

Word frequency and word order in freezes¹

GERTRAUD FENK-OCZLON

Abstract

The theoretical part of this analysis attempts to explain and predict sequencing phenomena by means of principles of cognitive economy: in order to avoid peaks of information that are difficult to process (Fenk and Fenk 1980), and also because of their easier 'lexical' and articulatory accessibility, the more frequent and therefore informationally poorer elements tend to occupy initial position (Fenk-Oczlon 1983a). In the context of 'freezes', this means 'more frequent (high-token frequency) before less frequent'. The arguments presented in this paper indicate that this new rule represents a principle that is superordinate to such old rules as 'short before long', 'the first word has fewer initial consonants', 'me-first principle', etc.

In order to discover whether these conclusions stand up to empirical examination, they were tested on the basis of 400 freezes using the corresponding statistical data (from Thorndike and Lorge, Josselson, Meier, Ruoff). The results include the following:

– With 84% correct predictions, the new rule achieves by far the highest accuracy.

– In paired comparisons (all possible combinations of five rules) no other rule achieves such a high degree of correspondence.

– In order to explain those freezes which represent an exception to our rule, recourse must be had primarily to the iconic coding of spatial-temporal relationships.

1. Introduction

Linguistics, and even contemporary linguistics, often shows a defensive attitude toward quantification and in particular toward frequency-based arguments. There may be understandable historical reasons for this attitude, but nevertheless it obstructs the view of often simple solutions to old questions. The 'old question' that is the subject of the present

analysis is that of the determinants of word order in more or less 'frozen' conjoined phrases. The explanatory principle discussed and applied here — informationally poor comes before informationally rich — has already proved its value as the answer to a related question — why old before new? (Fenk-Oczlon, 1983a): 'to be new' can easily be replaced by 'to have a lower degree of familiarity', or by 'to have a higher degree of informational content'; to place such elements at the beginning of a sentence would infringe the principle to be described in the following pages.

The application of this principle is developed first (sections 2 and 3) on a mainly theoretical level and then (section 4) in the form of an empirical investigation.

2. Why 'old' before 'new'? A cognitive view

A long series of psychological investigations (for example, Hasher and Chromiak 1977; Kausler and Puckett 1980; Hasher and Zacks 1984) indicate that, even without intending to, almost automatically and incessantly, we register frequencies and differences in frequency. A further, almost trivial explanatory principle of a large number of findings in cognitive psychology is that our cognitive resources are informationally restricted (see Fenk 1985, 1986).

Second, language can be considered as a subsystem of the cognitive apparatus (see Chomsky), or at least as a system which develops together (in 'coevolution'; see Holenstein 1978) with other cognitive functions. It is therefore an obvious step to link the above-mentioned cognitive principles to the present question, especially since such a link has already proved its value in a related problem: cross-linguistic studies indicate a relatively constant and economic flow of linguistic information that takes into account the constraints of the cognitive resources (Fenk-Oczlon 1983b; Fenk-Oczlon and Fenk 1985). On the other hand, within a series of elements in which the element in position n reduces the number of possible elements in position $n+1$, the first positions have a higher informational content — the later elements can be anticipated more easily. The consequence for the positioning of linguistic elements is as follows:

In order to avoid an overloading of cognitive capacity and to achieve a constant information flow, there is a tendency not to put elements with high informational content at the beginning of the sentence, a position which contributes itself to a high informational content (Fenk-Oczlon 1983a).

In my opinion the psychologist Underwood's (1971) comment on extra-linguistic phenomena, namely that the difference between 'old' and 'new' is actually based on frequency, can be expressed more precisely in terms of information theory and can be applied to linguistic phenomena and problems (for instance to the question of why the subject tends to occupy sentence-initial position, Fenk-Oczlon 1983a). And thus a possible answer is given to the question posed and left unanswered in, for instance, Keenan (1978: 303) — why old before new?

It is precisely for this reason that what has already appeared in the preceding discourse, that is, what is 'old' and familiar in the textual context, or in the given context of action or situational context, bears less subjective information than a 'new' element in the same context (Fenk-Oczlon 1983a). In this context it is more expectable, its analysis requires fewer cognitive costs. It is in the interest of an economical and constant flow of information to place such informationally poor elements at the beginning of a sentence (and perhaps also of a phrasal conjunct?), because as the sentence or phrase progresses there is in any case a significant reduction of information (= a constriction of the permissible possibilities for continuing).

3. A new rule for the sequence in binomials: 'high frequency before low frequency'. Are the old rules reducible to this?

3.1. On more or less 'frozen' conjuncts

Terms such as (irreversible) binomials, phrasal conjuncts, *koordinierte* (nominal) *Phrasen*, 'freezes', *Erstarrungen* are used to describe the following observed linguistic phenomenon: in conjoined expressions, such as *Messer und Gabel* 'knife and fork', *Lust und Laune* 'pleasure and mood', *peak and valley*, *kurz und gut* 'short and good', convention lays down a more or less binding sequence (the term 'freeze' is often used where the convention is absolutely binding). Many rules and principles have been suggested which ascribe the particular position (mostly first or second) to particular characteristics of the words or features of the words. The suggestions range from those based on the particular language in question (Abraham 1950) to universal principles (Jespersen 1942; Malkiel 1959; Cooper and Ross 1975; Pinker and Birdsong 1979; Ross 1980; Allen 1987; among others).

Malkiel (1959) proposes six forces that cause a particular sequence in irreversible binomials. Four of them are semantic. For instance, the element that occupies first place chronologically or in the social hierarchy,

or is 'stronger', also tends to take first place in binomials. The other two 'forces' are more phonologically based and concern the length of words (shorter forms before longer ones), as well as rhythmic factors and alliterations. Malkiel did not suggest a hierarchy of these six 'forces', nor of their effects; however, he found that the principle of placing 'short' before 'long' was at least a very strong principle and, in many cases, overruled semantic factors.

Cooper and Ross (1975) emphasise the importance of a semantic 'me-first' principle in freezes: concepts and qualities that describe the prototypical speaker, or best apply to him, tend to occupy first position. The elements in first position thus tend to be 'living', 'adult', 'male', 'animate', 'here', 'now', 'patriotic', 'agentive', 'present generation', 'positive', 'solid', etc. A large number of 'freezes' for which no semantic explanation seems to apply are explained on the basis of phonological constraints: in comparison to the second word, the first has (1) fewer syllables, (2) shorter vowels, (3) fewer initial consonants, (4) less obstruent initial consonants, (5) more closed or more front vowels, (6) more final consonants, (7) more obstruent final consonants.

Pinker and Birdsong (1979) examine experimentally five of these seven criteria — criteria (1), (2), (4), (5), (6) — for possible universal validity. They concluded that only the number of syllables (the first word has fewer syllables than the second word) and the vowel quality (the first word has a high front vowel, the second has a low back vowel) also showed cross-linguistic validity. Vowel length and the obstruency of the initial consonants are considered as at best relevant for particular languages (such as English). As far as concerns the number of final consonants, they find that the rule must be reversed; at any rate, the test persons showed a preference for coordinated (nonsense) words in which the word with more final consonants was in second position.

Cutler and Cooper (1978) showed in a 'phoneme-monitoring experiment' that the phonemes were recognized more quickly in the sequence 'monosyllabic before bisyllabic' than in the reverse order. The authors found no significant difference in reaction times for the recognition of phonemes in the sequence 'high vowel before low vowel' and vice versa.

Other principles, which largely concern the semantic aspect, and which to some extent overlap each other, that have been suggested in the literature, include the following: the first element of (irreversible) binomials is semantically unmarked (Mayerthaler 1981), prototypical (Mayerthaler 1981; Edmondson 1985; Kelly et al. 1986), lexically more accessible (Bock and Warren 1985). And as Allen (1987) concludes from a comparative assessment of different rules, the first word is characterized by a higher familiarity. Allen's use of 'familiarity' is intended to include

Ertel's (1977) 'closeness to the speaker's cognitive field' as well as Kuno's (1979) 'speaker's empathy' and should at least be considered as 'the most powerful determinant' (1987: 51). (Other rules suggested by Allen are partly reformulations, partly supplements, and partly a synopsis of rules such as 'male > female', 'positive > negative', 'vertical > horizontal', 'man > beast', etc., previously formulated by Malkiel [1959]; Cooper and Ross [1975], and others. While Allen considers many of the sequencing rules, particularly those concerning social status, to be linked to culture and subject to change, he sees the sequencing of the elements in their 'natural order' — for example, chronology of events — as 'universal sequencing conventions'.)

The dimension 'familiarity' is very close to our rule of 'high frequency before low frequency'. For 'familiarity' is a general construct for the interpretation of diverse effects (such as short recognition time, few guessing errors, rapid association, etc.) and implies a repetition of input and/or practice. 'Familiarity' can only be operationalized or experimentally varied by means of 'frequency of input and/or practice'.

3.2. *The old rules from the point of view of frequency*

Our attempt to explain the rules set out in the works of the authors already mentioned in terms of frequency will begin with the phonological rules proposed by Ross (1980) (section 3.2.1) and then extend to semantic rules (section 3.2.2), the aim being to show that 'high frequency before low frequency (HF > LF)' is the 'covering law'.

3.2.1. *A frequency interpretation of Ross's phonological rules.* For Ross (1980), six of the seven phonological rules originally set up by Cooper and Ross (1975) and modified slightly by Ross are an expression of the length contrast 'short/long': shorter vowels and fewer initial and final consonants, as well as greater obstruency of the final consonants, all contribute to a shorter first word. And the high front vowels of the first word could also be indirectly connected to shortness — they 'sound' smaller, because — and here he refers to Sapir (1929) — (high) front vowels tend to be associated with small things. At least this last interpretation sounds a little far-fetched. As far as concerns the lesser obstruency of the initial consonants, he admits that he could not find any link to shortness (Ross 1980: 47).

