaker A faces a variable Z where he or she has the choice between two or more semantically equivalent variants (regardless of whether they are lexical, morphological, or syntactic in nature), speaker A's choice will be affected by

(α) previous exposure to the variable Z, such that use of a specific variant (either by speaker A or by another speaker B, to whose output speaker A has been exposed) in previous discourse will make it more likely, all other things being equal, that the same variant will be used again by speaker A (henceforth: α-persistence; see example (2)); or by

(β) previous exposure to a linguistic pattern Z*, which is not necessarily variable but parallel to one of variable Z's variants, such that use of the linguistic pattern Z* (either by speaker A himself or by another speaker B, to whose output speaker A has been exposed) in previous discourse will make it more likely, all other things being equal, that the variant of variable Z which is parallel to the linguistic pattern Z* will be used by speaker A (henceforth: β-persistence; see example (3)).

Two points about the above definition ought to be stressed. First, it is assumed that persistence can pertain across speakers and across turns. For one thing, the psycholinguistic evidence on priming effects certainly supports this assumption (cf., for instance, Levelt and Kelter 1982). Also, repetition across turns and repetition of what another speaker says ('allo-repetition' in discourse-analytic terminology, cf. Tannen 1989; 'cross-speaker priming' or 'comprehension-to-production priming' in psycholinguistic terminology, cf. Branigan et al. 2000) are widely observed in discourse and serve important functions. Second, my definition not only includes dependencies between two occurrences of the same variable or two choice contexts (α-persistence), as in (2), but also dependencies between a variable and a linguistic pattern which is not necessarily a variable itself, but which shares one or more syntactic, morphological, or lexical properties with one of the alternating variable's variants (β-persistence), as in (3):

(2) Matt 'll find this out, and, I mean, we'll get involved in it. (Corpus of Spoken AmerikanEnglish, text "Bank Products")

(3) You go look, and every horse's hoof is shaped different. It doesn't matter. Every horse is gonna have a little different shape. (Corpus of Spoken American English, text "Actual blacksmithing")

The idea of β-persistence is that the occurrence of the spatial verb *go* (*you go look*), although not a future marker itself, might help trigger the nearby BE GOING TO future marker (*is gonna have* instead of *will have*) because it shares lexical and phonological substance with the BE GOING TO marker. In the present study's perspective, therefore, β-persistence is a fairly broad and powerful notion, potentially capturing all sorts of ties (lexical, syntactic, and even possibly phonological) between a given linguistic option and its linguistic context.[3] I should also stress that my definition of persistence not only captures *syntactic* or *structural* persistence, but also, e. g., lexical and morphological persistence (therefore, the psycholinguistic equivalent to persistence is not necessarily *syntactic* priming). This makes the present study's approach compatible, for instance, to studies of surface parallelism in morphological marking such as Poplack (1980), Scherre and Naro (1991), and Poplack and Tagliamonte (1993, 1996).

The loci where persistence effects can be investigated in a corpus-based approach are those identifiable occasions in the data where speakers demonstrably have the choice of using one variant or another. Crucially, the notion of 'choice' implies that there is rough semantic equivalence between the two options. This condition is met by the following alternations, which will serve as case studies in this paper:

- analytic vs. synthetic comparatives (*John is cleverer than Mary* vs. *John is more clever than Mary)*
- particle placement (*John looked up the word* vs. *John looked the word up*)
- BE GOING TO vs. WILL as future markers (*John will see Mary* vs. *John is going to see Mary*)

The principal tool used for investigating persistence in the present study is *binary logistic regression*, which is a VARBRUL-like multivariate analysis method. Logistic regression models estimate which of two outcomes – in this study, which of two alternative linguistic options – is more likely to occur given that one or more independent variables, which may be scalar, categorical, or both, influence that outcome. Hence, this paper's analyses point out, among other things, how usage of linguistic option A in a given slot will influence the odds that linguistic option B will be used next time there is a choice. The following information is provided by a logistic regression model (and will be reported in tables 1, 2, and 3 below):

The magnitude and the direction of the influence of each independent on the outcome. This information is provided by *odds ratios* (or exp(b) values) that are associated with each individual independent. Odds ratios indicate how the presence or absence of a feature (for categorical independents) or how a one-unit increase in a scalar independent influences the odds for an outcome. Because odds ratios can take values between 0 and $\propto$, three cases can be distinguished: (i) if exp(b) $<$ 1, an increase in the independent makes a specific outcome less likely; (ii) if exp(*b*) = 1, the independent has no effect whatsoever on the outcome; (iii) if exp(*b*) > 1, an increase in the independent makes a specific outcome more likely; results from tests of statistical significance of each exp(*b*) value will be reported. Note that significance levels are independent from effect sizes,[4] and that a general problem with multivariate analyses is constituted by inter-correlations between the factors, a phenomenon which is known as collinearity. Appendix A reports collinearity measures of the factors analyzed in the present study; as can be seen, collinearity is not a major issue in the datasets used in the present study.

Predictive efficiency of the model as a whole. The percentage of correctly predicted cases vis-à-vis the baseline prediction (*% correct (baseline)*) indicates how accurate the model is in predicting actual outcomes. The higher this percentage, the better the model fares in this endeavor.

Variance explained by, or explanatory power of, the model as a whole (R^2). The R^2 value can range between 0 and 1 and indicates the proportion of variance in the dependent variable (i. e., in the outcomes) accounted for by all the independent variables included in the model. Bigger R^2 values mean that more variance is accounted for by the model and that the model is substantially more significant. The specific R^2 measure which is going to be reported is the so-called *Nagelkerke R^2*, a pseudo R^2 statistic for logistic regression.

The statistical analyses which will be presented below generally include one binary dependent variable (i. e., which of two alternative options is actually used by a speaker, henceforth: CURRENT), and a number of independent, explanatory predictor variables. The two main independents which pertain to the domain of persistence are the following:

Which variant was employed in the variable preceding CURRENT (henceforth: PREVIOUS)? For the dependent variable under analysis, how was the last occurrence of the variable in the discourse (if there was one) realized, i. e., was the same option used or the alternative option? This is an independent which pertains to α-persistence.

Hypothesis: Use of a given option in PREVIOUS increases the likelihood that the same option will be used in CURRENT.

Textual distance between CURRENT *and* PREVIOUS *(henceforth:* TEXTDIST*).* There is evidence that production priming is stronger when subjects have been primed more recently (for instance, Branigan et al. 1999; Bock and Griffin 2000). TEXTDIST was measured in the *ln* of the number of interjacent words between PREVIOUS and CURRENT and is a proxy for recency of use of an alternating variable. This variable was modelled logarithmically and not, for instance, in a linear fashion because many psycholinguistic priming phenomena have been shown to decay in this way; 'forgetting' functions are rarely linear (see, e. g., Cohen and Dehaene 1998 with regard to inappropriate repetitions due to brain damage; McKone 1995 with regard to decreasing exponential decay of repetition priming, Gries (forthcoming) with regard to priming of prepositional and double-object datives).

Hypothesis: The smaller TEXTDIST, the more powerful persistence effects will be.

Depending on the alternation under analysis, a number of other persistence-related independents will be additionally introduced in the sections below. Note that in order to be able to state anything of interest about the magnitude of persistence effects, it is necessary to relate their scope to factors that have hitherto been claimed to influence the alternations under study in this paper; otherwise, one would not know exactly to what extent consideration of persistence improves the analyst's ability to explain linguistic variation. Inclusion of such factors is also advisable since this minimizes the likelihood that what appears to be a relevant factor is, in fact, a statistically spurious artefact of some other, not included predictor. For these reasons, this study's statistical modelling was not based on persistence factors only, but also on variables meant to tap the major 'traditional' factors known to play a role in the respective alternations. Crucially, however, I will not for a moment claim to have explained any one of the alternations exhaustively. Rather, my point will be to sketch, in somewhat programmatic terms, how persistence-related factors can fruitfully complement 'traditional' factors, and that a portion of what has been thought to be 'free' variation is actually not so free after all.

3. Data

The following data sources will be used:

The British National Corpus (BNC) contains a spoken section of about 10 million words which is subdivided into a *demographically sampled component* (henceforth: DS, spanning ca. 4.5 million words), consisting of "informal encounters recorded by a socially stratified sample of respondents, selected by age-group, sex, social class and geographic region" (Aston and Burnard 1998: 31), and into a *context-governed component* (henceforth: CG, ca. 5.5 million words in size) of formal encounters such as lectures, speeches, talks, etc. For the remainder of this study, the

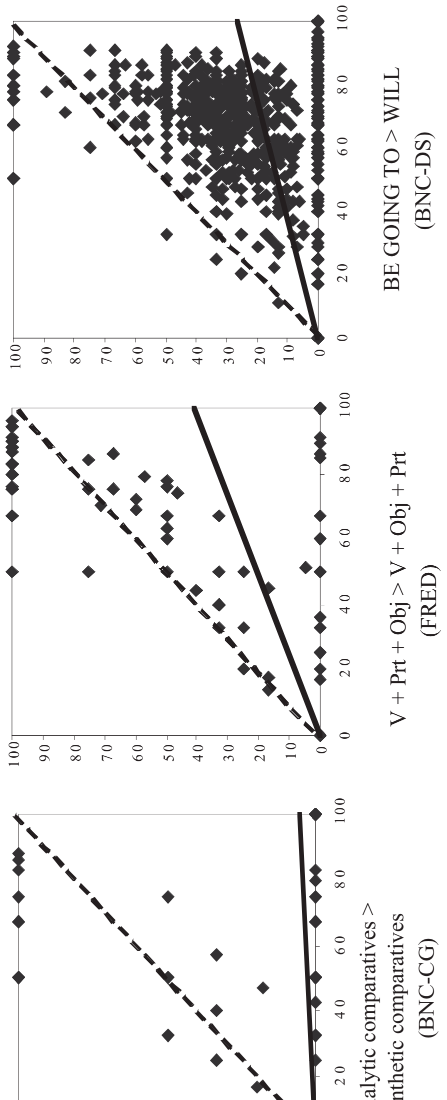
DS and CG sections of the BNC will be treated as separate corpora, the first of which contains informal British English and the second formal British English. The BNC-CG will be analyzed with regard to the alternation between synthetic and analytic comparison while the BNC-DS will be analyzed with regard to future time reference.

