

Studia Anglica Posnaniensia 45/2, 2009
doi: 10.2478/v10121-009-0019-0

THE INTERACTION OF FORM AND CONTENT:
SYNTACTIC CONSTRUCTIONS AND
GRAMMATICAL ENVIRONMENT

FRANČIŠKA TROBEVŠEK DROBNAK

University of Ljubljana, Slovenia

ABSTRACT

In diachronic linguistics, the Naturalness Theory purports to unravel the seemingly random distribution of linguistic variants at the early stage of their assertion, when no other tangible functional, contextual or pragmatic motivation exists. The paper presents the results of three empirical studies, which confirm that the complexity of grammatical environment is instrumental in the choice between elective morphosyntactic constructions. Special attention is paid to the relevance of traditional grammatical parameters in defining grammatical environment as complex.

1. Introduction

The subsequent paper is a review and a follow-up of a number of empirical studies (Trobevšek, 1990, 1994, 1999) into the use of selected pairs of morphosyntactic constructions in Old English and in present-day English texts, re-examining the parameters of their interaction with the immediate grammatical environment. The basic postulates of the research derive from the Theory of Naturalness in linguistics, which posits that, in addition to functional and pragmatic motivation, the assertion and subsequent expansion of linguistic variants to specific grammatical environment depends on their outer form. The focus of this paper is the investigation into which, if any, properties of the immediate grammatical environment correlate with the formal elaborateness or simplicity of linguistic variants, excluding from the equation, to the highest degree possible, the functional and pragmatic motivation for their actual use, as well as the already verified correlation between the form and content of linguistic structures themselves.

The pioneers of the Naturalness Theory were the founders of Natural Phonology (Stampe 1979; Donegan 1985) and Natural Morphology (Mayerthaler 1981; Dressler 1987). In syntax, it was applied by Rydén (1979) and, as a language-universal theory, by Dotter (1990) and Dressler (2000). The behaviour of morphosyntactic pairs of constructions in English and German language material was first empirically tested at the University of Ljubljana, Slovenia (Orešnik *et al.* 1990).

On synchronic level, the theory postulates that of two morphosyntactic variants one is formally more elaborate and the other one less elaborate. While the initial work applied the terms “stronger” and “weaker” (variants), Mayerthaler (1981, 1987) and Orešnik (2000, 2007) currently operate with the terms “naturalness scale”, “sem-values” and “sym-values”. The naturalness scale rests on the assumption that of two morphosyntactic variants one is more natural (<nat) from the speaker’s point of view, who is considered the focal point of communication (Orešnik 2007: 293). A typical <nat construction is formally less elaborate, bending to the principle of least effort (Havers 1931: 171). As early as in 1981, Mayerthaler (1981) divided the naturalness scale into the one referring to the symbolic (formal) properties (sym-values), and the one referring to the semantic properties (sem-values) of linguistic constructions. While the natural affinity for more economical linguistic variants can be explained in terms of fundamental Gricean pragmatics (Grice 1975), the naturalness theory aims to determine the role of “grammatical environment” in the speaker’s assessment of the “minimum [output of energy, both mental and physical,] compatible with achieving his ends” (Martinet 1962: 139).

On diachronic level, the initial hypothesis was that, *post status nascendi*, stronger variants are favoured in a relatively complex grammatical environment, and weaker variants in a relatively simple grammatical environment (Orešnik *et al.* 1990: 5-11). The author of this paper empirically tested the hypothesis by investigating the correlation between the complexity of grammatical environment and the assertion of expanded tenses in Old English, the loss of the preverbal *ge-* in Old English, and the distribution of participial non-finite clauses in Old English and in present-day English. The choice was based on the following two criteria:

- a) Each selected construction functions, or functioned at a given point in its history, as a (morpho)syntactic variant of another construction.
- b) In each pair of thus selected constructions, one is formally more elaborate, more transparent and easier to decode. In this paper, it is referred to as the <form construction. The other member of the pair is formally less elaborate, more economical and more opaque to decode. In this paper it is

defined as the >form construction. In contrast to Mayerthaler's <sym and <sem values, the terms <form and >form are assigned without any reference to their position within the naturalness scale.