It is possible that this rule might fit into the 'high frequency/low frequency' schema, and this may also apply to other rules. The question that must therefore be asked of every single rule in the following section

is, are the phonological principles diagnosed in the first word all merely concomitant to the fact that the first word is also the more frequent of the two conjuncts?

Rule a. The first word has fewer syllables than the second word.

This rule is an important principle for almost all authors. The replacement of 'short before long' (S > L) by 'high frequency before low frequency' (HF > LF) is obvious, since the most common words are generally also the shortest. This principle, described by Zipf (1929), was verified for German by Meier (1964: table on p. 241) in impressive detail. That frequency is not simply an epiphenomenon of shortness can be supported by observations of current language: as soon as words become frequent, they are shortened, for example, *Automobile* — *Auto* (German), United States of America — USA, etc. Once they are available in shortened form, they may become even more frequent, which may provoke a further shortening — as in US NAVY.

While a high negative correlation between number of syllables and frequency of a word is purportedly a linguistic universal, the following phonological rules — (b), (c), (d), (e), and (f) — are probably language-specific phonological characteristics (or specific to stress-timed languages?), which, however, again go hand in hand with a high frequency.

In this connection, it is interesting that the typical phonological characteristics of the first word, according to Ross (1980) — namely shortness of vowel, few initial consonants, greater sonority of initial consonants, etc. — are also the results of phonological processes which can be observed in casual rapid speech. What Donegan and Stampe (1979) call 'lenition processes' and what Dressler (1985) calls 'backgrounding processes', such as vowel reduction, consonant deletion, assimilation processes, and lenition of consonants, are usually characterized by the avoidance of articulatory effort and by an increase of perceptory effort.

However, in my opinion, principles of frequency and economy are also at work. Frequency of word for the speaker and expectability of the message for the hearer permit a reduction in the articulatory effort. Surprises for the hearer, the unfamiliar, cannot be reduced but tend to be emphasized and made prominent through greater articulatory effort (foregrounding processes).

According to Phillips (1984: 322), the (in my opinion) comparable phenomena of linguistic change affect the most frequent words most of all: 'CHANGES AFFECTING THE MOST FREQUENT WORD FIRST typically involve either vowel reduction and eventual deletion or assimilations — whether consonantal or vocalic, partial or complete.'

Rule b. The first word has less vowel length.

This criterion, which is found most of all in the analysis of English freezes, can in my opinion be linked with greater frequency without any difficulty. Thus Fidelholtz finds that the first syllable is reduced more often in the more frequent words than in the less frequent: '... the fact is apparent that vowels in frequently used words reduce more often than in relatively rare words' (Fidelholtz 1975: 208).

Rule c and rule d. The first word has fewer initial (c) and fewer final (d) consonants.

If, as a rule, increasing frequency leads to shortening, then it is at least plausible to explain a reduction of consonants, or, more accurately, fewer initial and final consonants, as a manifestation of higher frequency. At all events, a higher word frequency seems to go hand in hand with fewer initial and final consonants. This regularity is evidenced for German by Meier (1964), and an investigation, carried out by the present author, of the number of initial consonants using the frequency data in Josselson (1953) and Thorndike and Lorge (1944) shows that it also applies to Russian and English.

Of the 204 most frequent words in Russian (list 1 in Josselson 1953), only 36 words have two initial consonants (17.6%); and two words have three initial consonants (1%). In the 1993 words given in Josselson's lists 2-5 as the next frequent words, 555 (27.8%) have two initial consonants and 49 (2%) have three.

Of the 500 most frequent words in English (1-500 in Thorndike and Lorge), only 37 (7.4%) have two initial consonants, and two (0.4%) have three. Of the next frequent words (500-1000), 90 words (18%) have two initial consonants and four (0.8%) have three.

Rule e and rule f. The first word has less obstruent (more sonorant) initial consonants (e), but more obstruent (less sonorant) final consonants (f).

As already mentioned, in the experiments carried out by Pinker and Birdsong (1979), lesser obstruency of the initial consonants of the first word was only preferred by English native speakers; it was, therefore, only effective for this particular language. Frequency can also be made responsible for this language-specific phenomenon, namely by considering the lesser obstruency of initial consonants as a consequence of a reduction of articulatory effort (which is to be found above all in English?). 'As usage becomes more frequent, form becomes less accented, or more easily pronounceable, and vice versa' (Zipf 1929: 4). And the

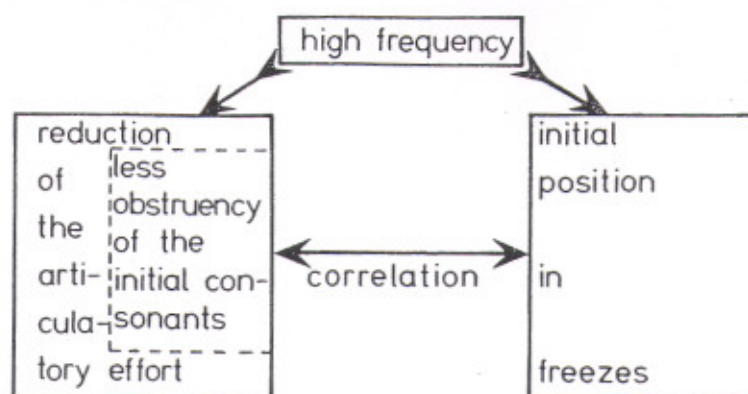


Figure 1. Frequency as the responsible factor for certain rules — for example, the correlation between word position and obstruency of the initial consonants

lenition of initial and intervocalic consonants is again typical of backgrounding processes that reduce the articulatory effort (see Dressler 1985). In Figure 1 an attempt is made to illustrate these connective structures in a way that also shows the explanatory principle presented in section 3.2.

However, the increase in obstruency of the final consonant (for example, devoicing) is also linked with a reduced articulatory effort (assimilation to the 'voiceless' pause), which is why this increase should tend to occur first of all in the more frequent words. The following reference in Phillips (1984: 321) can be interpreted as support for this assumption: although final devoicing in OE texts was rather sporadic, in 'all three scribes of the West Saxon Pastoral care' (Sweet 1871/1872) the extremely frequent word *sint* (= pres. indic. pl. of *wesan* 'to be') is written throughout (299 times) with (t), whereas with similar but less frequent words, such as *blind*, *find*, *wind*, *send*, there is no final devoicing.

An analysis of the relationship between the degree of obstruency (using Ross's [1980: 41] obstruency scale) and the frequency of initial consonants in English, which I carried out on the basis of frequency data from Thorndike and Lorge, gave the following results:

If we take all the words that begin with one of the glides [y], [w], [h] as our basis of 100%, then their share in the highest-frequency stage of Thorndike and Lorge — 'AA' (100 or more per million) — is 9% (The figures broken down give 16.7% for [y], 10.7% for [w] and 7.1% for [h].) For the words beginning with a liquid, a nasal, or a fricative, the percentages calculated in the same way are noticeably lower, namely 6.3% (= 5.5% for liquids, 6.5% for nasals, and 6.6% for fricatives; within the fricatives, [θ] is conspicuous by its high percentage of 15.5%). And of the words that begin with a stop, that is, with the most obstruent sound, only 4.5% are to be found in the highest-frequency stage.

This percentage distribution could have been guessed at for English

even without a systematic statistical analysis: of the particularly frequent personal pronouns *you, she, he, we, they*, not one begins with a stop.

As already mentioned, Ross (1980) was unable to link lesser obstruency of the initial consonants to shortness; however, the percentages just presented indicate a relationship to frequency, namely in the sense that the lesser obstruency of the initial consonants (at least in English) is an accompanying effect of a higher frequency of use.

Rule g. The first word tends to have a (high) front vowel, the second word a (low) back vowel.

Ross (1980) considers this to be a very strong principle and links it above all to the coding of spatial relationships. In many languages proximity is expressed by a word with a front vowel and distance by a word with a back vowel. For instance, German *hier/dort*, Russian *zdes'/tam*, French *ici/là*, Spanish *a qui/a la*, etc. Ross uses the following sequence from unrounded front vowels to rounded back vowels (1980: 42):

[i], [u], [e], [o], [ɛ], [ɔ], [a]

Such an interpretation is corroborated by the so-called 'echo words' (*Zwillingswörter*, reduplicative compounds). Here too it is well known (see Malkiel 1959; Mayerthaler 1977; Ross 1980) that front vowels dominate the first element. Examples drawn from Malkiel (1959) and Mayerthaler (1977) include *Piffpaff, Bimbam, Hickhack, klipp und klar, crisscross, splish-splash*.

However, these are not binomials in the narrow sense of the word, but linked syllables of which neither, or at best one, amounts to a word by itself or has an independent meaning. This means that the rule we posit for binomials, 'high frequency before low frequency', is neither applicable to nor testable against these: 'criss' or 'splish' do not exist on their own and therefore cannot be more or less frequent than 'cross' or 'splash'.