The Freiburg English Dialect Corpus (FRED) will serve to investigate variation in particle placement. Its aim is to strengthen research on morpho-syntactic variation in the British Isles (cf. Kortmann 2002). The corpus spans 2.5 million words of running text and consists of samples (mainly transcribed so-called "oral history" material) of dialectal speech from a variety of sources. Most of these samples were recorded between 1970 and 1990; in most cases, a fieldworker interviewed an informant about life, work, etc., in former days. The informants are typically elderly people with a working-class background. Speech styles are relatively formal due to the interview situation, though they are probably less formal than the settings in the formal BNC-CG. Because particle placement is a quite complex alternation that necessitates manual coding (unless the data set is syntax tagged, which FRED is not), my analysis of particle placement in FRED is based on a manageable subset of the corpus.[5]

Therefore, the present study's database spans several data sources, three registers (formal spoken English, informal spoken English, and interview situations), and different varieties of English. This selection of data is an attempt at ensuring that persistence is not restricted to a specific spoken register, data source, or variety. Instead, the present study is going to demonstrate that no matter what corpus (or what spoken register or variety) is being looked at, persistence effects are empirically observable.

4. Results

A rough measure of the extent of persistence in the data is provided by scatterplots such as the ones in figure 1 (cf. Sankoff and Laberge 1978). These scatterplots display switch rates (in per cent, on the vertical axis) from A tokens to B tokens in relation to the share of B tokens of the sum of all A and B tokens (in per cent, on the horizontal axis). Every dot represents the switching behavior of one speaker. A switch is defined as occurring if, given two successive variable sites in discourse, a speaker switches from one variant to the other. Because what is at issue is the sequential configuration of the variable sites, scatterplots such as these display α-persistence: the lower the switch rates, the more powerful α-persistence is. Given two successive variable sites, the leftmost graph



BE GOING TO > WILL (BNC-DS)

V + Prt + Obj > V + Obj + Prt (FRED)

analytic comparatives > synthetic comparatives (BNC-CG)

Figure 1. *Switches between variants as a function of variant proportion (relative frequency of variant A to variant B switches, in %, on Y axis; relative frequency of variant B, in %, on X axis). Each dot represents one speaker. Dotted diagonal line represents null hypothesis that switch rate is proportional to variant proportions. Heavy line indicates linear trend*

in figure 1 displays switch rates from analytic comparison (e. g., *more proud*) in the first site to synthetic comparison (for instance, *cleverer*) in the second site (or at the next opportunity). The second graph displays switch rates from V + Particle + Object particle placement (e. g., *I looked up the word*) to V + Object + Particle particle placement (e. g., *I carried the garbage out*). The rightmost graph plots switch rates from BE GOING TO-based future marking (e. g., *I'm gonna see Jim*) to WILL-based future marking (e. g., *I will see Jim*).

If there were no persistence effects, switch rates would cluster close to the diagonal line, which would indicate that switch rates are proportional to the overall distribution of switched-to forms (B forms); this constitutes the null hypothesis. The more the dots are clustered *below* the diagonal line, the less speakers switch. Note that since these graphs are sensitive to intra-speaker persistence only (recall that they plot successive dependencies in the speech of individual speakers only), they present a rather conservative estimate of α-persistence.

For all three alternations, we can reject the null hypothesis – everywhere, the dots cluster more or less heavily *below* the diagonal line, and the slopes of the heavy regression lines are much flatter than the slope of the diagonal line.[6] Speakers in all three corpora switch markedly less between variant forms than pure chance would predict, ergo there is a clear pattern of α-persistence. Observe, however, that there are differences between the alternations with regard to the strength of the effect. The switch rate is overall highest – meaning that α-persistence is weakest – in particle placement in FRED, where the regression line is steepest. In contrast, the switch rate is overall lowest – meaning α-persistence is strongest – in comparison strategy choice in the BNC-CG, where the regression line is almost horizontal. The switch rate for future marking in the BNC-DS ranges in between. The following sections will subject this somewhat impressionistic measure of persistence to a more fine-grained, multivariate analysis.

4.1. Persistence in Comparison Strategy Choice

This section will construct a multivariate model of comparison strategy choice. As is well known, there are two strategies in English to form comparatives: synthetic comparison using *-er*, as in (4a), and analytic comparison using *more*, as in (4b).

(4) a. If the new, *friendlier* systems do come onto the market, … people will just learn to use them. (BNC-DS KRG 514)

b. You talked a lot about computers being *more friendly* in the future than in the past. (BNC-DS KRG 469)

The rule of thumb governing the alternation is that monosyllabic adjectives take synthetic comparison, adjectives with more than two syllables take analytic comparison, and disyllabic adjectives alternate in the comparison strategy they take (for instance, Quirk et al. 1985; Bauer 1994). It is precisely this variation in disyllabic adjectives (and, additionally, in some monosyllabic and trisyllabic ones) that this section will subject to analysis.[7] For reasons explicated above, this analysis also includes the following factors which have previously been claimed to be determinants of comparison strategy choice:

Length of the synthetically inflected form in syllables (henceforth: LENGTH*).* For instance, *cheaper* has a length of two syllables.

Hypothesis: As the (potential) length of the synthetically inflected form of the adjective increases, so does the likelihood that the adjective takes analytic comparison and not synthetic comparison.

Morphological properties of the adjective (henceforth: MORPHOLOGY*).* Does the adjective which takes comparison begin in *un-* (as *unhappy*) or end in *-y* (as *lucky*; coded 0 if such affixes are not present and 1 if they are)?

Hypothesis: Presence of such affixes increases the likelihood for synthetic comparison (see, e. g., Leech and Culpeper 1997: 358–359; Quirk et al. 1985: 462).

Stress placement (henceforth: STRESS*).* If the adjective which takes comparison is polysyllabic, is it stressed on the final syllable (e. g., *complete*; coded 1 for final stress and 0 otherwise)?

Hypothesis: If the adjective is stressed on the final syllable, synthetic comparison has been claimed to be more probable (for instance, Kuryłowicz 1964: 15).

Text frequency (henceforth: FREQUENCY*).* What is the text frequency of the base form of the adjective under analysis in the spoken section of the BNC? *Poor*, for instance, has a text frequency of 1,031 occurrences *pmw* in the spoken section of the BNC.

Hypothesis: Frequently used adjectives have a preference for synthetic comparison (Bolinger 1968: 120; Quirk et al. 1985: 463; Mondorf 2003).

Syntactic function (henceforth: SYNTAX*).* Does the adjective occur in attributive function or in another function (coded 0 for attributive function and 1 for other functions)? Cohen's kappa, which measures the proportion of the best possible improvement over chance, was used to evaluate intercoder reliability of this annotation. A second coder, a native speaker and trained linguist, re-coded a random subset ($N = 50$, ca. 10% of the entire sample); comparison of the two samples yielded a simple agreement rate of ca. 96% and an 'excellent' (cf. Orwin 1994) kappa value of ca. 0.90. See Appendix B for the coding scheme.

Hypothesis: According to the literature, when the adjective occurs in predicative rather than attributive function, analytic comparison is favored (Braun 1982: 116; Leech and Culpeper 1997: 366; Mondorf 2003: 286–287).

Premodification by degree modifiers (henceforth: DEGREEMOD*).* Is the adjective preceded by one of the following degree modifiers: *much, even, far, bit, little, lot, times, noticeably, slightly, marginally* (coded 1 for degree modifiers present and 0 otherwise)?

Hypothesis: If the adjective is preceded by a degree modifier, analytic comparison is more likely (Braun, 1982: 116; Leech and Culpeper 1997: 366).

Presence of verbal complements (henceforth: COMPLEMENT*).* Is the adjective followed by prepositional or infinitival complements (coded 1 for verbal complements present and 0 otherwise)?

Hypothesis: The presence of complements increases cognitive complexity of the syntagm, which is why speakers presumably prefer the more explicit analytic option in such environments (for instance, Mondorf 2003: 254).

In addition, one further, persistence-related predictor (in addition to PREVIOUS and TEXTDIST) was included in logistic regression:

Presence of a token triggering analytic comparison in the preceding context (henceforth: more-trigger*).* Referring to sites not necessarily alternating, this predictor is meant to tap β-persistence. MORE-TRIGGER was coded 1 when the token *more* occurred in a context of 25 words (an arbitrary threshold) prior to CURRENT, and 0 otherwise.

Hypothesis: An occurrence of the token *more* (not necessarily in an analytic comparative, but in generic contexts such as *Tom ate more than Mary*) will help trigger analytic comparison, rather than synthetic comparison, at the next opportunity. The variable is also sensitive to whether parallelism in coordinated adjective phrases (cf. Leech and Culpeper 1997; Lindquist 2000; Mondorf 2000) obtains.