2. The method

For each pair of constructions, basic and control samples were compiled from Old English and/or present-day English texts. The basic samples illustrate the grammatical environment of the constructions under observation, and the control samples the grammatical environment of their respective linguistic variants

The grammatical environment of selected constructions was determined as to its scope (e.g. sentence, clause, matrix verbal phrase) and as to the observables. Initially, these were traditional grammatical categories, which were assigned binary values, one defined as constituting a more complex (<com), and the other as constituting a less complex (>com) grammatical environment. In the absence of other reliable criteria, the attribute <com was assigned to the marked values of grammatical categories, as proposed by the Prague School (Jakobson 1932) and by natural morphologists (Mayerthaler 1981). Unlike Mayerthaler's sem-values, with which they share the common markedness grounds, the <com and >com labels are used for the assessment of the complexity of grammatical environment of chosen constructions, and not for the evaluation of the (cognitive or grammatical) complexity of constructions themselves.

For each of the observables, the probability rates were computed for parameters to assume the <com value in the environment of the <form construction (P_a) and of the >form construction (P_b) respectively. The equation applied was (Pavlić 1985)¹

$$P(a, b) = \frac{n(a, b)}{n}$$

The probability rate of favourable events in the basic samples was compared with the probability rate of favourable events in the control samples. The index which shows the statistical significance of the difference between the two probability rates was computed:²

$$Id = \frac{P(a) - P(b)}{sd}$$

¹ a,b = favourable event, the presence of the <com value of the parameter;

² The difference between two values is considered statistically significant if I(d) exceeds 2 (Pavlić 1985).

with

$$sd = \sqrt{\frac{P \times q \times (n(a) - n(b))}{n(a) \times n(b)}}$$

The initial hypothesis that <form constructions are favoured in <com grammatical environment, and that >form constructions are favoured in >com grammatical environment, was presumed validated if the probability of favourable events in the basic sample was found significantly higher than the probability of favourable events in the control sample.

3. Expanded tenses in the Old English *Orosius*³

The first test of the validity of the hypothesis was the analysis of the use of expanded tenses in the Old English translation of *Historiarum Adversus Paganos* by the historian and theologian Orosius. The results of the research were published in Trobevšek (1990).

The Old English construction consisting of the verb *beon/wesan* + present participle *-ende* can be found in the oldest Old English manuscripts. While there are only four instances in *Beowulf* (Klaeber 1950), it is relatively frequent in texts translated from Latin. Mossé (1938: 53-54) believed that the introduction of this construction was due to the influence of Latin, and that it was motivated by the disappearance of the old Germanic system of the verbal aspect. Nickel, however, was sceptical of any such correlation (1966: 83-207).

In *Orosius*, dating probably from about CE 890, expanded tenses are relatively frequent and independent from the Latin constructions: of 237 instances in the Old English text, 131 correspond to simple verbal tenses or have no equivalent in the Latin original (Mossé 1938: 66; Nickel 1966: 112), and 154 are rendered as simple tenses in either or both of the Modern English translations (Barrington 1773; Thorpe 1873, as quoted in Nickel 1966: 330-351). Given this electiveness of expanded and nonexpanded tenses in *Orosius*, it is acceptable to treat them as syntactic variants, the former as the <form and the latter as the >form construction.

The basic sample consisted of 237 clauses with the verb in one of the expanded tenses. The control sample consisted of 855 clauses with the verb in a simple tense (present or preterite). Every third such clause was taken from Sweet's edition of *Orosius* (1883, reprinted in 1956). The probability rates of

³ The term "expanded tenses" used in previous studies and in this paper is taken from Nickel (1966). It refers to the Old English periphrasis involving *beon/wesan* + present participle in *-ende*. Unlike the terms progressive/continuous tenses, it makes no allusions to its potential functions.

any chosen grammatical category to assume a particular value was computed for the basic and for the control sample, following the method described above. Subordinate clauses, negative propositional modality, preterite tense, non-indicative (subjunctive or imperative) mood, imperfective aspect, plural number and transitivity of the verb were presumed as <com grammatical environment. Additionally, the type of the complement to the transitive verb was observed. The prepositional phrase was assumed to constitute <com environment.