As already mentioned in 3.2.1, because this principle does not fit into the short/long schema, Ross used a rather artificial-sounding argument: he suggests that a word with a front vowel sounds 'smaller' or 'shorter'. The two following possible explanations appear to me not as far fetched:

a. Perhaps the symbolism of the sound takes into account the following factor: the further away a sound (thunder, a shout) is, the deeper and more 'muffled' it is perceived; and the sound that reaches us directly from its source, and thus reaches us first, sounds 'clearer' than one that comes to us indirectly and thus later. We are also reminded of the Doppler effect; the whistle of an approaching railway locomotive sounds higher in pitch, but when it has passed, the sound seems lower. Such an explanation based upon principles of physics would also create

a link between two attempts at explanation proposed by Ross — spatial deixis and sound symbolism.

b. Perhaps one ought also to take into account the following phonetic rule: 'As far as the vowels are concerned, their duration appears to be correlated with tongue height: other factors being equal, a high vowel is shorter than a low vowel' (Lehiste 1970: 18). And in the present case perhaps frequency can again be brought into play: for the frequent is encoded more briefly, and the use of shorter vowels is a proven means of achieving this brevity. Statistical data support this interpretation: the 20 most frequent words in German are all monosyllabic, and the syllable contains an [i] in eight cases, an [e] in six, an [u] in two, an [a] in two, an [o] in one and an [ei] likewise in one. That is to say that 16 of the 20 most frequent words have a (high) front vowel in the first (and only) syllable.

The fact that to the ears of the participants in an experiment using nonsense words (Pinker and Birdsong 1979) the sequence 'front vowel before back vowel' sounded better may simply mean that, perhaps because of the principles noted in (a) and/or (b), this sequence was more familiar to the test persons from their preexperimental experience.

Interim conclusion. Six of the seven phonological rules proposed by Ross fit more easily under the superordinate rule 'HF > LF' than under Ross's suggested superordinate rule 'S > L'. This is particularly noticeable in the case of rule (e) (less obstruency of the initial consonants of the first word), which Ross was unable to fit into his 'short > long' principle. If we take 'shortness' as an economically motivated correlate of frequency and accordingly attempt to submit rule (e) to the high frequency > low frequency principle, then the empirical data provide the justification.

Rule (g) '(high) front vowel before (low) back vowel', can be understood either as the result of sound-symbolic coding or possibly (in German) by the fact that high front vowels are, as a result of their shortness, found particularly frequently in particularly frequent words.

3.2.2. *A frequency interpretation of semantic rules*

Rule h. The 'me-first' principle.

According to Ross (1980: 48), of two conjuncts, the first position is occupied by the one which best refers to, describes or embodies the prototypical speaker, the 'normal case'. However the 'normal case' of the speaker is the statistically normal case, and the first element of two conjuncts is (normally) also the more frequent. It might be interesting to discover the reasons for the greater frequency of particular speaker characteristics, be they of a 'biological', psychological, or sociocultural

nature (see on this topic above all Mayerthaler 1981), or to examine to what extent these 'reasons' themselves are of a frequency 'nature' (Fenk and Fenk-Oczlon 1987). An obvious starting point is at any rate the relative frequency of the first word — first, it is a variable that can be explained independently of specific theoretical constructions, and for this reason it can — as a second stage — be related to a broad spectrum of interpretative and theory-dependent dimensions, for instance to the phonological properties that characterize the first word in the sense that they are found there relatively frequently (see above).

Ross takes the following 'metaphor' as the basis for the interaction between the semantic and phonological-phonetic properties of the individual conjuncts: *Je mehr Ton, desto mehr Bedeutung* 'The more sound, the more meaning' (Ross 1980: 48). However, in most cases this ('iconic') principle — 'formal complexity correlates with semantic complexity' — which has also been formulated by Jakobson (1971), Anttila (1972), Mayerthaler (1981), and others, can be considered as the consequence of differences in frequency (Fenk-Oczlon 1986, 1987b) or as an 'economically motivated index of familiarity' (Haiman 1983: 802). Particular factors such as 'natural salience' or the cultural importance of particular concepts lead to a higher frequency of the lexical units that represent them, and this higher frequency again leads to the shortness of the linguistic forms (Witkowski and Brown 1983).

Rule i. The first word is more prototypical and lexically more easily accessible.

This rule extends beyond the 'me-first' principle not only in that it is words which denote 'prototypical speaker qualities' that occupy initial position, but also that, in conjoined phrases, prototypical category members are generally placed before the less typical category members. Kelly et al. (1986: 59) find support for this wider claim in their experimental results:

The first experiment showed that sentences were systematically changed in recall to allow prototypical instances of categories to be mentioned before nonprototypical instances. In the second experiment, sentences in which the prototype preceded the nonprototype were judged more natural than sentences with the opposite order. Finally, an examination of dictionary definitions of categories found that prototypes tended to occur before nonprototypes.

The authors were aware of the fact that the prototypicality of a word correlated with its frequency. In order to exclude the possibility that in fact frequency was the decisive factor for the effects found, they checked

whether the words which tended to be placed first in experiment 1 were also the most frequent. They found that this was not always the case. However, this sort of verification is subject to the objection that it is inappropriate to draw on mere token frequency (Fenk-Oczlon 1987/1988): for whether an element is classed as more 'prototypical' depends on the 'frequency of instantiation'. This can be defined as 'someone's subjective estimate of how often they have experienced an entity as a member of a particular category' (Barsalou 1985: 631). More generally, it depends on the context in question (Roth and Shoben 1983). It can therefore be assumed that the element that is experienced as most frequent in a particular context and in connection with a particular category name will also occupy initial position.

What does it then mean when for instance Kelly et al. (1986: 71) suggest 'gun' as prototypical for the category 'weapon' and 'stone' as a nonprototypical example (therefore 'guns' before 'stones'), although 'stone' in absolute terms is more frequent? At any event, this does not mean a yawning gap between 'prototypicality' and 'frequency': for in the context of 'weapon', in the context of hunting and fighting, 'gun' is indeed the more frequent word.

Seen in this light, the following result does not hold any surprises: 'mean levels of meaning-preserving recall were higher (...) for sentences in which the prototype was more frequent than the nonprototype' (Kelly et al. 1986: 66). Here one has to consider cases in which a high absolute (=in all contexts) frequency coincides with a high context-specific frequency.

When Kelly et al. see the preference for the sequence 'prototype before nonprototype' in conjoined phrases as determined primarily by the greater lexical accessibility of the prototypes — as Bock (1982) and Bock and Warren (1985) had earlier seen lexical accessibility as central for the ordering in conjoined phrases — then the connection to frequency becomes even more obvious. For frequency is the central factor in almost all theories on lexical accessibility; the more frequently a word is found in a particular context, the faster lexical decisions can be made in these contexts (for example, Bradley 1978; Gordon 1985), and the quicker objects can be correctly named (Oldfield and Wingfield 1965).

A second interim conclusion. The two 'semantic' rules (h, i) are closely interconnected: both concern the 'prototypical' speaker, and on closer analysis 'prototypicality' reveals itself as a frequency phenomenon. For the lexical accessibility mentioned in rule i we do not need a theoretical analysis: the *conditio sine qua non* for a high level of accessibility is high

familiarity on the basis of a high frequency of input and/or practice. (Further 'semantic' rules — 'chronology of the events', 'upward before downward', etc., will be analyzed in the empirical section).

4. An empirical investigation

Up to now empirical-statistical material has been used as evidence that many of the 'old' rules are connected with frequency in some way or other. The statistical investigation set out in the following section pursues a different aim: it is intended first to confront the efficiency of our new rule with that of the old (4.1), and second to discover which of the other rules best apply to those freezes that represent an exception to our rule (4.2). At the same time it will reveal that in a few rare cases, in which the tendency of 'high frequency before low frequency' clashes with certain iconic principles, it is the latter that prevail (4.3).

4.1. Which rule achieves the better accuracy?

The token frequency of the elements (the words of 400 freezes, largely drawn from the relevant literature)² was ascertained on the basis of appropriate word-frequency lists (Thorndike and Lorge 1944 for English; Josselson 1953 for Russian; Meier 1964 and Ruoff 1981 for German; in addition, Brown, cited in Hofland and Stig 1982, for US English). Thus every individual freeze could be tested as to whether it corresponded to the rule 'high frequency before low frequency' (HF > LF), which is expected to be a universal rule. In addition, investigation was made as to which of the following rules the freezes obeyed: 'the first word has fewer syllables than the second' (S > L); 'the first word has fewer initial consonants than the second' (C > CC); 'the first word has high front vowels, the second low back vowels' (FV > BV); 'semantic principles' (me-first principle, chronology of events). Other rules suggested in the literature were not examined, given the small number of instances or the problematical decision as to applicability in the concrete case.

The following observations are intended to insure that the investigation is reproducible:

– If the frequency of several word forms was given for one word stem, the frequencies were added. In some cases, in which the frequency differences between the two conjuncts were glaring, this extra work was dispensed with.

– In the applications of the rule $FV > BV$ I have followed Ross's (1980) suggested ordering of the vowels (see section 3.2.1, rule g).

– The database used for English was primarily Thorndike and Lorge. It was only in those cases where the two conjuncts fell within the same frequency class in Thorndike and Lorge (AA or A) that use was made of the data in Brown (the values in question are marked with B in the Appendix).

– The database used for German was primarily Meier. Ruoff was only drawn on for words not dealt with in Meier (the values in question are marked with R in the Appendix). This system was also maintained for such cases where Ruoff would have supplied frequencies that were more 'favorable' to the hypothesis (for example for *Pech und Schwefel*, no. 200 in the Appendix), with one exception: before I had decided upon this system, the binominals *zu Wasser und zu Land* (no. 296 in the Appendix) and 'land and sea' (no. 294), were discussed as an example of a 'switching over' — depending on the frequency in the different languages (Fenk-Oczlon 1987a) — and they still remain illustrative in this respect (see section 4.3). The data for no. 296 come from Ruoff, however; the data in Meier, as transpired later, give a different picture. Sometimes, therefore, the existence of a 'switching over' depends on the frequency statistics used, unfortunately.