Analysis of the BNC-CG yielded a database of in all $N = 533$ relevant adjectives taking either synthetic or analytic comparison. Table 1a[8] dis-

Table 1. *Comparison strategy choice in the BNC-CG: logistic regression estimates*

	odds ratio (exp(*b*))
a. 'traditional' predictors	
LENGTH	0.12 ***
MORPHOLOGY(1)	0.17 ***
STRESS(1)	0.21 ***
FREQUENCY	0.99 ***
SYNTAX(1)	0.31 ***
DEGREEMOD(1)	1.09
COMPLEMENT(1)	0.30 *
b. persistence-related predictors	
PREVIOUS(ANA)	0.03 ***
PREVIOUS(ANA) * TEXTDIST	1.34 **
MORE-TRIGGER(1)	0.25 *
model intercept	4,470 ***
N	533
model chi-square	369.08 ***
Nagelkerke R^2	0.672
% correct (% baseline)	85.4 (57.8)

* significant at $p < .05$, ** significant at $p < .01$, *** significant at $p < .005$. Predicted odds are for synthetic comparison.

plays how the 'traditional' predictors hitherto discussed in the literature influence comparison strategy choice in this database (for simplicity, interactions between these 'traditional' predictors were not included in the model). Except for DEGREEMOD, all of these are selected as significant. LENGTH, SYNTAX, and COMPLEMENT have the hypothesized effect: increased length of the adjective taking comparison, usage of the adjective in non-attributive function, and the presence of verbal complements all make synthetic comparison less likely, as the odds ratios smaller than 1 indicate. But MORPHOLOGY, STRESS, and FREQUENCY turn out not to have the effect claimed in the literature: my analysis finds that (i) the presence of affixes such as − *y* or *un-* on the adjective,[9] (ii) stress on the final syllable, and (iii) high text frequency of the adjective make synthetic comparison *less* likely. This is not the place to extensively discuss or re-evaluate previous claims about these predictors; yet, two remarks can be made: first, mine is − to the best of my knowledge − the first multivariate analysis of comparison choice. Such analyses may correct for statistical artefacts that may go unnoticed in univariate analyses and therefore lead to false claims. Second, my findings derive from spoken data (most previous research on comparison is

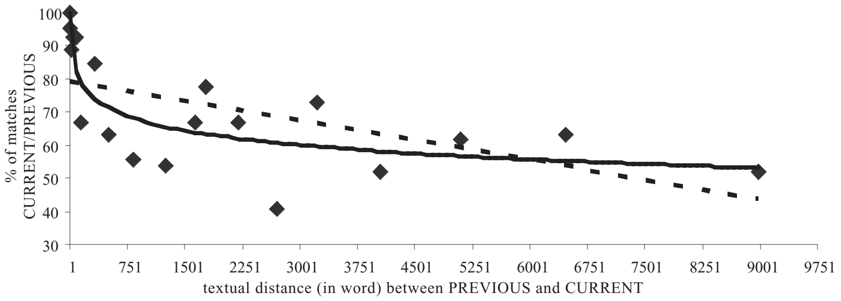


Figure 2. *Percentage of persistent pairs (i. e., PREVIOUS / CURRENT pairs where the same comparison strategy is used) as function of textual distance (in words) between CURRENT and PREVIOUS. Heavy line represents logarithmic estimate of the relationship, dotted line represents linear estimate of the relationship*

based on written data), in which some predictors may behave differently from how they do in written registers.

Table 1b shows how persistence interacts with comparison strategy choice. The main effect of PREVIOUS is associated with an odds ratio of 0.03. This means that if analytic comparison was employed in the first slot of a pair of successive variable sites in discourse, the odds for synthetic comparison in the second variable site shrink by 97%. In other words, there is a very marked tendency to avoid switching between comparison strategies and instead to go for the option used previously. In this study's terminology, this is α-persistence.

Note, however, that the extent of α-persistence between PREVIOUS and CURRENT is actually dependent on the textual distance (TEXTDIST) between PREVIOUS and CURRENT. This is evidenced by the significant interaction term PREVIOUS(ANA) * TEXTDIST: statistically, for every one-unit increase in TEXTDIST, the odds ratio associated with the main effect of PREVIOUS is changed by a multiplicative factor of 1.34. Another way of saying this is that, exactly as hypothesized, α-persistence between two successive variable sites in discourse weakens as textual distance between these sites increases. Figure 2 scrutinizes this relationship by plotting the non-logged textual distance (in words) between PREVIOUS and CURRENT against the percentage of persistent PREVIOUS / CURRENT pairs,[10] visually confirming two things: first, the percentage of matched pairs clearly does not bob around randomly. Instead, the more recently a comparison strategy choice has been made, the more likely speakers are to go for the same comparison strategy at the next opportunity. Recency of use thus clearly plays a role. Second, and also as hypothesized, the forgetting function that describes this relationship is logarithmic rather than linear (recall that I had modeled TEXTDIST logarithmically in logistic regression): the heavy, logarithmic regression line fits the data much better (adjusted $R^2 = 0.67$, F (1, 17) = 37.12, $p < 0.001$) than the linear estimate (dotted line; adjusted $R^2 = 0.30$, F (1, 17) = 8.60, $p < 0.01$).

We had also postulated that β-persistence played a role in comparison strategy choice such that if the trigger *more* was used up to 25 words prior to a slot for which a comparison strategy choice had to be made, speakers would be more likely to go for analytic comparison than they would otherwise. This, too, is confirmed by the analysis: if the above condition is met, the odds for synthetic comparison in CURRENT decrease by 75% (exp(*b*) = 0.25). Therefore, the odds for analytic comparison in CURRENT increase if the trigger *more* has been used recently.

Finally, let us determine statistically to what extent consideration of persistence helps us to understand comparison strategy choice. Collectively, the predictors displayed in table 1 explain a very decent two thirds of the observable variance ($R^2 = 0.672$). If the model had to exclusively rely on 'traditional' predictors (table 1a), explained variance would be 61% only. Crucially, the persistence-related predictors (PREVIOUS, TEXTDIST, and MORE-TRIGGER) in table 1b enhance model chi-square significantly (step chi-square = 47.39, df = 3, $p < 0.001$), and account for an extra 6% of observable variation that would be left unaccounted for otherwise. On the whole, the model in table 1 predicts 85.4% of speakers' linguistic choices correctly.

4.2. Persistence in Particle Placement

Let us next investigate particle placement with regard to persistence. 'Particle placement' refers to the variation observable in the particle / direct object word order in transitive, separable phrasal verbs in English ("type II transitive phrasal verbs" in Quirk et al.'s [1985: 1153] diction). Consider (5a), where the verb and its particle are separated by the direct object, and (5b), where they are adjacent:

(5) a. Mary *looked* the word *up*.
b. Mary *looked up* the word.

While the two word order patterns in (5a) and (5b) are certainly semantically equivalent, they are different formally and probably pragmatically and discourse-functionally (see, for instance, Bolinger 1971; Fraser 1965, 1966, 1974; Gries 2003b). A vast number of factors influencing particle placement have been suggested in the literature. Of these 'traditional' factors, I included in my analysis the factors that have been shown empirically to have substantial explanatory value for the alternation (see, for instance, Gries 2003a: table 6):

Definiteness of the direct object (henceforth: DEFINITEDO*).* Does the phrasal verb construction under analysis contain a direct object that is determined by a definite determiner (coded 1 for a definite determiner present, and 0 otherwise)? A test of intercoder reliability of the annotation of this feature, which was computed by having the author and a second scorer (a trained linguist) code a sample of $N = 102$ phrasal verb constructions for the feature, yielded a simple agreement rate of ca. 96% and an 'excellent' (cf. Orwin 1994) Cohen's kappa value of ca. 0.91. See Appendix B for the feature's coding scheme.

Hypothesis: If the determiner of the direct object is definite, there is a preference for the V + Particle + Object pattern (for instance, Gries 2003a: table 2). A typical example is *Fletcher and me went to* bring in *the sheep* (FRED WES019).

News value of the direct object (henceforth: NEWSVALUEDO*).* This variable is meant to assess the news value of the direct object. It was coded 0 if the referent of the direct object is not mentioned in the preceding five sentences, and 1 otherwise (i. e., if it is discourse-old).

Hypothesis: If the direct object is discourse-new, the V + Particle + Object pattern is more likely (for instance, Bolinger 1971).

Length of the direct object in syllables (henceforth: SYLLABLESDO*).* This independent measures length, or weight, of the direct object in syllables.

Hypothesis: The longer the direct object, the greater the preference for the V + Particle + Object pattern (the essence of this predictor boils down to Behaghel's [1909/1910] principle of 'end weight'). For instance, the direct object in *they filled* the bucket *up* (FRED SFK011) commands three syllables.

Complexity of the direct object (henceforth: COMPLEXITYDO*).* Does the direct object of the phrasal verb contain embedded clauses (coded 1 for embedded clauses present in the direct object, and 0 otherwise)?

Hypothesis: This is another predictor that is related to end weight. Presence of embedded clauses in the direct object will make the V + Particle + Object pattern more likely (for instance, Gries 2003a: table 2). A typical example is *pick out the ones* that you are going to use for seed (FRED HEB021).

Literalness of the phrasal construction (henceforth: LITERALNESS*).* Does the phrasal verb under analysis have a rather literal/spatial meaning, or a rather idiomatic meaning (coded 0 if the construction has a rather idiomatic meaning and 1 if it has a rather literal meaning)? All verb occurrences were coded individually, taking into account their respective context. Because coding for this feature reliably is not trivial, a test of intercoder reliability was, once again, performed. After initially poor Cohen's kappa values in the 0.5 range, re-coding by a trained linguist of a random sample of ca. 10% ($N = 102$) of the present database yielded, after quite some training, a simple agreement rate of ca. 87% and a moderately satisfactory Cohen's kappa value of 0.74 (see Appendix B for the feature's coding scheme).

Hypothesis: Constructions with more literal or spatial meanings will prefer the V + Object + Particle pattern (for instance, Biber et al. 1999: 933). A typical example for a literal phrasal construction is *bring the garbage out*, a typical example for an idiomatic meaning is *to figure out something*.

Presence of a directional prepositional phrase after the VP (henceforth: DIRECTIONALPP*).* Is the phrasal VP followed by a directional prepositional phrase (coded 1 if it is, and 0 if it is not)?

Hypothesis: If the phrasal VP is followed by a directional prepositional phrase, we expect a preference for the V + Object + Particle pattern (for instance, Fraser 1966). A typical example would be *We were sending cattle off* to the mainland (FRED LAN012).