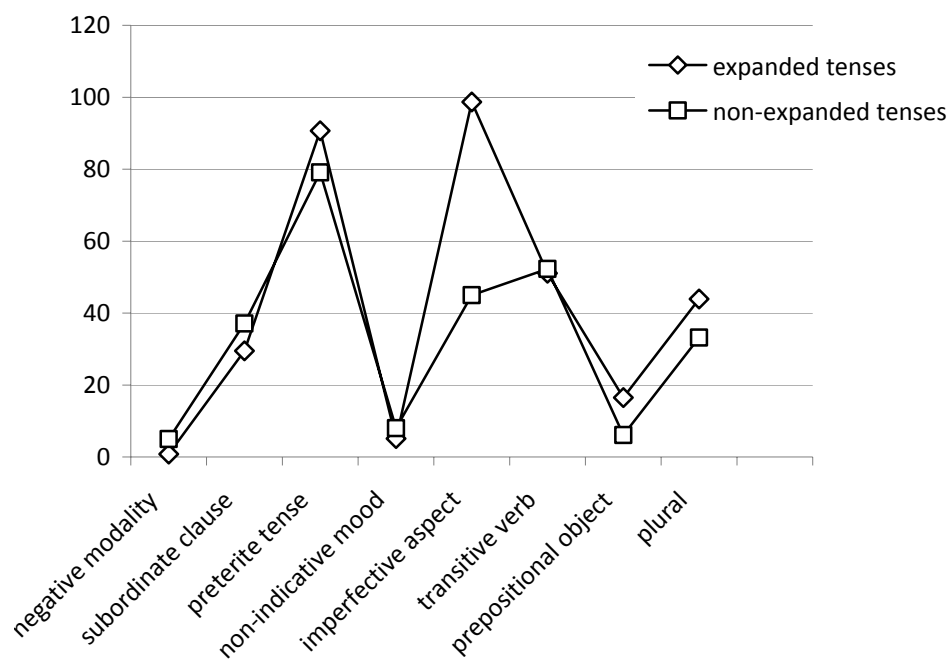


Figure 1. Probability of <com values of parameters in the environment of expanded and non-expanded tenses in Old English

Four out of eight grammatical parameters (the tense, the aspect, the type of object and the number) assumed the <com value in the environment of the <form construction, while four parameters (the propositional modality, the type of clause, the mood, the transitivity) assumed the <com value in the environment of >form construction. From the point of view of the hypothesis stated in section 1, the predictions were confirmed in the case of the tense, the aspect, the type of verbal complementation (object) and the number. The propensity of <form for <com grammatical environment is clearly indicated, albeit not consistently so.⁴

⁴ The validity of the results for the propositional modality and the mood is reduced by the low

4. The preverbal *ge-* in *Orosius* and in Old English gospels

The second empirical test involved the study of the grammatical environment of Old English verbs with the prefix *ge-* in comparison with the grammatical environment of Old English simplex verbs (Trobevšek 1994).

The Old English preverbal *ge-* has traditionally been related to the Gothic *ga-*, which is in turn explained as corresponding to Latin *cum-/com-* (< Indo-European **kom-*). Most linguists, like Grimm (1878: 829), Mossé (1938: 1-13) and Samuels (1949: 81-90), agree that the original meaning of the Gothic prefix, 'zusammen', had faded out in Old English, and that its function was to imply completion, pluperfect or the future perfect. Lindemann (1970: 19-25) examined 45,000 simplex and prefixed verbs and concluded that the OE *ge-* seems to "do something" to the information conveyed by the verb, but the exact nature of this "additional information" depends on the context.

The gradual withdrawal of the preverbal *ge-* from Old English is generally ascribed to the collapse of the Germanic aspectual system, which was eventually to be replaced by expanded tenses in modern English (Mossé 1938: 42). There seems to be no contention with the fact that, at least in the late Old English period, the preverbal *ge-* was weakened to the point that its use was no longer tangibly contrastive with simplex verbs (Trobevšek 1994: 126). Consequently, it can be treated as the <form linguistic variant of corresponding simplex verbs.

The research samples were taken from the Old English *Orosius* and from the Old English translation of the *Gospels*. The first basic sample consisted of all 820 occurrences of *ge-*verbs in Sweet's (1883 [1959]) edition of *Orosius* (past participles excluded), and the first control sample was made of 1000 simplex verbs in the same edition (every other simplex verb was taken, with the exception of modal verbs and the verbs *beon/wesan*, *habban* and *weorðan* when used with the past participle). The sampling covered 294 of 298 pages of the edition. The second set of samples was taken from Skeat's (1871) edition of the Old English translations of the *Gospel* according to St Mark and of the *Gospel* according to St John (1878). The basic sample consisted of all the occurrences (377) of *ge-*verbs in the two *Gospels*, the control sample consisted of all simplex verbs (565) in the same two texts.