The result of a comparison $HF > LF$ with the competing rules. The accuracy of $HF > LF$ is significantly higher than that of the other rules (see Figure 2). Only 16% of the freezes represent an exception to this rule. The next-best rule (rule 'sem.')

fails to apply to more than 60% of freezes, and this despite the fact that this rule actually stands for a whole group of rules ('semantic principles'). The rule $S > L$ is even less successful, although it is closely connected with our rule in that the more frequent is mostly encoded as the shortest. The rule $S > L$ cannot even be applied to many freezes (244 in number) since the first word and the second word, measured in terms of the number of syllables, are equally long. The rule $FV > BV$ applies to 28.0% of the cases, and $C > CC$ to only 17.5% of freezes.

4.2. Paired comparisons between the individual rules

The proof that the rule we suggest has by far the highest explanatory power, or explains the most cases, does not of course exclude the possibility that other tendencies — although to a lesser degree — might also be in operation. An indication of this should be discovered by paired comparisons of the rules under investigation.

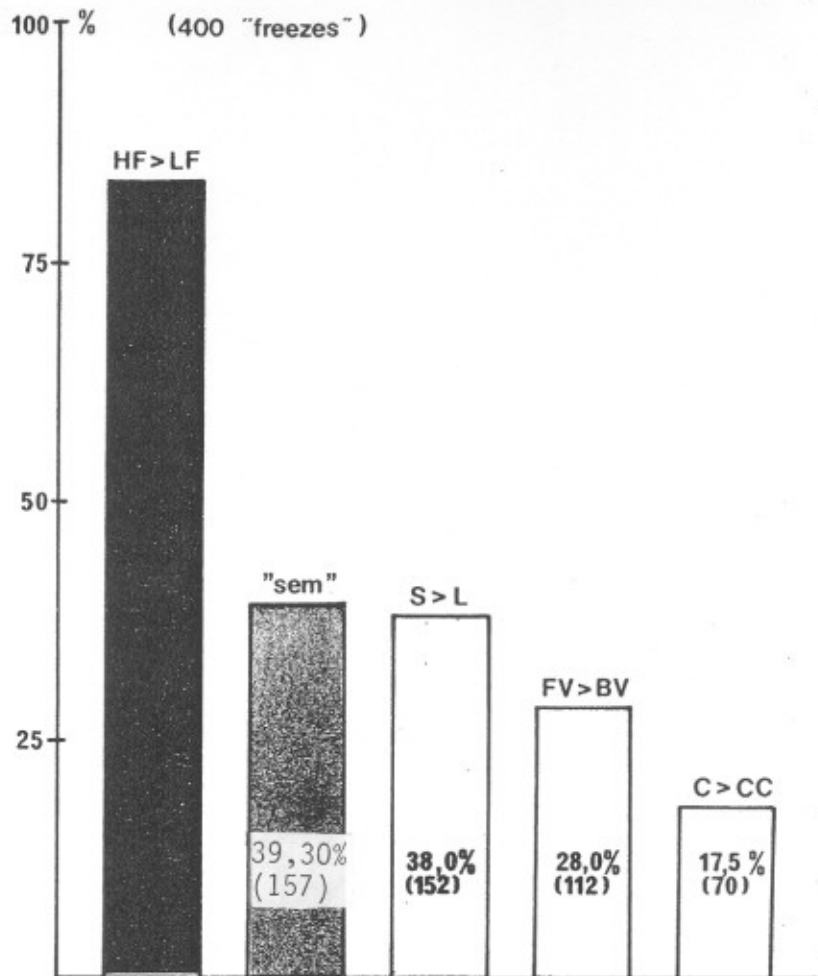


Figure 2. A comparison of the five rules as far as concerns the success quota in percentage (in brackets the number of cases to which the rule applies)

The results of the paired comparisons.

a. The comparison of HF > LF with each of the other rules separately, as set out in Figure 3, shows that a large percentage of the cases explained by the other rules are also explained by our rule. The other findings can be seen from a comparison between Figure 3 and Figure 4.

- The overlap between our rule and each of the competing rules is larger than the overlap in the paired comparisons between the competing rules.

- In the paired comparisons between the competing rules (Figure 4), and in contrast to the paired comparisons between our and the competing rules (Figure 3), the overlaps (the ranges of correspondence) are smaller than the ranges of discrepancy. Table 1 sets out the exact percentages of congruence.

b. In comparison to the rule HF > LF (Figure 3), the 'sem.' rule ('semantic principles') has the largest independent explanatory field: it 'explains' 28 freezes which our new rule HF > LF does not cover.

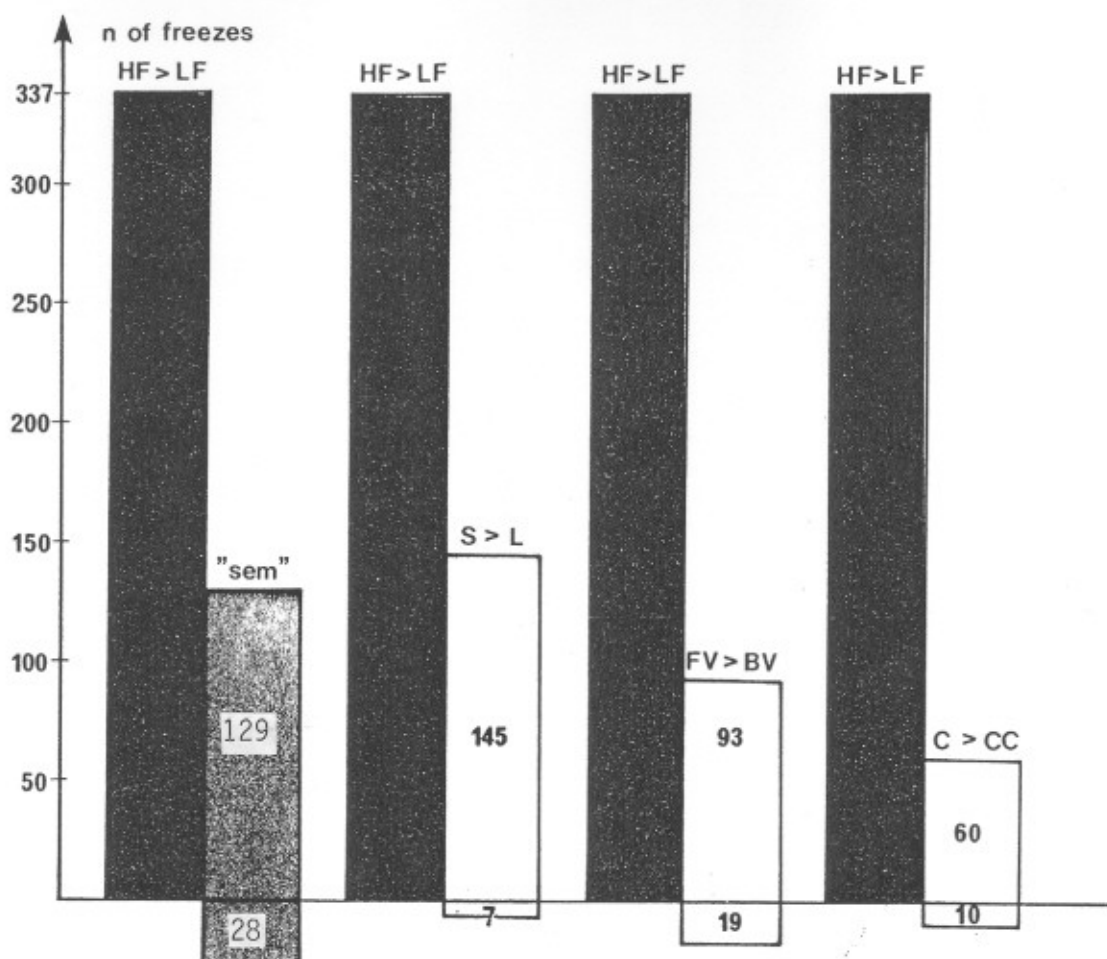


Figure 3. The range of correspondence of the rule $HF > LF$ with that of the other four rules

4.3. Frequency versus semantic principles

Since the principles subsumed under the rule 'sem.' can also to some extent be interpreted in terms of frequency (see 3.2.2), the following question had to be asked: do the 28 freezes that are explained by the 'sem.' rule but not by the rule $HF > LF$ have anything in common that might indicate a tendency that operates independently of the frequency tendency?

a. *Temporal relationships.* In the following freezes (in brackets the relevant number in the Appendix), that is, in 9 of the 28 cases, it is obviously a case of the 'iconic' coding of the temporal sequence:

bred and born (81); *now and again* (143); *birth and death* (211); *now and then* (241); *sooner or later* (242); *past and present* (243); *früher oder später* 'sooner or later' (245); *gestern und heute* 'yesterday and today' (247); *Anfang und Ende* 'beginning and end' (383).