Distinctive collostruction strength of the phrasal construction (henceforth: DISTINCTIVENESS*).* Biber et al. (1999:933) were not the first to point out that "there is considerable variability among individual phrasal verbs in their preference for […] particle placement". To account for this variability, this section's analysis incorporated results from Gries and Stefanowitsch's (2004) 'distinctive collexeme analysis' in which they extracted 700 verbs from the ICE-GB corpus and determined the collostructional strengths associated with them (i.e., basically whether and to what extent each of these verbs prefers the V + Object + Particle or V + Particle + Object pattern) by means of a statistical analysis. My analysis operationalized Gries and Stefanowitsch's findings through the scalar variable DISTINCTIVENESS, which can take values between 0 and 100 and is based on Gries and Stefanowitsch's findings.[11]

Hypothesis: The higher a phrasal verb's DISTINCTIVENESS score, the more marked this phrasal verb's preference for the V + Object + Particle pattern. To illustrate: *find out* (as in *the examiner'd* find *these little faults* out, FRED SAL030) has a comparatively high DISTINCTIVENESS score (99.99) and is therefore strongly associated with the V + Object + Particle pattern. The opposite is true for the verb *send back*, which has a comparatively low distinctiveness score (1.49).

FRED dialect area (henceforth: FRED-AREA*).* This independent accounts for variation in particle placement between dialect areas in FRED. If left unaccounted for, this variation would cause unnecessary noise in the analysis.

Besides PREVIOUS and TEXTDIST, my examination of persistence in particle placement also utilized the following predictors:

Same verb lemma in both previous and current (henceforth: VLEMMAID*).* This variable checks whether two neighboring transitive phrasal verb constructions involve the same phrasal verb (though not necessarily the same verb form; coded 1 if the lemma was the same, and 0 if it was not). Pickering and Branigan (1999) showed that production priming is stronger when the priming verb lemma and the target verb lemma match; Gries (forthcoming) also obtained the effect through corpus study.

Hypothesis: If the verb lemma of two successive phrasal verb variables matches, α-persistence is even stronger than it would be otherwise. A typical example of two successive phrasal sites where both the verb lemma and the placement strategy matches would be *they* take *your beams* out … *we were doing so bad in the mill they* took *your beams* out (FRED LAN009).

Length of the sentence in which the variable under analysis is embedded (henceforth: SENTENCELENGTH*).* Sentence length (in words) will be taken to be a proxy for syntactic complexity of the environment where CURRENT is embedded (see Szmrecsanyi 2004 for why considering sentence length a proxy for syntactic complexity is justified).

Hypothesis: As syntactic complexity of the context where CURRENT is embedded increases, online processing constraints become more acute. Hence, we expect an interaction between PREVIOUS and SENTENCELENGTH such that for increasing values of SENTENCELENGTH, persistence effects grow more potent due to their facilitatory effect on online processing (cf. Tannen 1987, 1989 on cognitive efficiency of repetitiveness).

Analysis of the FRED subset yielded *N* = 1,048 phrasal verb constructions (it should be added that phrasal verb constructions whose object was pronominal were not included in the dataset since pronominal objects near-categorically yield the V + Object + Particle pattern). A logis-

Table 2. *Particle placement in FRED: logistic regression estimates*

	odds ratio (exp(*b*))
a. ‘traditional’ predictors	
DEFINITEDO(1)	1.47 *
NEWSVALUEDO(1)	1.49 *
SYLLABLESDO	0.67 ***
COMPLEXITYDO(1)	0.06 *
LITERALNESS(1)	2.56 ***
DIRECTIONALPP(1)	4.97 *
DISTINCTIVENESS	0.99 ***
FRED-AREA	– – ***
b. persistence-related predictors	
PREVIOUS(*V* + *Part* + *NP*)	0.17 **
PREVIOUS(*V* + *Part* + *NP*) * TEXTDIST	1.04
PREVIOUS(*V* + *Part* + *NP*) * VLEMMAID(1)	0.39 *
PREVIOUS(*V* + *Part* + *NP*) * SENTENCELENGTH	1.02 ***
model intercept	1.23
N	1,048
model chi-square	335.99 ***
Nagelkerke R^2	0.412
% correct (% baseline)	84.8 (76.3)

* significant at $p < .05$, ** significant at $p < .01$, *** significant at $p < .005$. Predicted odds are for the V + Object + Particle pattern.

tic regression model on this database was then estimated, the results of which are shown in Table 2. All of the predictors traditionally cited to explain particle placement (Table 2a) play out as expected in this model (note that for reasons of simplicity, interactions between ‘traditional’ predictors were not included in the model), and the results are roughly compatible with Gries’s (2003b) multivariate analysis of particle placement. If the direct object is definite (DEFINITEDO), or if it is discourse-old (NEWSVALUEDO), the odds for the V + Object + Particle pattern increase by ca. 50 %. Increasing length of the direct object (SYLLABLESDO) reduces the odds for the V + Object + Particle pattern by a considerable one third for every one-syllable increase in the direct object (exp(*b*) = 0.67). Even more impressing is the effect of the presence of embedded clauses in the direct object (COMPLEXITYDO): presence of an embedded clause in the direct object reduces the odds for the V + Object + Particle pattern by more than 90 % (exp(*b*) = 0.06). By contrast, if the phrasal verb under analysis has a rather literal, or spatial meaning, as in *he brought the garbage out* (LITERALNESS), the odds for the V + Object + Particle pattern are multiplied by a factor of ca.

2.5, and if a directional prepositional phrase follows the phrasal verb phrase (DIRECTIONALPP), the odds for the V + Object + Particle pattern increase almost five-fold (exp(*b*) = 4.97). Gries and Stefanowitsch's (2004) collostruction strength scale (DISTINCTIVENESS) also turns out to be a significant predictor of particle placement. For each one-unit increase in the scalar variable DISTINCTIVENESS, the odds for the V + Object + Particle pattern decrease by 1 %. This relationship has the expected direction. On the whole, the predictor DISTINCTIVENESS accounts for ca. 5 % of the observable variance in particle placement in my data. Note that Gries and Stefanowitsch's (2004) 'distinctive collexeme' scores were derived from the ICE-GB, a corpus of spoken and written standard British English. Given that these scores were applied to a corpus of English dialects, the share of variance accounted for by the variable is actually considerable. Finally, dialect areas significantly help predicting particle placement in FRED. Taking the Southeast as statistical baseline area, there is a significant dispreference for the V + Object + Particle pattern in the Hebrides, and also one, though slighter, in the Midlands. In the North of England, the V + Object + Particle pattern tends to be more frequent than elsewhere.

Table 2b provides regression estimates on how the persistence-related predictors PREVIOUS, VLEMMAID, TEXTDIST, and SENTENCELENGTH affect particle placement. The main effect of previous is associated with an odds ratio of 0.17. This means that when in two successive phrasal verb slots, the V + Particle + Object pattern is employed in the first slot, the odds that speakers will switch to the V + Object + Particle pattern in the second slot are 84 % lower than the odds that they will not switch. However, PREVIOUS interacts with the other predictors such that PREVIOUS' effect size actually changes for different values of VLEMMAID, TEXTDIST, and SENTENCELENGTH.

First, observe that the interaction term PREVIOUS VLEMMAID is associated with an exp(*b*) value of 0.39. This means that if the same phrasal verb lemma is used in two successive variable sites, the main effect of PREVIOUS (0.17) is multiplied by a factor of 0.39. In other words, α-persistence is even stronger when two successive phrasal verb constructions involve the same verb lemma than it would be otherwise.[12] Therefore, much like Pickering and Branigan (1999) and Gries (forthcoming), my analysis finds that priming is stronger if prime and target involve the same verb lemma, as in *he 'd* fill *all their bags* up … *he wouldn't* fill *our bags* up (FRED LND001).

Second, while the interaction term PREVIOUS TEXTDIST is associated with an exp(*b)* value greater than one (which is the expected effect direction), the term is not selected as significant in logistic regression. However, figure 3 – which plots matching PREVIOUS/CURRENT pairs

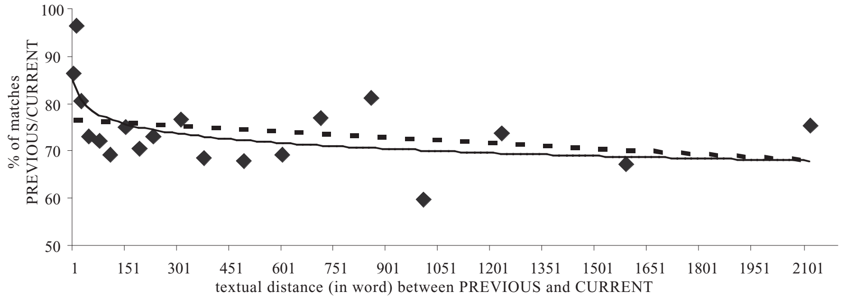


Figure 3. *Percentage of persistent pairs (i. e.,* PREVIOUS / CURRENT *pairs where the same particle placement strategy is used) as function of textual distance (conceptualized in terms of 20-tiles) between* CURRENT *and* PREVIOUS. *Heavy line represents logarithmic estimate of the relationship, dotted line represents linear estimate of the relationship*

against non-logged textual distance between PREVIOUS and CURRENT – strongly suggests that in fact, the hypothesized relationship between PREVIOUS and TEXTDIST obtains. For one thing, figure 3 documents that the likelihood that CURRENT matches previous' particle placement pattern increases when textual distance between previous and current decreases – and thus, that persistence is stronger when exposure to the last variable slot has been recent. Moreover, this forgetting function is best modeled as logarithmic rather than linear, both intuitively and statistically (heavy line: adjusted $R^2 = 0.33$, F (1, 17) = 9.72, $p < 0.01$; dotted line: adjusted $R^2 = 0.04$, F (1, 17) = 1.68, $p > 0.05$).