The parameters chosen as the indicators of the complexity of grammatical environment were the same as in study of expanded tenses, with the following minor modifications:

occurrence of the negative modality and the subjunctive mood in both samples (below 10 percent). The absolute prevalence of the imperfective aspect in the basic sample could be semantically/grammatically motivated and also has to be interpreted with caution.

- The parameter “type of clause” was assigned the values: independent (simple) sentence, main clause with subordinate clause(s), subordinate clause.
- The parameter “aspect” was excluded since the use of *ge*-verbs could be grammatically motivated.
- The parameter “person” was added.
- The parameter “number” was observed within the third person only.

The probability of each grammatical parameter to assume a given value was computed for both samples. Subordinate clauses, preterite tense, transitivity, non-indicative mood, non-third person and plural number were provisionally assigned the value <com. Their frequencies are given in Figure 2:

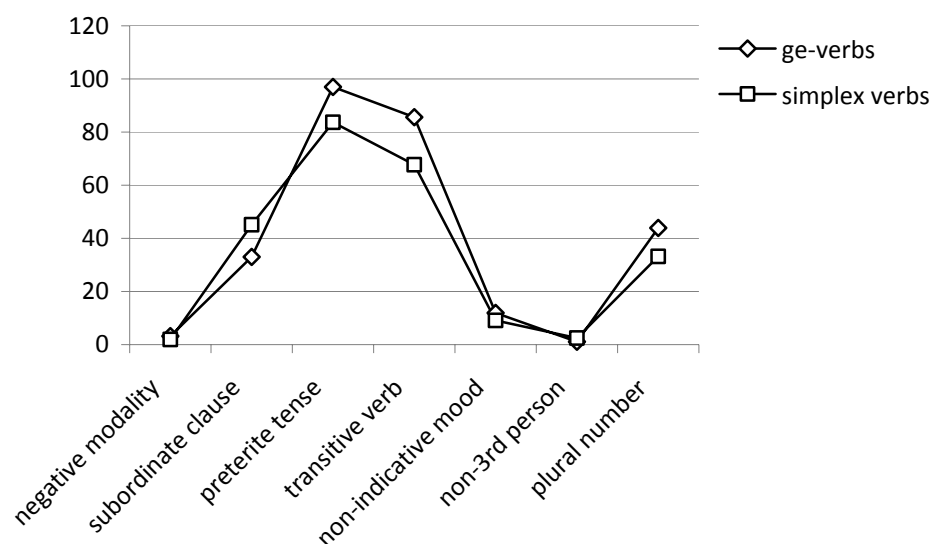


Figure 2: Probability of <com values of grammatical parameters in the environment of *ge*-verbs and of simplex verbs in *Orosius*

In the samples taken from Skeat’s *Gospels*, the probability rates of individual values of parameters in the basic and in the control sample were computed for the same parameters. The probability rates of their <com values are given in Figure 3:

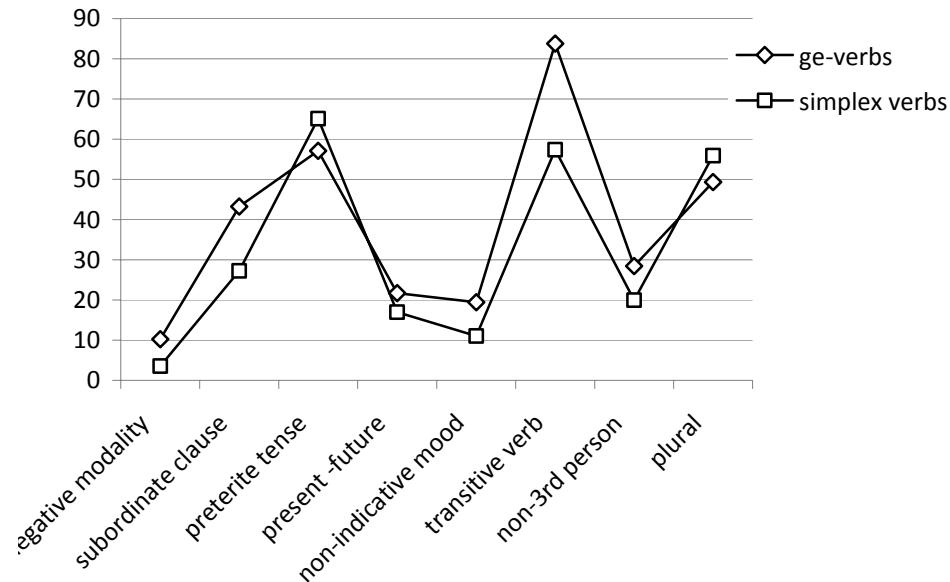


Figure 3: Probability of <com values of grammatical parameters in the environment of *ge*-verbs and of simplex verbs in Old English *Gospels*