(Whether *bred and born*, taken from Malkiel [1959], fits into the chrono-

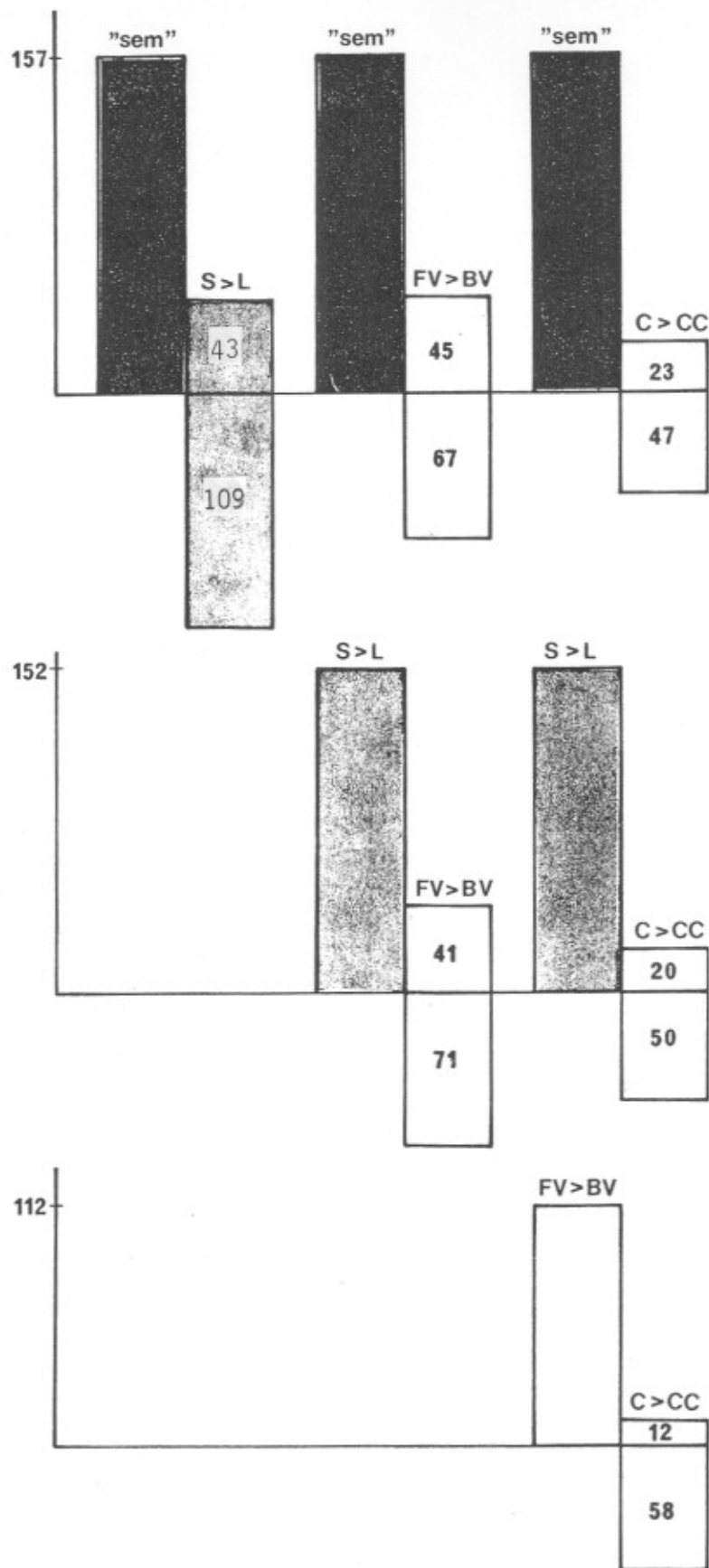


Figure 4. The range of correspondence between the competing rules to HF > LF (in analogy to and as supplement to Figure 3)

Table 1. *The explanatory field that the weaker rule in each case shares with the compared rule (in percentages of the weaker rule)*

	HF > LF	S > L	C > CC	FV > BV	'sem.'
HF > LF					
S > L	95.4				
C > CC	85.8	28.6			
FV > BV	83.0	36.6	10.7		
'sem.'	82.2	28.3	32.9	40.2	

logy schema depends on which of the possible meanings one ascribes to the word *bred*. Nevertheless, several native speakers inform me that the reverse form, *born and bred*, is more familiar to them. This reversal would then be no exception to our HF > LF rule.)

The chronology interpretation fits a few other freezes, but not so unambiguously, for example in *knowledge and action* (321) and *Knall und Fall* 'bang and fall' (59). In both cases the chronological sequence is unclear: an action need not (only) result from particular knowledge — it can be carried out in order to acquire particular knowledge. And although game does fall (*Fall*) after the bang (*Knall*) from the gun, an object falls before it bangs on the floor. Other objections could be made about *parent and child* (256) and *actor and action* (302). The argument that parents or actors must be present before they produce children or carry out actions can be countered by the fact that parents and actors logically only become parents and actors as a result of their children and their actions. And even if we accept the chronology interpretation, it must be admitted that for nos. 256 and 302 a different interpretation principle sounds plausible: both could be interpreted as iconic codings of a 'producer/product' relationship that is not restricted to temporal aspects. And in *parent and child*, social and politeness conventions may also play a part.

b. *Spatial relationships*. In the binomials in question the spatially near (nos. 230, 231, 238, and 239), the item in front (no. 297), or the item above or higher (nos. 322 and 322) comes first:

here and there (230); *this and that* (231); *hier und da* 'here and there' (238); *zdes'i tam* 'here and there' (Russian) (239); *front and back* (297); *peak and valley* (322); *upper and lower* (332).

c. *Movement (phenomena experienced in the categories of space and time)*. It is to be noted that in the relevant freezes (nos. 216, 330, and 331) the upward is placed before the downward:

rise and fall (216), *upstairs and downstairs* (330), *ascending and descending* (331).

We can further generalize that if one takes as starting point the (observer of a) subject acting, operating, and moving in space, then the cases noted under (b) and (c) can also be interpreted as the iconic coding of chronology relationships [see a]):

The more distant (less close) reaches the subject later; the item in front lies within the field of vision. And if one wants to see what is happening behind, one must turn (or turn one's head) (see Mayerthaler 1981). And for someone or something to come down, he or it must usually have gone up first (of its own or of another's effort) — a point of view that particularly corresponds to a being for which movement on the ground is the norm and therefore represents the reference point (for the undertaking, the analysis, or the description of changes).

In brief: by means of the concept of planned movement and movement anticipation, one can explain almost all freezes whose sequencing is in some way or other 'semantic' and does not fit our frequency rule; and this explanation is in terms of a coding of chronological aspects.

How then can we assess the freezes which are an exception to our rule? Three points seem essential here:

- A sequence that is in contradiction to our 'high frequency before low frequency' does not necessarily contradict the general economy principles.
- An iconic coding of before-and-after relationships may be cognitively simpler (for speaker and hearer), even if this coding begins with the less frequent element.
- In by far the greater number of cases in which our frequency principle conflicts with one or other 'semantic' principle (such as the me-first principle) the frequency principle prevails.

This is shown on the one hand by the large areas of overlapping of the corresponding columns in Figure 3, and, in more detail, for instance, by the fact that our frequency principle quite definitely deals with a large number of freezes that are classed in the literature (Cooper and Ross 1975) as exceptions to the 'me-first' principle that are difficult to explain.

Thus the 'male > female sequence' in freezes is reversed when the frequency relationships are reversed: compare the following freezes:

goose and gander (270), *duck and drake* (271), *Ente und Enterich* (272), *Gans und Gänserich* (273), *bride and groom* (274), *Braut und Bräutigam* (275).

Likewise the exception to the rule 'human > nonhuman', *Pferd und Reiter*

'horse and rider' (378) could be explained by the greater frequency of *horse*. Perhaps more instructive are those freezes in which words of similar meaning (= semantically similar) in different languages take either first or second place depending on frequency in the language in question, undisturbed by semantic or me-first principles:

No. in Appendix	Element 1 (token fr.)	Element 2 (token fr.)	No. in Appendix	Element 1 (token fr.)	Element 2 (token fr.)
294	land and (218)	sea (95)	296	zu Wasser und 'on water and (140)	zu Land on land' (55)
229 (Russian)	živoj ili 'living or (86)	mertvyj dead' (62)	313	dead or (AA)	alive (A)
310	living or (B)(194)	dead (B)(174)	314	tot oder (512)	lebendig (429)

5. General discussion

5.1. *The constant flow of information — one principle for the explanation of context-dependent and 'context-independent' sequencing*

The opposition in the heading between context-dependent and 'context-independent' is of course a simplification. In truth we are dealing with a continuum between dependency of direct (linguistic and situational) context on the one hand, and fairly fixed, frozen, context-invariant sequencing on the other hand. In the speech situation of the moment, we can thus distinguish between different degrees of freedom as far as concerns the sequence of elements in coordinated (nominal) phrases:

a. Smallest degree of restriction: these are not 'freezes', nor, because for instance a discourse is just beginning, are we concerned with factors of the direct context of discourse which restrict the freedom of choice. In this case the 'lexical accessibility' in the sense of Bock and Warren (1985) remains the decisive factor. Or, put in more general terms, the decisive factor is the different familiarity of the two elements, but familiarity on the basis of the learning history of the speaker, going back beyond the present context. The familiarity of a word is essentially determined by the token frequency of a word independent of context or going beyond context, as is to be found in dictionaries of word frequency. This tendency is always involved — but it can be affected (neutralized or emphasized) by other factors under the conditions described in (b) and (c).

b. Greater degree of restriction: restriction by the factors mentioned in (a) and by the current context. The element frequently mentioned in the previous discourse will occupy first place in coordinated (nominal) phrases, in the sense of a link to the main subject, to what has been frequently mentioned in the previous discourse and is thus (or as a result of the 'extralinguistic' context) present and does not first have to be drawn from the depths of memory.

c. Greatest degree of restriction: restriction of choice by old, fixed, 'frozen' conjuncts, or 'freezes'. Here the very combination of the two elements is familiar. One can of course ask why the two words were once frozen into this order. In my opinion the answer is again in the tendency mentioned in (a) and analyzed in the present work: the more familiar word — familiar as a result of a higher token frequency independent of context — was fixed in first position.

Conclusion. The economy principle of constant information flow provides a plausible explanation for the order within more or less frozen conjuncts. As a linguistic sequence progresses, the number of possible continuations becomes more and more restricted; that is, there is a reduction of the uncertainty or of the information. In order to avoid peaks of information, there is a tendency to place the element that is already more familiar or which contains less information in the less easily predictable sequence-initial position — for example, at the beginning of (part of) a sentence, of a freeze, etc. What is familiar is easier for the speaker to call up and is more expectable and therefore more easily perceivable for the hearer.