Third, the interaction PREVIOUS SENTENCELENGTH has a moderate exp(*b*) value of 1.02. Therefore, when sentence length increases – and hence, when syntactic complexity of the environment surrounding CURRENT increases –, the impact of PREVIOUS on CURRENT decreases. This finding is not expected in that it implies that α-persistence *weakens* in syntactically complex environments. Recall that given the discourse-analytic literature (e. g., Tannen 1987, 1989), one would have expected the opposite.

To conclude the discussion of particle placement, I will turn to what is gained analytically by considering persistence in analyses of particle placement. The model as displayed in table 2 accounts for a moderate 41.2% of the observable variation in the FRED subset and predicts correctly 84.8% of speakers' choices. If the model did not include persistence-related factors, explained variance would be 37.2% only, and the model would predict correctly only 83.1% of speakers' particle placement decisions. Hence, persistence explains an extra 4% of the observable variation, and improves predictive efficiency (albeit not in a statistically significant way) by 1.7%. These increases are associated with a statistically significant enhancement of model chi-square (step chi-square = 39.73, df = 4, $p < 0.001$).

4.3. Persistence in Future Marker Choice

This study will now move on to an analysis of what role persistence plays in future marker choice. English possesses two highly grammaticalized syntactic options for overtly expressing futurity, BE GOING TO and WILL. Each of these paradigms also has variant forms (which will not be distinguished for the remainder of this section): *be going to*, as in (6a), and *gonna*, as in (6b); *will*, as in (6c), *'ll*, as in (6d), and *won't*, as in (6e).[13]

(6) a. I *am going to* go to London tomorrow.
b. I'm *gonna* go to London tomorrow.
c. I *will* go to London tomorrow.
d. I'*ll* go to London tomorrow.
e. I *won't* go to London tomorrow.

Although it is clear that there are semantico-pragmatic nuances in meaning between the above variants, and between BE GOING TO and WILL as future markers in general, many researchers now argue that there is rough semantic equivalence between these options.[14] In exactly this spirit, I will set out to model the variation in future reference in a strictly variationist way. Variation in English future time reference that is neither semantically conditioned nor extralinguistic is not exceedingly well researched. Of the handful or so of factors discussed in the literature, I included in my analysis one that is relatively straightforward to operationalize:

Contexts of negation (henceforth: NEGATION*).* Is the future marker negated by *not*, or by a *not*-contracted auxiliary (coded 0 for affirmative contexts and 1 for negated contexts)?

Hypothesis: Berglund (1999, 2000) and Szmrecsanyi (2003) report that BE GOING TO is preferred over WILL in contexts of negation.

As for persistence-related predictors, my analysis modelled the following predictors (in addition to PREVIOUS and TEXTDIST):

Was PREVIOUS *in the same turn as* CURRENT *(henceforth:* SAMETURN*)? Was it produced by the same speaker that produced* CURRENT *(henceforth:* SAMESPEAKER*)?* These binary independents are about whether the effect size of persistence is sensitive to turn taking (coded 1 if PREVIOUS and CURRENT are in the same turn [SAMETURN], or if PREVIOUS and CURRENT are produced by the same speaker [SAMESPEAKER], and 0 otherwise).[15] Note that, naturally, the two variables are somewhat correlated with TEXTDIST, but not at a level which would be problematic for regression analysis (cf. the collinearity measures in Appendix A).

Hypothesis: Persistence effects across turns and persistence effects within turns have different strengths, and depending on whether persistence comes about through self-repetition or allo-repetition, persistence effects vary in size.

Type-token ratio of the lexical environment where current *is embedded (henceforth:* TTR*).* TTR will be considered a proxy for lexical density. 'Lexical environment' refers to a textual context of 50 words before and 50 words after CURRENT. The factor complements the discussion of the variable SENTENCELENGTH in the previous case study.[16]

Hypothesis: The larger TTR, and hence, the higher lexical density of the context where CURRENT is embedded, the more sizable the persistence effects (cf. Tannen 1987 on why parallel patterns might be preferred in lexically dense contexts due to processing efficiency advantages). We thus expect a significant interaction effect between TTR and PREVIOUS.

Presence of the verb to go in the preceding context (henceforth: GO-TRIGGER*).* Do the tokens *go, goes, went, going*, or *gone* occur anywhere in a context of 75 words (an arbitrary threshold) to CURRENT?

hypothesis: The presence of the verb *to go* may trigger a BE GOING TO based future marker through lexical priming or similar mechanisms, a triggering effect which would qualify as β-persistence in the terminology of the present study.

Table 3a presents a logistic regression estimate on how well negation predicts future marker choice in the BNC-DS, based on $N = 33{,}558$ relevant observations. As can be seen, contexts of negation clearly favor usage of BE GOING TO, much as claimed by, e. g., Berglund (1999, 2000): when CURRENT occurs in a negated context, the odds for a WILL marker decrease by a substantial 92 % ($\exp(b) = 0.08$). Yet, the predictor NEGATION accounts for only 4.8 % of the observable variation in future marker reference. One reason is that negation is a marked phenomenon, and thus relatively rare compared to affirmative contexts – thus, the variable NEGATION has a very limited scope in the dataset.

Table 3. *Future marker choice in the BNC-DS: logistic regression estimates*

	odds ratio (exp(*b*))
a. 'traditional' predictors	
NEGATION(1)	0.08 ***
b. persistence-related predictors	
PREVIOUS(BGT)	0.01 ***
PREVIOUS(BGT) * TTR	1.07 ***
PREVIOUS(BGT) * SAMETURN(1)	0.64 ***
PREVIOUS(BGT) * SAMESPEAKER(1)	0.77 ***
PREVIOUS(BGT) * TEXTDIST	1.15 ***
GO-TRIGGER(1)	0.94 **
model intercept	0.00
N	35,558
model chi-square	1,932.11 ***
Nagelkerke R^2	0.120
% correct (% baseline)	72.0 (68.6)

* significant at $p < .05$, ** significant at $p < .01$, *** significant at $p < .005$. Predicted odds are for will marking.

The quality of the regression model is improved significantly (step chi-square = 1932.11, df = 6, $p < 0.001$) when persistence-related predictors are factored in. This step increases R^2 by some 7 per cent points, so that the model in table 2 – crude as it still is – now explains 12 % of the observed variance in future marker reference. Predictive efficiency is increased from 70.5 % to 72.0 %, a differential which is statistically significant (chi-square = 4.20, df = 1, $p < 0.05$). Table 3b gives logistic regression estimates of how the persistence-related predictors affect future marker choice. To begin with, consider the main effect of previous, which is associated with an odds ratio of 0.01: when a BE GOING TO marker is used in a given slot, the odds that a WILL marker will be used next time are diminished by 99 % (conditioned on the interactional factors being zero). In other words, a given BE GOING TO marker is, due to α-persistence, highly likely to be followed by another BE GOING TO marker instead of a WILL marker.

However, as the several interaction terms with PREVIOUS indicate, the exact strength of α-persistence is dependent on the values of several variables. First, the exp(*b*) value associated with the interaction PREVIOUS TTR indicates that for every one-unit increase in TTR, the main effect of PREVIOUS changes by a multiplicative factor of 1.07. Contrary to my hypothesis, α-persistence between two variable sites hence *weakens* as lexical density increases; given Tannen (1987), one would have ex-

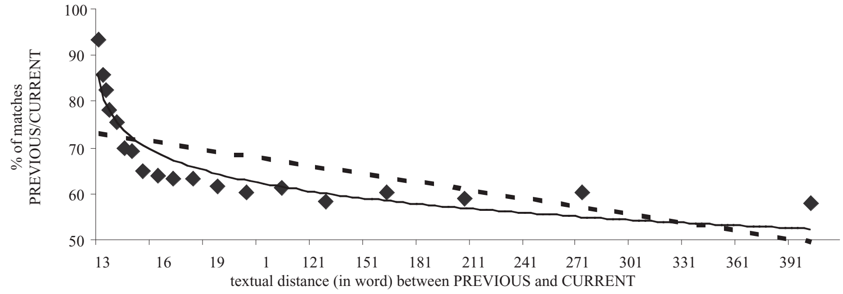


Figure 4. *Percentage of persistent pairs (i. e.,* PREVIOUS / CURRENT *pairs where the same future marker is used) as function of textual distance (conceptualized in terms of 50-tiles) between* CURRENT *and* PREVIOUS*. Heavy line represents logarithmic estimate of the relationship, dotted line represents linear estimate of the relationship*

pected repetitiveness to be functionally exploited to relax informationally dense contexts. A tentative explanation for this finding is that higher lexical density is indicative of discourse that involves better planning and monitoring; this may arguably weaken the effect of a partly subconscious phenomenon such as persistence. Second, the interactions involving the turn-by-turn variables SAMETURN and SAMESPEAKER are both associated with quite similar exp(*b*) values of 0.65 and 0.77, respectively. Thus, the effect of a previous future marker choice on an upcoming choice (e. g., α-persistence) is stronger when (i) the previous future marker occurrence was in the same turn as the upcoming slot, and when (ii) the previous future marker occurrence was produced by the same conversational party that is faced with the upcoming choice. As hypothesized, persistence within turns is stronger than persistence across turns, and intra-speaker persistence ('self-repetition', or 'production-to-production priming' in psycholinguistic parlance) is stronger than inter-speaker persistence ('allo-repetition', or 'comprehension-to-production priming' in psycholinguistic parlance).

Third, we again obtain a significant interaction between PREVIOUS and TEXTDIST such that for every one-unit increase in the *ln* of textual distance between PREVIOUS and CURRENT, the main effect of PREVIOUS on CURRENT is changed by a multiplicative factor of 1.15 − which is equivalent to saying that as textual distance between two variable sites increases, α-persistence between these sites weakens. Figure 4 visualizes the nature of this relationship by plotting the percentage of PREVIOUS-CURRENT matches against textual distance. One can see that this relationship is nicely logarithmic (or decreasing exponential): a logarithmic regression fits the data much better (adjusted $R^2 = 0.86$, F (1, 17) = 109.29, $p < 0.001$; heavy line) than a linear regression (adjusted $R^2 = 0.34$, F (1,17) = 10.19, $p < 0.01$; dotted line). Once again, there is evidence that α-persistence is subject to a forgetting function, and that this function is logarithmic.