The comparison of the results of the analysis of both sets of samples reveals consistency in the case of negative modality, transitivity and non-indicative mood. In both sets of samples, the prediction that *ge*-verbs would persist longer in <com environment was confirmed for these three parameters. The results differ in the case of the type of clause, tense, person and number. Further research was clearly suggested, with more caution applied to: (a) defining constructions as <form and >form syntactic variants; (b) pairing off different values of grammatical parameters as representing simple (>com) or complex (<com) grammatical environment.

5. Participial non-finite clauses

The one contention against defining expanded tenses and *ge*-verbs as the <form variants of non-expanded tenses and simplex verbs respectively, could be the possible functional motivation (i.e. aspectual information) for their use. No such motivation can be found in the choice between a participial non-finite clause and a corresponding adjectival or adverbial finite subordinate clause.

The early stage of the assertion of participial non-finite clauses was in Old English (Kisbye 1971: 24-27), possibly under the influence of Latin.⁵ They are still interchangeable with adverbial and adjectival subordinate clauses in present-day English, which qualifies the two constructions as syntactic variants. The participial non-finite clauses (PNF) can be assigned the >form value, and the corresponding adjectival and adverbial finite subordinate clauses (FSC) <form value. Their respective partiality for specific grammatical environment can be empirically tested.

The list of parameters as the indicators of the complexity of grammatical environment was revised. The parameter “type of clause” was removed, since one of the variant constructions observed was to be a finite subordinate clause. Three parameters were added: the [+/-animate] status of the subject of the matrix verb, the [+/-stative] status of the matrix verb, and the syntactic function of PNF or FSC in the superordinate linguistic construction. All these additional parameters appear relevant in the process of creolization, which is commonly recognized as manifesting the interdependence of the form and of the content of communication (Todd 1974; Bickerton 1981).⁶ The parameter transitivity was assigned three values: [-transitive], [+transitive] and, as a sub-set of [+transitive], the value [+ditransitive] for verbs with three complements (the subject, the direct and the indirect object). The parameter tense was combined with the “stativity” of the verb. All parameters were observed in the “environment” of the two constructions, so that the computed values pertain to the matrix verb (in the case of PNF) or to the verb in the main clause (in the case of FSC).

Two sets of basic and control samples were formed, one taken from an Old English text, and one from a present-day English text. The first set was taken from Skeat’s (1871) edition of the Old English *Gospel* according to St Mark. The basic sample consisted of all 114 clauses containing a non-finite clause with the present participle, the control sample consisted of 252 clauses to which adjectival or adverbial clauses were subordinated.⁷ The second set of samples was taken from *Murder in the Calais Coach* by Agatha Christie (1934). The basic sample consisted of 250 instances of PNF, the control sample of 300 instances of main clauses with adjectival or adverbial subordinate clauses.

⁵ Only non-finite clauses with the present participle were included in the study.

⁶ When the number encodement, which has no formal expression in pidgin languages, becomes mandatory in creoles, the morphologization favours *human* > *animates* > *count nouns* > *mass nouns*, and also *subject* > *direct object* > *indirect object* > *locative* > *genitive*. The primary tense system, based on [+/-anterior] is dependent on the stative or non-stative status of the verb. Bickerton reports that the default tense of the zero form of stative verbs is [-anterior], and the default tense of non-stative verbs is [+anterior] (Bickerton 1975: 461).

⁷ Only adjectival and adverbial clauses which could be replaced with corresponding non-finite clauses with the present participle were chosen.

The hypothesis was (re)formulated as follows: (a) The values of grammatical parameters in the environment of PNF will differ from the values of grammatical parameters in the environment of FSC. (b) PNF, as the >com variant, will be favoured in >com grammatical environment, and resisted in <com grammatical environment, which will still favour the <form of FSC. (c) The contrast between the complexity of the grammatical environment of PNF and FSC will be more pronounced in the Old English than in the present-day English samples.