As applied to freezes, this principle results in the rule 'high frequency before low frequency', and in the empirical comparison described here, this rule has far exceeded in accuracy all other rules suggested previously. Nevertheless, it cannot claim to be the only determinant of the sequence in freezes: it is clear that iconic tendencies can prevail in the coding of temporal and/or spatial relationships if they come into conflict with our frequency tendency — which, however, does not occur very often.

In 1975, Cooper and Ross suggested that 'conjuncts which are easier to process tend to occupy place 1 in a freeze' (1975: 92, see also 103). And the present analysis has specified and shown those conditions under which a unit is more easily processed: it is above all more easily processed if it has — at least in similar contexts — become familiar as a result of frequent use.

Received 29 August 1988
Revised version received
2 January 1989

University of Klagenfurt

Appendix

	Word 1		Word 2	f 1	f 2	rule HF> LF	rule S> L	rule C> CC	rule FV> BV	rule sem.
1.	Angst	und	Schrecken	542	119	+	+	+	-	-
2.	Max	und	Moritz	121	84	+	+	-	-	-
3.	Furcht	und	Schrecken	937	119	+	+	+	+	-
4.	Glanz	und	Gloria	407	34	+	+	-	-	-
5.	Grund	und	Boden	3146	2166	+	+	-	+	-
6.	Kraut	und	Rüben	42	64	-	+	-	-	-
7.	(mit) List	und	Tücke	150	58	+	+	-	-	-
8.	(hinter) Schloß	und	Riegel	(R)3	(R)1	+	+	-	-	-
9.	(mit) Sinn	und	Verstand	1454	829	+	+	-	+	-
10.	Wind	und	Wetter	616	529	+	+	-	+	-
11.	auf	und	davon	80944	3909	+	+	+	-	+
12.	dumm	und	dusselig	133	0	+	+	-	-	-
13.	groß	und	kräftig	1568	295	+	+	-	-	-
14.	klar	und	deutlich	1725	885	+	+	-	-	-
15.	kurz	und	bündig	1739	19	+	+	-	-	-
16.	nie	und	nimmer	4834	368	+	+	-	-	-
17.	recht	und	billig	(R)197	(R)28	+	+	-	-	-
18.	still	und	leise	935	540	+	+	-	+	-
19.	Busch	und	Tal	60	296	-	-	-	+	-
20.	Jacke	wie	Hose	25	13	+	-	-	-	-
21.	Mann	und	Maus	7069	53	+	-	-	-	+
22.	Stock	und	Stein	224	486	-	-	-	-	-
23.	kurz	und	klein	1739	640	+	-	+	-	-
24.	drunter	und	drüber	37	95	-	-	-	-	-

25.	ganz	und	gar	13584	6375	+	-	-	-	-
26.	immer	und	ewig	11664	660	+	-	-	+	-
27.	trick	or	treat	(B)15	(B)26	-	-	-	-	-
28.	stress	and	strain	(B)107	(B)31	+	-	-	-	-
29.	hem	and	haw	14	1	+	-	-	+	-
30.	angst	und	bang	542	42	+	-	+	-	-
31.	Ach	und	Krach	1628	19	+	-	+	-	-
32.	Saft	und	Kraft	98	2806	-	-	+	-	-
33.	Schutz	und	Trutz	946	21	+	-	+	-	-
34.	Weg	und	Steg	3257	23	+	-	+	-	-
35.	singen	und	springen	550	306	+	-	+	-	-
36.	recht	und	schlecht	(R)197	(R)112	+	-	+	-	-
37.	weit	und	breit	5768	294	+	-	+	-	-
38.	ab	und	zu	4559	172625	-	-	+	-	-
39.	sea	and	ski	AA	6	+	-	+	-	-
40.	by hook	or by	crook	47	10	+	-	+	-	-
41.	long	and	strong	(B)756	(B)202	+	-	+	-	-
42.	price	and	pride	AA	A	+	-	-	-	-
43.	sink	or	swim	(B)23	(B)15	+	-	+	-	-
44.	Handel	und	Wandel	1007	173	+	-	-	-	-
45.	Rat	und	Tat	(R)7	(R)3	+	-	-	-	-
46.	Sack	und	Pack	(R)75	(R)5	+	-	-	-	-
47.	Schritt	und	Tritt	(R)7	(R)2	+	-	-	-	-
48.	wear	and	tear	(B)36	(B)11	+	-	-	-	+
49.	(mit) Kisten	und	Kasten	146	226	-	-	-	+	-
50.	Tür	und	Tor	957	320	+	-	-	+	-
51.	singen	und	tanzen	550	184	+	-	-	+	-
52.	dick	und	fett	284	197	+	-	-	+	-
53.	voll	und	ganz	1844	13548	-	-	-	+	-
54.	by guess	and by	gosh	AA	6	+	-	-	+	-

Word 1		Word 2	f 1	f 2	rule HF> LF	rule S> L	rule C> CC	rule FV> BV	rule sem.
55.	Tag	und	Nacht	6751	4082	+	-	-	+
56.	dies	und	das	8983	124232	-	-	+	-
57.	dieses	und	jenes	11113	1103	+	-	+	+
58.	(durch) dick	und	dünn	505	462	+	-	-	+
59.	Knall	und	Fall	26	2993	-	-	-	+
60.	Bus	und	Bahn	(R)10	(R)7	+	-	+	-
61.	Männer	und	Maschinen	1517	299	+	+	+	+
62.	Herr	und	Hund	28884	463	+	-	-	+
63.	Stein	und	Bein	976	456	+	-	-	-
64.	push	and	pull	A	AA	-	-	-	+
65.	safe	and	sane	AA	11	+	-	-	-
66.	aches	and	pains	28	AA	-	-	+	-
67.	aid	and	abet	A	1	+	+	-	-
68.	all	and	any	17799	6532	+	+	-	+
69.	at	or	near	26250	1338	+	-	+	-
70.	beam	and	rafter	42	7	+	+	-	-
71.	big	and	little	(B)360	(B)831	-	-	-	+
72.	hunger	and	thirst	37	23	+	-	-	-
73.	buckle	and	thong	8	4	+	-	-	-
74.	bump	and	grind	16	18	-	-	+	-
75.	by	and	large	(B)50306	(B)361	+	-	-	-
76.	cap	and	gown	A	38	+	-	+	-
77.	(to give one) cards	and	spades	A	13	+	-	+	-
78.	cheap	and	nasty	A	7	+	+	-	-
79.	birds	and	bees	AA	A	+	-	-	-

80.	bit	and	blow	(B)101	(B)69	+	-	+	+	-
81.	bred	and	born	14	AA	-	-	-	+	+
82.	chalk	and	cheese	14	AA	-	-	-	-	-
83.	hot	and	heavy	130	110	+	+	-	-	-
84.	house	and	home	(B)591	(B)547	+	-	-	-	-
85.	life	and	limb	AA	38	+	-	-	-	-
86.	poor	and	pert	AA	2	+	-	-	+	-
87.	loud	and	long	(B)20	(B)756	-	-	-	-	-
88.	part	and	parcel	AA	23	+	-	-	-	-
89.	pots	and	pans	(B)33	(B)16	+	-	-	+	-
90.	line	and	level	AA	A	+	-	-	-	-
91.	rant	and	rave	1	9	-	-	-	-	-
92.	rough	and	ready	A	AA	-	+	-	-	-
93.	(mit) Kind	und	Kegel	2115	29	+	+	-	+	+
94.	(von) Ruf	und	Rang	632	226	+	-	-	+	-
95.	big	and	black	(B)360	(B)203	+	-	+	+	-
96.	heart	and	soul	AA	A	+	-	-	-	-
97.	rats	and	mice	37	34	+	-	-	-	-
98.	town	and	gown	AA	38	+	-	-	-	-
99.	Hände	und	Füße	1583	333	+	-	-	-	+
100.	den'	i	noc'	(L)1	(L)2	+	-	-	+	+
101.	brain	and	brawn	A	1	+	-	-	+	-
102.	facts	and	figures	(B)447	(B)209	+	+	-	-	-
103.	friend	and	foe	AA	23	+	-	-	-	+
104.	(ohne) Furcht	und	Tadel	937	154	+	+	-	+	-
105.	cool	and	coy	AA	2	+	-	-	+	-
106.	shirt	and	tie	(B)27	(B)23	+	-	-	-	-
107.	cup	and	saucer	AA	A	+	+	-	-	-
108.	knife	and	fork	A	31	+	-	-	-	-
109.	ham	and	cheese	(B)19	(B)9	+	-	-	-	-
110.	lamb	and	salad	45	28	+	+	-	+	-

	Word 1		Word 2	f 1	f 2	rule HF> LF	rule S> L	rule C> CC	rule FV> BV	rule sem.
111.	salt	and	pepper	AA	27	+	+	-	-	-
112.	joy	and	sorrow	AA	A	+	+	-	-	+
113.	blood	and	thunder	AA	46	+	+	-	-	-
114.	flesh	and	blood	A	AA	-	-	-	+	-
115.	(to) chop	and	change	29	AA	-	-	-	-	-
116.	death	and	destruction	AA	37	+	+	-	-	-
117.	fair	and	foolish	AA	36	+	+	-	-	-
118.	fast	and	furious	AA	21	+	+	-	-	-
119.	fat	and	fulsome	AA	17	+	+	-	-	-
120.	fine	and	dandy	AA	3	+	+	-	-	-
121.	fire	and	water	(B)187	(B)442	-	-	-	+	-
122.	(a dress) fits	and	flatters	AA	25	+	+	+	+	-
123.	fun	and	games	A	AA	-	-	-	-	-
124.	fuss	and	bother	11	30	-	+	-	-	-
125.	ghosts	and	goblins	32	14	+	+	-	-	-
126.	glow	and	glitter	A	23	+	+	-	-	-
127.	gold	and	silver	(B)52	(B)29	+	+	-	-	-
128.	hale	and	hearty	7	18	-	+	-	+	-
129.	hares	and	hounds	18	23	-	-	-	-	-
130.	(to agree like) harp	and	harrow	20	5	+	+	-	-	-
131.	health	and	happiness	AA	A	+	+	-	-	-
132.	high	and	dry	(B)498	(B)68	+	-	+	-	-
133.	(over) hill	and	dale	AA	13	+	-	-	+	+
134.	horse	and	rider	AA	20	+	+	-	-	-
135.	(to run) hot	and	heavy	130	110	+	+	-	-	-