The model in table 2 also includes GO-TRIGGER, a predictor meant to tap β-persistence. I had hypothesized that a generic occurrence of the verb *to go* could trigger a BE GOING TO future marker instead of a WILL marker at the next opportunity. This hypothesis is, indeed, borne out: when a form of the verb lemma *go* (as in *Mary went to school*) was used anywhere in a context of up to 75 words prior to a future marker slot, this decreases the odds for a WILL marker in CURRENT by 6% (exp(*b*) = 0.94). Much like *more* can trigger analytic comparison, therefore, generic *go* can trigger a BE GOING TO future marker through β-persistence.

5. Summary and Conclusion

By examining three grammatical alternations in English in three spoken registers (formal spoken English, colloquial English, and dialect speech in interview situations), I hope to have delivered convincing evidence that persistence is a factor worth considering in variationist research, and that (naturalistic) data derived from diverse corpora can match (experimental) psycholinguistic data. In the case studies analyzed in this paper, consideration of the factor clearly enhanced the explanatory power of our modeling of speakers' linguistic choices. Models omitting persistence would leave a substantial share of the observable variation unaccounted for, or even erroneously identify it as 'free' variation although it is clearly patterned.

More specifically, this study would seem to have suggested that any given variable site in discourse is sensitive to two major, hitherto rather neglected characteristics of the site's contextual environment. First, successive variable sites in discourse influence each other. For one thing, we saw that switch rates between two alternative options are considerably lower than chance switch rates. Along the same lines, logistic regression estimates showed that when a given option A was employed in the first of two successive variable sites in discourse, the odds that option B is used in the second site are reduced substantially − according to my analysis, by between 83% and 99%, depending on the alternation under analysis. I have termed this type of persistence *α-persistence*. At the same time, the effect size of α-persistence is itself a function of several determinants such as (i) textual distance between two successive variable sites (persistence decays with increasing textual distance); (ii) for purely structural alternations such as particle placement, whether two successive variable sites involve the same verb lemma (if they do, persistence is more powerful); (iii) turn-taking mechanisms: whether two successive variable sites are in the same conversational turn, and whether they are produced by the same conversational party (in both cases, persistence is stronger if the answer is yes); and (iv) syntactic and lexical complexity of the contextual environment. A second way in which persistence interferes with speakers' choices is the following: given a variable site where speakers have a choice between two or more options, that choice is not only influenced by other variable sites. It is also affected by non-variable linguistic patterns that share structural, lexical, or other characteristics with one of the choice options. This is what the present study has termed *β-persistence*. We established, for instance, that a non-comparative occurrence of *more* (as in *I would like more soup*) can help trigger an analytic comparative (which necessarily involves the token *more*) in a variable site nearby, much like a generic form of the verb *go* can trigger a BE GOING TO future marker.

More corpus-based research has to be carried out to uncover further determinants of persistence, both intralinguistic and extralinguistic. For instance, I point out elsewhere (Szmrecsanyi forthcoming) that persistence is sensitive to speaker characteristics such as sex and age. Quite fascinatingly, persistence appears to be a phenomenon where things such as memory limitations in elderly speakers (cf. Zurif et al. 1995) and greater innate fluency of female speakers (cf. Bortfeld et al. 2001) might interface with linguistic variation.

How is persistence relevant to linguistic theory and practice? To begin with, the relatively strong empirical showing of the phenomenon plays methodical havoc with a standard assumption underlying most empirical linguistic research: namely, that an occurrence of a linguistic pattern can and should be considered the result of a new throw of the dice, and that it can be investigated in isolation and out of the wider discourse context. This is, first, a problem for qualitative linguistic inquiry where, often, a data fragment is investigated asking, 'why did the speaker use this specific option, instead of the alternative one, here?'. The present study leaves us good reason to think that the answer might often be as simple as 'because the speaker had just used that option – or some trigger – before'. Secondly, persistence also poses a problem to some varieties of quantitative linguistic research in that text frequencies of some linguistic pattern may be misleading unless, for instance, textual distances between the individual hits are factored in. A brief example will illustrate this: Szmrecsanyi (2003: table 2) claimed that in the BNC-DS, the distribution of BE GOING TO and WILL/SHALL is roughly 28:72 (cf. Berglund 1999 for similar figures). This figure, of course, does not take into account persistence. If the researcher chooses, for instance, to exclude all cases where two successive future marker slots are located in the same turn (because, as we have seen, persistence is quite powerful in such contexts), the distribution changes to roughly 30 : 70, a difference which is highly significant (chi-square = 37.1, df = 1, $p < 0.005$). If the researcher further opts to exclude all cases where textual distance between two future marker hits is less than 150 words – after this textual distance, one can be quite sure that persistence effects have dissipated –, the ratio changes to 32 : 68, which is, again, a highly significant difference (chi-square = 71.7, df = 1, $p < 0.001$). In other words, accounting for persistence in this case returns distributions which tend more towards a 50 : 50 distribution. While more extreme cases could be constructed, the above example goes to show that text frequencies may be distorted by persistence, and persistence – much like restarts, for instance – is not a factor that researchers interested in text frequencies would normally like to have included in their statistics.

On a more general level, persistence has the potential to be of theoretical interest to linguists engaged in very diverse research programs. Certainly, the present study has demonstrated that persistence, as an explanatory factor, is immediately relevant to all those who seek to account for the choices speakers make in the spirit of variationism or probabilistic grammar. Along somewhat different lines, persistence may be thought of as a type of short-term entrenchment. 'Entrenchment' (originally a Cognitive Grammar term) is a mechanism due to which the effect of discourse frequency on mental representations is such that these representations are strengthened through their activation in use (cf. Langacker 1987: 59 f.). It is true that entrenchment is understood as being a mechanism operating over longer intervals of time, possibly a speaker's lifetime – in contrast, persistence is a phenomenon that probably dissipates after a few minutes. Yet, persistence as well is due to linguistic patterns, or representations thereof, being activated through use; in this way, it may make sense to refer to persistence as "micro-entrenchment", and to entrenchment as "macro-persistence". Cognitive grammar aside, persistence is obviously interesting to mainstream functionalists since issues such as online processing constraints, economy, and discourse management are, as we have seen, involved in motivating surface structure. But also for less mainstream, more extreme functionalists who view grammar as an emergent system of meaningful repetition and as a "vast collection of hand-me-downs that reaches back in time to the beginnings of time" (Hopper 1998: 150), persistence should be a worthwhile phenomenon to consider. Maybe surprisingly, the existence of the phenomenon can even be seen as underlining the validity of the generative enterprise, for two reasons. First, persistence or parallelism in surface structure can potentially yield linguistic outcomes that are dysfunctional –

Scherre and Naro (1991: 30), for instance, have noted that due to speakers' inclination to maintain surface parallelism, morphological "markers tend to occur precisely when they are not needed and tend not to occur when they would be useful". Thus, persistence and functional factors can very well work against each other, for instance, in contexts where functional factors would license some option A, but due to persistence it is option B that is actually used. *Ex negativo*, this might lead one to argue that grammar cannot be motivated functionally alone; hence the need for formal analysis. Second (and relatedly), the fact that speech generation is sometimes heavily inertial and mechanical (insofar as the human speech processing system is skewed towards repetition) can be construed as evidence, albeit somewhat indirect, for the autonomy of syntax hypothesis. The point is that if speakers sometimes cannot help being persistent and repetitive (a claim that the present study certainly has not contradicted), the cognitive module which is responsible for syntax must be, to some extent at least, self-contained. Also note that behaviourists – had they not disappeared from the linguistic scene long ago – would find the stimulus-response pattern of persistence, repetitiveness, and prime-target pairs intriguing. Last but not least, persistence could also have implications for historical linguistics: the multiplicative and self-enforcing effect of persistence, coupled with logarithmic forgetting functions, might very well be involved in the S-curve patterns so often observable in language change. This is a point that would certainly merit scrutiny in future research.

Appendix A: Correlations between factors

Table 4 reports Variance Inflation Factors (VIFs) for all variables that were entered into logistic regression. VIFs measure the strength of interrelationships among explanatory variables in a multivariate model. Increasing VIFs indicate increasing regression coefficients, which may result in more unstable estimates. VIFs exceeding a value of 10 are commonly considered to indicate multicollinearity, but values above 2.5 may already be a cause for concern.

Appendix B: Coding schemes

Syntactic function (attributive vs. predicative) of adjectives: "Code '0' for attributive function (e. g., *the green house is there, I like red cars*). Code '1' for predicative function (e. g., *the house is green, the car seems nice, Jim became angry*)."

Definiteness of the direct object of transitive phrasal verbs: "Code definite direct objects of phrasal verbs as '1' (e. g., *Jim looked up the word*). Direct

Table 4. *Variance Inflation Factors (VIFs)*

comparison strategy choice		future marker choice		particle placement	
variable	*VIF*	*variable*	*VIF*	*variable*	*VIF*
LENGTH	3.26	DEFINITEDO	1.02	NEGATION	1.00
MORPHOLOGY	3.45	NEWSVALUEDO	1.07	PREVIOUS	1.01
STRESS	1.87	SYLLABLESDO	1.04	TTR	1.01
FREQUENCY	1.09	COMPLEXITYDO	1.04	SAMETURN	1.61
SYNTAX	1.07	LITERALNESS	1.09	SAMESPEAKER	1.18
DEGREEMOD	1.04	DIRECTIONALPP	1.03	TEXTDIST	1.44
COMPLEMENT	1.56	DISTINCTIVENESS	1.07	GO-TRIGGER	1.01
PREVIOUS	1.12	FRED-AREA	1.12		
TEXTDIST	1.04	PREVIOUS	1.09		
MORE-TRIGGER	1.03	TEXTDIST	1.13		
		VLEMMAID	1.12		
		SENTENCELENGTH	1.05		

objects tend to be preceded by a definite article, or by some kind of genitive or possessive pronoun. Code indefinite direct objects of phrasal verbs as ‘0’ (e. g., *Jim looked up a word*). Indefinite objects tend to be preceded by an indefinite article, or by no article at all.”