The grammatical environment of PNF and FSC was first analysed in the Old English samples. The frequencies of <com values of grammatical parameters are given Figure 4:

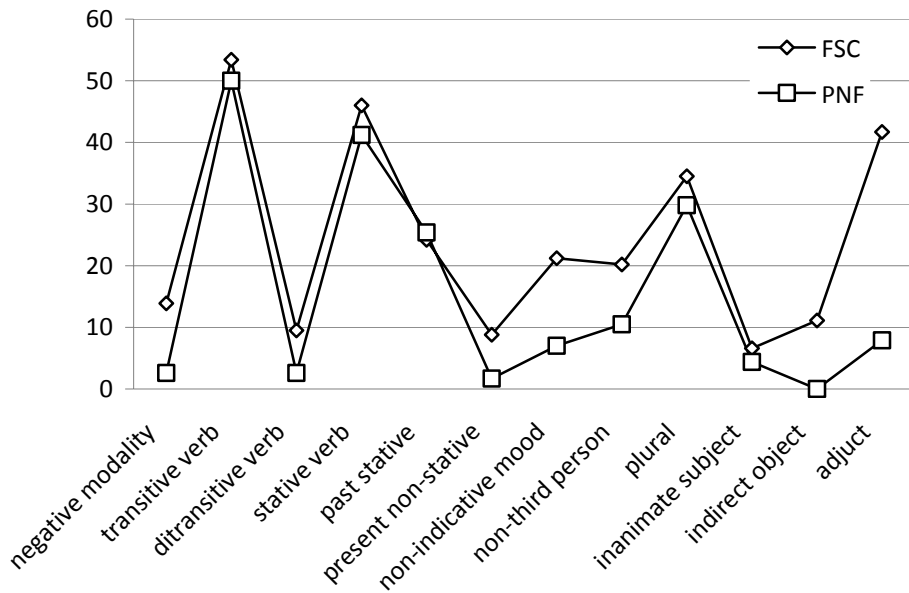


Figure 4: Probability of <com values of grammatical parameters in the environment of PNF and FSC in Old English *Gospels*

The grammatical environment of NPF and FSC in present-day English is presented in Figure 5:

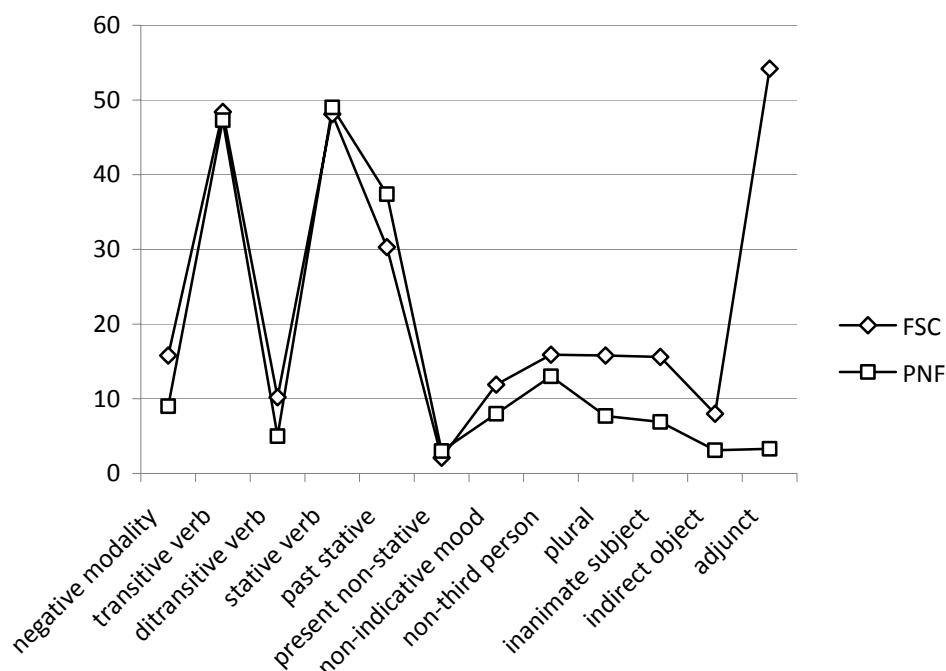


Figure 5: Probability of <com values of grammatical parameters in the environment of FSC and PNF in present day English