136.	lean	and	lanky	A	1	+	+	-	+	-
137.	low	and	lonely	AA	35	+	+	-	-	-
138.	let	and	hindrance	AA	3	+	+	-	-	+
139.	meek	and	mild	7	32	-	-	-	+	-
140.	neat	and	clean	29	AA	-	-	+	-	-
141.	neck	and	crop	AA	A	+	-	+	+	+
142.	(in every) nook	and	cranny	9	2	+	+	+	+	-
143.	now	and	again	(B)314	(B)578	-	+	-	-	+
144.	oil	and	vinegar	AA	12	+	+	+	-	-
145.	part	and	parcel	AA	23	+	+	-	-	-
146.	peace	and	quiet	AA	A	+	-	-	+	-
147.	(to) pick	and	choose	AA	A	+	+	-	+	-
148.	(to sit on) pins	and	needles	43	34	+	+	-	-	-
149.	poor	but	honest	AA	A	+	+	-	+	-
150.	pot	and	kettle	47	27	+	-	-	-	-
151.	pure	and	simple	A	AA	-	-	-	-	-
152.	(to) live or lie	at	rack and manger	29	7	+	+	-	-	-
153.	rags	and	tatters	28	7	+	+	-	-	-
154.	root	and	branch	A	AA	-	-	+	+	-
155.	rough	and	tumble	A	30	+	-	-	-	-
156.	words	and	pictures	(B)548	(B)230	+	+	-	-	-
157.	sex	and	slaughter	26	14	+	+	+	+	-
158.	sin	and	shame	(B)53	(B)21	+	-	-	+	+
159.	(to) slip	and	slide	A	36	+	-	-	+	-
160.	slow	and	steady	(B)60	(B)41	+	+	-	-	-
161.	full	and	equal	AA	A	+	+	-	-	-
162.	sound	and	fury	AA	29	+	+	-	-	-
163.	straight	and	narrow	(B)114	(B)63	+	+	-	-	-
164.	strong	and	stormy	AA	21	+	+	-	-	-
165.	stuff	and	nonsense	A	19	+	+	-	-	-

	Word 1		Word 2	f 1	f 2	rule HF> LF	rule S> L	rule C> CC	rule FV> BV	rule sem.
166.	to	and	fro	AA	14	+	-	+	+	-
167.	tried	and	tested	AA	A	+	+	-	-	-
168.	views	and	volleys	AA	6	+	+	-	+	-
169.	waifs	and	strays	1	18	-	-	+	-	-
170.	whys	and	wherefores	AA	43	+	+	-	-	-
171.	wind	and	weather	(B)84	(B)69	+	+	-	+	-
172.	(by one's) wits	and	fists	A	23	+	-	-	-	-
173.	head	and	shoulders	(B)467	(B)112	+	+	-	+	+
174.	head	or	tail	(B)467	(B)24	+	-	-	-	+
175.	blood	and	iron	(B)121	(B)43	+	+	-	-	-
176.	bright	and	shiny	AA	7	+	+	-	-	-
177.	time	and	winter	(B)1601	(B)93	+	+	-	-	-
178.	rhyme	and	reason	14	AA	-	-	-	-	-
179.	fresh	and	frisky	AA	1	+	+	-	-	-
180.	heroes	and	heroines	A	7	+	+	-	-	+
181.	Greek	and	Latin	(B)61	(B)50	+	+	-	+	-
182.	warp	and	woof	12	2	+	-	-	-	-
183.	warp	and	weft	12	1	+	-	-	-	-
184.	hip	and	thigh	18	13	+	-	-	-	-
185.	horse	and	cow	AA	A	+	-	-	-	-
186.	cat	and	mouse	A	34	+	-	-	+	+
187.	sun	and	moon	(B)112	(B)60	+	-	-	-	+
188.	chapter	and	verse	A	38	+	-	-	+	-
189.	(a) gentleman	and	(a) scholar	AA	20	+	-	-	+	-
190.	brush	and	comb	A	19	+	-	-	-	-
191.	hammer	and	tongs	34	6	+	-	-	+	-

192.	classes	and	masses	AA	A	+	-	-	-	-
193.	salary	and	wages	(B)43	(B)42	+	-	-	-	-
194.	Hammer	und	Sichel	73	27	+	-	-	-	-
195.	Freud	und	Leid	2766	653	+	-	-	-	+
196.	Lust	und	Laune	839	306	+	+	-	+	-
197.	Nacht	und	Nebel	3907	334	+	+	-	-	-
198.	Messer	und	Gabel	230	131	+	-	-	+	-
199.	Sonne	und	Mond	1625	378	+	-	-	-	+
200.	Pech	und	Schwefel	38	349	-	+	+	-	-
201.	Land	und	Leute	4487	3452	+	+	-	-	-
202.	Krieg	und	Frieden	2313	1954	+	+	-	-	+
203.	wirklich	und	wahrhaftig	4393	313	+	+	-	+	-
204.	boot	and	sole	37	25	+	-	-	+	-
205.	snow	and	ice	(B)59	(B)45	+	-	-	-	-
206.	fame	and	riches	A	20	+	+	-	-	-
207.	fox	and	hounds	25	23	+	-	-	-	-
208.	coat	and	pants	AA	6	+	-	-	-	-
209.	man	and	beast	AA	A	+	-	-	-	+
210.	horse	and	buggy	AA	7	+	+	-	+	+
211.	birth	and	death	A	AA	-	-	-	-	+
212.	challenge	and	response	25	18	+	-	-	-	+
213.	marriage	and	divorce	A	19	+	-	-	-	+
214.	question	and	answer	(B)397	(B)152	+	-	-	+	+
215.	(from) start	to	finish	(B)154	(B)39	+	+	-	-	+
216.	rise	and	fall	(B)102	(B)147	-	-	-	-	+
217.	wedding	and	reception	35	19	+	-	-	-	+
218.	stop	and	shop	(B)120	(B)63	+	-	-	-	+
219.	heaven	and	hell	AA	31	+	-	-	-	+
220.	buy	and	sell	(B)111	(B)85	+	-	-	-	+
221.	love	and	hate	AA	A	+	-	-	-	+
222.	(to rain) cats	and	dogs	(B)23	(B)75	-	-	-	-	-

	Word 1		Word 2	f 1	f 2	rule HF> LF	rule S> L	rule C> CC	rule FV> BV	rule sem.
223.	meat	and	potatoes	AA	A	+	+	-	+	-
224.	potatoes	and	gravy	A	8	+	-	+	-	-
225.	Hören	und	Sehen	(R)2	(R)1	+	-	-	-	-
226.	mehr	oder	minder	18549	847	+	+	-	-	+
227.	Reih'	und	Glied	(R)18	(R)3	+	-	+	-	-
228.	kreuz	und	quer	298	118	+	-	-	-	-
229.	živoj	ili	mertvyj	86	62	+	-	-	+	+
here:										
230.	here	and	there	(B)750	(B)2725	-	-	-	+	+
231.	this	and	that	(B)5146	(B)10594	-	-	-	+	+
232.	hither	and	thither	27	25	+	-	-	-	+
233.	hither	and	yon	27	22	+	-	-	+	+
234.	come	and	go	(B)630	(B)626	+	-	-	-	+
235.	Kommen	und	Gehen	5300	3557	+	-	-	-	+
236.	da	und	dort	23497	5314	+	-	-	-	+
237.	hier	und	dort	16667	5314	+	-	-	+	+
238.	hier	und	da	16667	29497	-	-	-	+	+
239.	zdes'	i	tam	4	2	-	-	-	+	+
240.	in	and	out	(B)21337	(B)2096	+	-	-	+	+
now:										
241.	now	and	then	(B)1314	(B)1377	-	-	-	-	+
242.	sooner	or	later	(B)17	(B)397	-	-	-	+	+
243.	past	and	present	(B)281	(B)377	-	+	+	-	-

244.	heute	und	morgen	5259	2344	+	-	-	+	+
245.	früher	oder	später	2552	2961	-	-	-	+	+
246.	morgen	und	übermorgen	2344	75	+	+	-	-	+
247.	gestern	und	heute	1224	5259	-	-	-	-	-
248.	rano	ili	pozdno	62	60	+	-	-	-	+
present generation:										
249.	father	and	grandfather	AA	36	+	+	+	-	+
250.	son	and	grandson	AA	8	+	+	+	-	+
251.	Vater	und	Großvater	4509	228	+	+	+	-	+
252.	Enkel	und	Urenkel	209	19	+	+	-	-	+
253.	Kinder	und	Kindeskinder	2564	10	+	+	-	-	+
adult:										
254.	man	and	boy	(B)1207	(B)242	+	-	-	-	+
255.	father	and	son	(B)183	(B)166	+	-	-	-	+
256.	parent	and	child	A	AA	-	-	-	-	+
257.	mother	and	daughter	(B)227	(B)86	+	-	-	-	+
258.	Mutter	und	Tochter	3744	1714	+	-	-	+	+
259.	Onkel	und	Neffe	297	96	+	-	+	-	+
260.	Tante	und	Nichte	540	303	+	-	-	-	+
261.	cow	and	calf	A	14	+	-	-	-	+
262.	cat	and	kitten	A	35	+	+	-	-	+
263.	mare	and	foal	14	1	+	-	-	-	+
264.	man	and	woman	(B)1207	(B)242	+	+	-	-	+
265.	husband	and	wife	(B)131	(B)228	-	-	-	-	+
266.	boys	and	girls	(B)143	(B)142	+	-	-	-	+
267.	Bruder	und	Schwester	2262	1187	+	-	-	+	+
268.	brat	i	sestra	(L)2	(L)3	+	-	-	-	+
269.	ladies	and	gentlemen	(B)28	(B)21	+	+	-	-	+
270.	goose	and	gander	45	2	+	+	-	+	-