Literal vs. idiomatic meanings of phrasal verbs : “If the phrasal verb is literal, code ‘1’. Literal phrasal verbs are verbs where the meaning of the whole verb is the semantic sum of the verb and the particle. Often, literal phrasal verbs are phrasal verbs where some spatial movement is involved (for instance, *to bring in* is the semantic sum of *to bring* and *in*; also, some spatial movement is involved). If the phrasal verb is idiomatic, code ‘0’. A phrasal verb is idiomatic if the meaning is more than the semantic sum of verb and particle (if one needs to have learned its idiomatic meaning, therefore). Most often, idiomatic phrasal verbs are *not* spatial (for instance, *to figure out* means something else than the semantic sum of *to figure* and *out*; also, there is no spatial movement involved).”

Notes

* I am indebted to Manfred Krug and Stefan Th. Gries for many helpful comments and suggestions, and to Daniela Kolbe, Nuria Hernandez, and Tamsin Anderson for helping me establish intercoder reliability of my corpus annotation. Part of the work on which this paper is based was also supported by the Deutsche Forschungsgemeinschaft (DFG, grant No. Ko 1181/1−1,2). The author can be contacted at the English Department, University of Freiburg, Rempartstr. 15, D-79098 Freiburg, Germany. E-mail: < benedikt.szmrecsanyi@uni-freiburg.de>.

[1] In this study, quotes from the BNC will be identified by the respective text identifier plus line; quotes from FRED and the Corpus of Spoken American English

will be identified by their respective text identifiers only (the format of these corpora does not support static line numbers).

2. I avoid using psycholinguistic terminology ('priming', 'prime', 'target', etc.) a priori because corpus-based study may be inappropriate to explicitly investigate psycholinguistic mechanisms such as production priming effects (see Branigan et al. 1995 on this point, but Gries forthcoming for a dissenting opinion). In naturalistic data, speakers' output may exhibit persistence effects for reasons of rhetoric, politeness (for instance, Tannen 1982, 1987, 1987), or thematic coherence, to aid the process of gap filling in creating and processing elliptical utterances (for instance, Matthews 1979), to open up question-answer pairs (for instance, Levelt and Kelter 1982), because speakers feel like intentionally repeating items from previous discourse, or because they have been primed in preceding discourse – but it is not easily possible to disentangle the above motivations through corpus study in a waterproof fashion.
3. One referee wondered whether β-persistence does not go against some previous psycholinguistic findings, particularly against Bock and Loebell (1990), who found that the infinitive phrase in *Susan brought a book to study* did not prime the prepositional phrase in *Susan brought a book to Stella* as well as another prepositional phrase did (such as *The defendant told a lie to the crowded courtroom*). Yet, the present study's notion of β-persistence would seem to make exactly such a claim: that the token *to* in the infinitive phrase *to study* would facilitate a prepositional dative (*to Stella*) instead of a double-object dative. But, crucially, Bock and Loebell (1990) did not argue that the infinitive marker *to* could never prime the dative preposition *to*; they only stated that in their experiment, prepositional datives were far better syntactic primes than infinitive phrases, which has implications for theories about *syntactic* priming. In analogy, I have no intention of conveying the impression that β-persistence is more potent than α-persistence – the present study's point is that β-persistence is a statistical tendency which is observable in naturalistic data, and no claim is made as to precisely which psycholinguistic mechanism(s) might or might not be responsible this statistical tendency. As such, the notion of β-persistence is, in fact, fully consonant with Bock and Loebell's view that "people tend to say the same thing on successive occasions, [but, BS] it is rarely obvious what constitutes 'the same thing'" (Bock and Loebell 1990: 29).
4. This means that independents associated with very small effect sizes (for instance, 0.99) may be selected as significant in logistic regression, while independents with big effect sizes (for instance, 0.03) need not necessarily turn out as significant.
5. This subset consisted of the following texts: LAN008–LAN014, NBL001, NBL003, NBL006, NBL007, NBL008, WES001, WES002, HEB001–HEB041, SAL001–SAL039, WAR001, DUR001–DUR003, LAN001–LAN007, KEN006–KEN008, KEN014, LND001, LND002, SFK011–SFK033. These comprise ca. 1,000,000 words and thus ca. 40% of the entire FRED corpus; dialect areas included in the sample are the Hebrides, the Midlands, the North of England, and the Southeast.
6. For analytic vs. synthetic comparison, the regression line appears to be incredibly horizontal, given the distribution of the dots. Note, however, that the majority of dots sitting on the X axis represent more than only one speaker (often many more), to which the regression is of course sensitive.
7. More specifically, I will investigate comparative forms of the following 112 adjectives, which have been shown to take both synthetic and analytic comparison in previous research (Bauer 1994; Biber et al. 1999; Leech and Culpeper 1997: 125–132; Mondorf 2003: 251–304; Quirk et al. 1985): *able, acute, afraid, akin, ample,*

apt, aware, bitter, bizarre, blunt, bold, brittle, cheap, cheeky, clear, clever, common, compact, complete, correct, costly, cosy, crazy, cruel, curt, dead, deadly, dense, empty, exact, extreme, feeble, fierce, fit, fond, free, friendly, full, gentle, guilty, handsome, handy, humble, hungry, intense, just, keen, kindly, likely, little, lively, lonely, lovely, lowly, lucky, mature, mellow, narrow, nimble, noble, obscure, odd, pale, pleasant, polite, poor, precise, profane, profound, prone, proud, queer, quiet, rare, ready, real, remote, rich, right, risky, robust, rude, secure, severe, sexy, shallow, sick, silly, simple, sincere, slender, slow, sober, solid, sound, stable, stupid, subtle, sure, tender, trendy, tricky, true, ugly, unhappy, unwise, used, wealthy, wicked, worthy, wrong, yellow.

[8] In this table and in the following, the value in brackets following categorical independents indicates which category of the independent has been tested. Therefore, MORPHOLOGY(1) tests the presence (as opposed to the absence) of affixes on an adjective; this presence is associated with an exp(*b*) value of 0.17.

[9] I should add that there is a minor collinearity issue with LENGTH and MORPHOLOGY (cf. the correlation measures in Appendix A) such that adjectives ending in − *y* or starting in *un-* tend to be longer than other adjectives. Therefore, the effect of the extra length which suffixes such as − *y* or *un-* add to an adjective seems to be the dominant one in regression, making analytic comparison more likely.

[10] Figure 2 − exactly as figures 3 and 4 below − is based on 19 measuring points. These have been arrived at by dividing the observed textual distance between PREVIOUSand CURRENT into 20-tiles, i. e., into 20 equal groups (which have 19 cut-off points); the percentage of matches between PREVIOUS and CURRENT was then determined separately for each 20-tile.

[11] I am indebted to Stefan Th. Gries and Anatol Stefanowitsch for giving me access to the complete list. This is how the *p* values of which their list consists of were transformed mathematically into a 0−100 scale: *p* values of verbs that Gries and Stefanowitsch found to have a preference for *V* + *Particle* + *Object* were subtracted from + 2; *p* values of verbs that have a preference for *V* + *Object* + *Particle* were multiplied by −1. Subsequently, all values were multiplied by + 50. Thus, low values close to 0 indicate that the verb under analysis has a preference for the *V* + *Particle* + *Object* pattern, and high values close to 100 indicate that the verb under analysis has a preference for the *V* + *Object* + *Particle* pattern.

[12] Mathematically, if VLEMMAID is 1, PREVIOUS is associated with an odds ratio of 0.17 × 0.39 = 0.066 instead of 'only' 0.17.

[13] Due to exceedingly low frequencies and its marginal status in present-day spoken English (cf. Kjellmer 1998: 155−186; Tottie 2002: 37−58; Trudgill 1984: 32−44), *shall* (and *shan't*) have not been considered in the present study.

[14] For instance, it has been claimed that using one or the other option "has a scarcely perceptible effect on meaning" (Quirk et al. 1985: 218), which is why "it is difficult to discover any simple sentences in which either *will* yields a clearly definable sense which *going to* does not" (Hall and Hall 1970: 138−139). Similarly, Danchev et al. (1965: 375−386) argue for overall synonymy, and Palmer (1974: 163) asserts that "in most cases, there is no demonstrable difference between *will* and *be going to*". Haegemann (1989: 291−317) has argued that whatever the difference is between BE GOING TO and WILL, it must be pragmatic rather than truth-conditionally semantic.

[15] The reason why SAMETURN and SAMESPEAKER were not included in the regression on comparison strategy choice is that adjectives which can take both types of comparison are relatively rare (average textual distance between two com-

parison choice contexts in the BNC-CG is over 2,000 words). Virtually categorically, SAMETURN would thus have been 0. As for particle placement, the nature of the data source analyzed (FRED) does not lend itself for analysis of SAMETURN and SAMESPEAKER: an interviewer usually asks brief questions, and the interviewee responds in a quite monologic way. Both variables, therefore, would have been practically always 1.

16. It is for reasons of space that, rather eclectically, SENTENCELENGTH is used as an explanatory variable in particle placement and TTR in future marker choice. The results would not be dramatically different if it were the other way round, or if both variables were included in both models.

References

Abbi, Anvita
1985 Reduplicative structures: A phenomenon of the South Asian linguistic area. In Acson, Veneeta, and Richard Leed (eds.), *For Gordon H. Fairbanks (Oceanic Linguistics, Special publication 20)*. Honolulu: University of Hawaii Press, 159–171.