The results reveal that the following grammatical environment resisted the spread of PNF in Old English: negative modality, transitive and especially ditransitive (matrix) verbs stative (matrix) verbs, non-stative (matrix) verbs in the present tense, non-indicative mood, non-third person(s), plural subject of the (matrix) verb, inanimate subject, PNF performing the function of the indirect object or the adjunct in the superordinate structure. In contrast to Old English samples, PNF has become the preferred choice in the environment of stative verbs and in the present tense of non-stative verbs in present-day English samples. The differences in the probability rates of parameter values in the environment of both syntactic variants are less pronounced in the present-day English samples than in the Old English samples, with the exception of the parameters present stative verbs, past stative verbs, (in)animate subject and the function of adjunct.

6. Conclusion

Linguistic innovation and the early stage of its proliferation are usually studied from the point of view of the (multiple) meanings and functions of the new lin-

guistic entity in different linguistic material and context. Initially, the innovation frequently behaves as a linguistic variant of another (older) construction, without any consistent semantic and/or grammatical contrast between them. Since the Naturalness Theory explores the correlation between the formal properties and the complexity of linguistic variants, as well as of their respective grammatical environment, its postulates can be applied in the analysis of the (statistical) bias of the speakers to choose between two linguistic variants.

The most vulnerable aspect of the research is that it rests, among other postulates, on the assumption that certain values of grammatical environment are more complex than others. The decision to start out with the marked values of parameters as proposed by the Prague School and by Natural Morphologists was logical in the absence of any other reliable criteria. The simultaneous verification of the theory and of its apparatus requires a deductive empirical approach, so as to minimize the possibility of verifying the hypothesis with selective data. It took a number of similar studies to discover which parameter values display consistency, so that they can be used as stable yardsticks in determining the veracity of the initial prediction that <com grammatical environment favours <form constructions and >com environment favours >form constructions.

Of all the studies, the one exploring the assertion and distribution on participial non-finite clauses yielded the most consistent results. The predictions formulated above proved valid:

- Of 17 parameter values observed in the grammatical environment of participial non-finite clauses and finite subordinate clauses, 13 differ significantly ($p < 0.05$) both in the Old English and in the present-day English samples.
- In the Old English samples, participial non-finite clauses, the less elaborate of the two constructions, are preferred in the environment of affirmative propositional modality, intransitive verbs, non-stative verbs, stative verbs with present time reference, non-stative verbs with past time reference, indicative mood, third person, singular subject, animate subject, third persons, singular number, as the subject of the matrix verb, as the direct object of the matrix verb. All of these values constitute less complex grammatical environment, both according to the markedness criteria of the Prague School and the sem-values of Natural Morphologist. The results are also consistent with the principles that govern the process of morphologization in Creole languages. In the more complex grammatical environment, at contrasting marked values of parameters, participial non-finite clauses were disfavoured and finite subordinate clauses are the preferred choice.
- The differences in the probability rates of parameter values in the environment of both syntactic variants are less pronounced in the present-day English samples than in the Old English samples for 13 of 17 parameters.

Of all the studies, the one exploring the assertion and distribution of participial non-finite clauses yielded the most consistent results. They prove that, in addition to the correlation between the form and the content of linguistic constructions themselves, the nature of the immediate grammatical environment affects the choice between formally contrastive linguistic variants. This influence is more pronounced at the early stage of the proliferation of linguistic innovation, before other factors, such as functional motivation, become predominant. When determining the complexity of grammatical environment, the traditional marked values should be considered in combination with the findings of other branches of linguistics, e. g. language typology and Creole studies.

REFERENCES

PRIMARY SOURCES

Christie, Agatha

1934 *Murder in the Calais coach*. New York: Dodd Mead & Co.

Skeat Walter W. (ed.)

1871 *The Gospel according to Saint Mark. In Anglo-Saxon and Northumbrian versions synoptically arranged, with collations exhibiting all the readings in all the MSS.* Cambridge: Cambridge University Press.

1878 *The Gospel according to Saint John In Anglo-Saxon and Northumbrian versions synoptically arranged, with collations exhibiting all the readings in all the MSS.* Cambridge: Cambridge University Press.

Sweet, Henry (ed.)

1883

[1959] *King Alfred's Orosius*. (EETS 79.) London: Trübner.