Aston, Guy, and Lou Burnard
1998 The BNC Handbook: Exploring the British National Corpus with Sara. Edinburgh: Edinburgh University Press.

Bauer, Laurie
1994 *Watching English change: an introduction to the study of linguistic change in standard Englishes in the twentieth century*. London: Longman.

Behaghel, Otto
1909/10 Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25.

Berglund, Ylva
1999 Utilising Present-day English corpora: A case study concerning expressions of future. *ICAME Journal* 24, 25–63.
2000 "You're gonna, you're not going to": A corpus-based study of colligation and collocation patterns of the *BE going to* construction in Present-day spoken British English. In Lewandowska-Tomaszczyk, Barbara and Patrick James Melia (eds.), *PALC'99: Practical applications in language corpora: papers from the 2. international conference at the University of Lódz, 15–18 April 1999*. Frankfurt a. M. Berlin: Peter Lang.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan
1999 *Longman grammar of spoken and written English*. Harlow: Longman.

Bock, Kathryn
1986 Syntactic Persistence in language production. *Cognitive Psychology* 18, 355–387.

Bock, Kathryn, and Helga Loebell
1990 Framing Sentences. *Cognition* 35, 1–39.

Bock, Kathryn, and Zenzi Griffin
2000 The Persistence of Structural Priming: Transient Activation or Implicit Learning? *Journal of Experimental Psychology: General* 129, 177–192.

Bolinger, Dwight
1961 Syntactic blends and other matters. *Language* 37, 366–381.
1968 *Aspects of Language*. New York: Harcourt.
1971 *The Phrasal Verb in English*. Cambridge, MA: Harvard University Press.

Bortfeld, Heather, Silvia Leon, Jonathan Bloom, Michael Schober, and Susan Brennan

2001 Disfluency rates in conversation: effects of age, relationship, topic, role and gender. *Language and Speech* 44, 123−147.

Branigan, Holly, Martin Pickering, and Alexandra Cleland

1999 Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin and Review* 6, 635−640.

2000 Syntactic Coordination in Dialogue. *Cognition* 75, 813−825.

Branigan, Holly, Martin Pickering, Simon Liversedge, Andrew Stewart, and Thomas Urbach

1995 Syntactic Priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research* 24, 489−506.

Braun, Albert

1982 Studien zur Syntax und Morphologie der Steigerungsformen im Englischen. *Schweizer Anglistische Arbeiten*, volume 110. Bern: Francke.

Cohen, Laurent, and Stanislas Dehaene

1998 Competition between past and present: Assessment and interpretation of verbal perseverations. *Brain* 121, 1641−1659.

Danchev, A., A. Pavlova, M. Nalchadjan, and O. Zlatareva

1965 The Construction *going to + inf.* in Modern English. *Zeitschrift für Anglistik und Amerikanistik* 13, 375−386.

Estival, Dominique

1985 Syntactic Priming of the Passive in English. *Text* 5, 7−21.

Fraser, Bruce

1965 *An examination of the verb-particle construction in English.* Massachusetts Institute of Technology: Unpublished PhD-Thesis.

1966 Some remarks on the verb-particle construction in English. In Dinnen, Francis (ed.), *Problems in Semantics, History of Linguistics, Linguistics and English.* Washington D.C.: Georgetown University Press.

1974 Review of Dwight Bolinger, "The Phrasal Verb in English". *Language* 50, 568−575.

Gries, Stefan Th.

2003a Grammatical variation in English: A question of 'structure vs. function'? In Rohdenburg, Günter, and Britta Mondorf (eds.), *Determinants of Grammatical Variation in English.* Berlin: Mouton de Gruyter, 155−174.

2003b. *Multifactorial Analysis in Corpus Linguistics: A study of particle placement.* Continuum.

forthc. Syntactic priming: A corpus-based approach.

Gries, Stefan Th., and Anatol Stefanowitsch

2004 Extending collustructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9, 97−129.

Haegeman, Liliane

1989 *Be going to* and *will*: A Pragmatic Account. *Journal of Linguistics* 25, 291−317.

Hall, R., and Beatrice Hall

1970 A Note on *will* vs. *going to*. *Linguistic Inquiry* 1, 138−139.

Hartsuiker, Robert, and Casper Westenberg

2000 Word order priming in written and spoken sentence production. *Cognition* 75, B27-B39.

Hopper, Paul

1998 Emergent Grammar. In Tomasello, M. (ed.), *The new psychology of language: cognitive and functional approaches to language structure.* Mahwah, NJ: Lawrence Erlbaum Associates, 155−176.

Kempley, S. T., and John Morton
1982 The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology* 73, 441–445.

Kjellmer, Goran
1998 On Contraction in Modern English. *Studia Neophilologica* 69, 155–186.

Kortmann, Bernd
2002 New prospects for the study of dialect syntax: Impetus from syntactic theory and language typology. Barbiers, Sjef, Leonie Cornips, and Susanne van der Kleij (eds.), *Syntactic Microvariation.* Amsterdam: Meertens Instituut, 185–213.

Kuryłowicz, Jerzy
1964 *The Inflectional Categories of Indo-European*. Heidelberg: C. Winter.

Langacker, Ronald
1987 *Foundations of Cognitive Grammar, Volume 1: Theoretical prerequisites.* Stanford: Stanford University Press.

Leech, Geoffrey, and Jonathan Culpeper
1997 The Comparison of Adjectives in Recent British English. In Nevalainen, Terttu, and Kahlas-Tarkka (eds.), *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*. Amsterdam: Rodopi, 125–132.

Levelt, Willem and Stephanie Kelter
1982 Surface Form and Memory in Question Answering. *Cognitive Psychology* 14, 78–106.

Lindquist, Hans
2000 *Livelier* or *more lively*: syntactic and contextual factors influencing the comparison of disyllabic adjectives. *Papers from the Nineteenth International Conference on English Language Research on Computerized Corpora, ICAME 1998*.

Matthews, R. J.
1979 Are the grammatical sentences of a language a recursive set? *Synthese* 40, 209–224.

McKone, Elinor
1995 Short-term implicit memory for words and non-words. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21, 1108–1126.

Meyer, David and Roger Schvaneveldt
1971 Facilitation in recognizing pairs of words: Evidence of dependence between retrieval operations. *Journal of Experimental Psychology* 90, 227–234.

Mondorf, Britta
2003 Support for *more*-support. In Rohdenburg, Günter, and Mondorf, Britta (eds.), *Determinants of Grammatical Variation in English*. Berlin, New York: Mouton de Gruyter, 251–304.

Orwin, Robert
1994 Evaluating coding Decisions. In Cooper, Harris, and Larry Hedges (eds.), *The Handbook of Research Synthesis*. New York: Russel Sage Foundation, 139–162.

Palmer, Frank
1974 *The English verb*. London: Longman.

Pickering, Martin, and Holly Branigan
1999 Syntactic priming in language production. *Trends in Cognitive Science* 3, 136–141.

Poplack, Shana
1980 The notion of the plural in Puerto Rican English: Competing constraints on (s) deletion. In Labow, William (ed.), *Locating Language in Time and Space*. New York: Academic Press, 55–67.

Poplack, Shana, and Sali Tagliamonte

1993 The zero-marked verb: Testing the creole hypothesis. *Journal of Pidgin and Creole Languages* 8, 171–206.

1996 Nothing in context: Variation, grammaticization and past time marking in Nigerian Pidgin English. Changing meanings, changing functions. In Baker, Philip, and Anand Syea (eds.), *Papers relating to grammaticalization in contact language*. London: University of Westminster Press, 71–94.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik

1985 *A Comprehensive Grammar of the English Language*. London, New York: Longman.

Sankoff, David, and Suzanne Laberge

1978 Statistical Dependence among Successive Occurrences of a Variable in Discourse. In Sankoff, David (ed.), *Linguistic Variation: Models and Methods.* New York: Academic Press, 119–126.

Scherre, Maria, and Anthony Naro

1991 Marking in discourse: "Birds of a feather". *Language Variation and Change* 3, 23–32.

Szmrecsanyi, Benedikt

2003 *Be going to* versus *will/shall*: Does Syntax matter? *Journal of English Linguistics* 31, 295–323.

2004 On Operationalizing Syntactic Complexity. In Purnelle, Gérard, Cédrick Fairon, and Anne Dister (eds.), *Le poids des mots. Proceedings of the 7th International Conference on Textual Data Statistical Analysis. Louvain-la-Neuve, March 10–12, 2004*. Louvain-la-Neuve: Presses universitaires de Louvain, 1032–39.

forthc. *Persistence Phenomena in the Grammar of Spoken English.*

Tanenhaus, Michael, H. P. Flanigan, and Mark Seidenberg

1980 Orthographic and phonological activation in auditory and visual word recognition. *Memory and Cognition* 8, 513–520.

Tannen, Deborah

1982 Oral and literate strategies in spoken and written narratives. *Language* 58, 1–21.

1987 Repetition in conversation: Toward a poetics of talk. *Language* 63, 574–605.

1989 *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

Tottie, Gunnel

2002 Non-Categorical Differences between American and British English: some Corpus Evidence. In Modiano, Marko (ed.), *Studies in mid-atlantic English.* Gävle: University of Gävle Press, 37–58.

Trudgill, Peter

1984 Standard English in England. In Trudgill, Peter (ed.), *Language in the British Isles.* Cambridge: Cambridge University Press, 32–44.

Weiner, Judith and Labov, William

1983 Constraints on the agentless passive. *Journal of Linguistics* 19, 29–58.

Zurif, Edgar, David Swinney, Penny Prather, Arthur Wingfield, and Hirma Brownell

1995 The allocation of memory resources during sentence comprehension: Evidence from the elderly. *Journal of Psycholinguistic Research* 24, 165–182.