SECONDARY SOURCES

Bickerton, Derek

1975 *Dynamics of Creole system*. Cambridge: Cambridge University Press.

Booij, Geert – Christian Lehman – Joachim Mugdan (eds.).

2000 *Morphologie: Ein internationales Handbuch zur Flexion und Wortbildung*. Berlin: Gruyter.

Cole, Peter – Jerry Morgan (eds.)

1975 *Syntax and semantics 3: Speech acts*. New York: Academic Press.

Donegan, Patricia Jane

1985 *On the natural phonology of vowels*. New York – London: Garland.

Dotter, Franz

1990 *Nichtarbitrarität und Ikonizität in der Syntax*. Hamburg: Buske.

Dressler, Wolfgang U. – Willi Mayerthaler – Oskar Panagl – Wurzel Wolfgang U. (eds.)

1988 *Leitmotifs in Natural Morphology*. Amsterdam: Benjamins.

- Dressler, Wolfgang U.
2000 "Naturalness", in: Geert Booij – Christian Lehman – Joachim Mugdan (eds), 288-296.
- Grice, Herbert P.
1975 "Logic and conversation", in: Peter Cole – Jerry Morgan (eds.), 41-58.
- Grimm, Jakob
1878 *Deutsche Grammatik*. Göttingen: Dieterichsche Buchhandlung.
- Havers, Wilhelm
1931 *Handbuch der erklärenden Syntax. Ein Versuch zur Erforschung der Bedingungen und Triebkräfte in Sytax und Stilistik*. Heilderberg: Winter.
- Jakobson, Roman
1932 "Zur Struktur des russischen Verbums", in: *Charisteria Guilelmo Mathesio Quinquagenario: A Discioulis et Circuli Linguistici pragensis Sodalibus oblate* (Cercle Linguistic de Prague.): 74-86.
- Klaeber, Friedrich
1950 *Beowulf and the fight at Finnsburg*. (3rd edition.) Boston: D. C. Heat.
- Kysbie, Torben
1971 *An historical outline of English syntax*. Aarhus: Akademisk Boghandel.
- Lindemann, John W. R.
1970 *Old English preverbal ge-: Its meaning*. Charlottesville: The University Press of Virginia.
- Martinet, André
1962 *A functional view of Language*. Oxford: Oxford University Press.
- Mayerthaler, Willi
1981 *Morphologische Natürlichkeit*. Wiesbaden: Athenaion.
1988 "System-independent morphological naturalness", in: Wolfgang U. Dressler – Willi Mayerthaler – Oskar Panagl – Wolfgang U. Wurzel (eds.), 25-57.
- Mossé, Ferdinand
1938 *Histoire de la forme périphrastique être + participe présent en germanique*. Paris: University of Paris.
- Nickel, Gerhard
1966 *Die Expanded Form im Altenenglischen. Vorkommen, Funktion und Herkunft beon/wesan + Partizip praesens*. Neumünster: Wachholtz.
- Orešnik, Janez – Andrej Snedec – Karmen Teržan – Frančiška Trobevšek Drobnak
1990 "Introduction to the subsequent three papers in the present volume", *Linguistica* 30: 5-12.
- Orešnik, Janez
2004 *Naturalness in (morpho)syntax. English examples*. Ljubljana: Slovenska akademija znanosti in umetnosti.
- Pavlič, Ivan
1985 *Statistička teorija i primjena*. Zagreb: Tehnička knjiga.
- Ryden, Mats
1979 *An introduction to the historical study of English syntax*. Stockholm: Almqvist & Wiksell.
- Samuels, Michael L.
1950 "The ge- prefix in the Old English gloss to the *Lindisfarne Gospel*", *Transactions of the Philological Society* 1949: 62-116.

Stampe, David

1979 *A dissertation on Natural Phonology*. New York – London: Garland.

Todd, Loreto

1974 *Pidgins and Creoles*. London: Routledge and Kegan.

Trobevšek Drobnak, Frančiška

1990 “Expanded tenses in the Old English *Orosius*”, *Linguistica* 30: 13-46.

1994 “The Old English preverbal *ge-* in the light of the theory of language changes as strengthening or weakening”, *Studia Anglica Posnaniensia* 28: 123-142.

1999 “Linguistic change: The grammatical environment of participial non-finite clauses in Old English and in present-day English”, *Acta Neophilologica* 32: 99-109.