	Word 1		Word 2	f 1	f 2	rule HF> LF	rule S> L	rule C> CC	rule FV> BV	rule sem.
271.	duck	and	drake	49	8	+	-	+	-	-
272.	Ente	und	Enterich	42	0	+	+	-	-	-
273.	Gans	und	Gänserich	38	0	+	+	-	-	-
274.	bride	and	groom	41	16	+	-	-	-	-
275.	Braut	und	Bräutigam	391	209	+	+	-	-	-
positive:										
276.	positive	or	negative	15	10	+	-	-	-	+
277.	plus	or	minus	10	2	+	+	-	-	+
278.	all	or	none	(B)3001	(B)108	+	-	+	+	+
279.	now	or	never	(B)1314	(B)20	+	+	-	-	+
280.	more	or	less	(B)2216	(B)438	+	-	-	-	+
281.	happy	or	unhappy	AA	39	+	+	-	-	+
282.	like	or	dislike	AA	23	+	+	-	-	+
283.	many	or	few	(B)1030	(B)601	+	-	-	-	+
284.	win	or	lose	(B)55	(B)58	-	-	-	+	+
singular:										
285.	singular	and	plural	12	1	+	-	+	+	+
286.	one	or	two	(B)3297	(B)1412	+	-	-	-	+
287.	first	and	second	(B)1361	(B)375	+	+	-	-	+
288.	once	or	twice	AA	A	+	-	+	-	+
animate:										
289.	people	and	things	(B)902	(B)701	+	-	-	-	+
290.	men	and	machines	(B)781	(B)54	+	+	-	-	+

friendly:											
291.	for	or	against	(B)9488	(B)626	+	+	-	-	+	
292.	support	or	oppose	AA	A	+	-	-	-	+	
293.	accept	or	refuse	(B)189	(B)86	+	-	+	-	+	
solid:											
294.	land	and	sea	(B)218	(B)95	+	-	-	-	+	
295.	field	and	stream	(B)346	(B)62	+	-	+	-	+	
296.	(zu) Wasser	und (zu)	Land	(R)140	(R)55	+	-	-	-	-	
front:											
297.	front	and	back	(B)221	(B)967	-	-	-	+	+	
298.	front	and	rear	AA	A	+	-	-	-	+	
299.	bow	and	stern	A	36	+	-	+	-	+	
300.	fore	and	aft	11	4	+	-	-	+	+	
agentive:											
301.	speaker	and	hearer	25	6	+	-	-	-	+	
302.	actor	and	action	28	AA	-	-	-	-	+	
303.	subject	and	object	(B)242	(B)130	+	-	-	-	+	
304.	employer	and	employee	23	21	+	-	-	-	+	
305.	Fuchs	und	Henne	74	12	+	+	-	+	+	
power source:											
306.	bow	and	arrow	A	37	+	+	-	-	+	
307.	car	and	driver	AA	40	+	+	+	-	+	
308.	horse	and	carriage	AA	46	+	+	-	+	+	
309.	gin	and	tonic	9	4	+	+	-	+	+	
living:											
310.	living	or	dead	(B)194	(B)174	+	-	-	+	+	

	Word 1		Word 2	f 1	f 2	rule HF> LF	rule S> L	rule C> CC	rule FV> BV	rule sem.
311.	life	and	death	(B)715	(B)277	+	-	-	-	+
312.	live	or	die	(B)486	(B)192	+	-	-	+	+
313.	dead	or	alive	AA	A	+	+	-	-	-
314.	tot	oder	lebendig	512	429	+	+	-	-	-
315.	Leben	und	Sterben	8865	872	+	-	+	-	+
general:										
316.	form	and	substance	AA	A	+	+	-	-	+
317.	Form	und	Substanz	2406	293	+	+	-	-	+
318.	general	and	particular	AA	A	+	+	-	+	+
319.	abstract	and	concrete	6	15	-	-	+	-	+
320.	word	and	deed	AA	A	+	-	-	-	+
321.	knowledge	and	action	(B)145	(B)291	-	-	-	-	+
up vs. down										
322.	peak	and	valley	28	A	-	+	-	+	+
323.	up	and	down	(B)1896	(B)895	+	-	+	-	+
324.	hinauf	und	hinunter	634	204	+	+	-	-	+
325.	over	and	under	(B)1237	(B)707	+	-	-	+	+
326.	high	and	low	(B)498	(B)173	+	-	-	-	+
327.	above	and	below	(B)296	(B)145	+	-	+	-	+
328.	raise	and	lower	(B)179	(B)144	+	+	-	-	+
329.	top	and	bottom	AA	A	+	+	-	-	+
330.	upstairs	and	downstairs	9	20	-	-	+	-	+
331.	ascending	and	descending	29	49	-	-	+	-	+
332.	upper	and	lower	A	AA	-	-	+	-	+

333.	arms	and	legs	(B)121	(B)67	+	-	+	-	+
334.	auf	und	nieder	80944	1235	+	+	+	-	+
335.	hoch	und	tief	(R)120	(R)26	+	-	-	-	+
336.	auf	und	ab	80944	4559	+	-	-	-	+
337.	Berg	und	Tal	836	565	+	-	-	+	+
338.	drauf	und	dran	179	168	+	-	-	-	+
339.	oben	und	unten	2278	1048	+	-	-	-	+
340.	milk	and	honey	AA	A	+	+	-	+	-
341.	sugar	and	spice	AA	17	+	-	+	-	-
342.	pepper	and	onion	27	25	+	-	-	-	-
343.	fingers	and	toes	AA	35	+	-	-	+	+
344.	ball	and	chain	AA	A	+	-	-	-	-
345.	rough	and	tough	A	18	+	-	-	-	-
346.	hide	and	seek	A	AA	+	-	-	-	+
347.	space	and	time	(B)184	(B)1601	-	-	-	-	+
348.	bag	and	baggage	AA	11	+	+	-	-	-
349.	hit	or	miss	A	AA	-	-	-	-	-
350.	fair	and	square	116	90	+	-	+	-	-
351.	wax	and	wane	28	7	+	-	-	-	+
352.	betwixt	and	between	11	AA	-	-	-	-	-
353.	frisch	und	munter	729	312	+	+	-	+	-
354.	gesund	und	munter	818	312	+	-	-	-	-
355.	gut	und	teuer	6013	442	+	+	-	+	+
356.	süß	und	sauer	745	218	+	+	-	+	+
357.	Mord	und	Totschlag	264	21	+	+	-	-	-
358.	Schnee	und	Eis	326	238	+	-	-	-	-
359.	Sinn	und	Zweck	3655	3219	+	-	+	+	-
360.	Haut	und	Knochen	474	257	+	+	+	-	-
361.	Haut	und	Haar	(R)14	(R)9	+	-	-	-	-
362.	(ein) Herz	und	(eine) Seele	6985	3268	+	+	-	-	-

Word 1		Word 2	f 1	f 2	rule HF> LF	rule S> L	rule C> CC	rule FV> BV	rule sem.
363.	(unter) Dach	und	Fach (bringen)	255	192	+	-	-	-
364.	Milch	und	Honig	226	98	+	+	-	+
365.	gut	und	gerne	6013	3160	+	+	-	+
366.	Haus	und	Hof	2412	566	+	-	-	-
367.	Tisch	und	Sessel	1385	136	+	+	-	+
368.	Tisch	und	Bett	1385	764	+	-	-	+
369.	Obst	und	Gemüse	81	68	+	+	+	-
370.	Rock	und	Hose	285	53	+	+	-	-
371.	(bei) Wasser	und	Brot	2346	603	+	-	+	-
372.	Gold	und	Silber	993	671	+	+	-	-
373.	(in) Amt	und	Würden	(R)18	(R)1	+	+	+	-
374.	Haß	und	Neid	418	214	+	-	-	-
375.	Papier	und	Bleistift	1388	33	+	+	+	-
376.	Räuber	und	Gendarm	198	48	+	-	-	-
377.	(wie) Hund	und	Katz'	447	97	+	-	-	+
378.	Pferd	und	Reiter	1622	468	+	+	-	-
379.	Augen	und	Ohren (offenhalten)	6399	797	+	-	-	-
380.	mehr	oder	weniger	18549	3427	+	+	-	-
381.	hin	und	her	3575	2235	+	-	-	+
382.	heiß	und	kalt	1393	1347	+	-	-	+
383.	Anfang	und	Ende	1713	3183	-	-	-	-
384.	Leben	und	Tod	8865	3484	+	-	-	+
385.	Worte	und	Taten	7782	3853	+	-	-	+
386.	Mittel	und	Wege	2569	2357	+	-	-	+
387.	Salz	und	Pfeffer	240	50	+	+	-	-
388.	groß	und	klein	18813	7983	+	-	-	-

389.	Körper	und	Geist	1940	3242	-	-	-	-	-
390.	nose	and	throat	AA	A	+	-	+	-	-
391.	each	and	every	(B)877	(B)491	+	+	-	+	-
392.	fears	and	anxieties	AA	22	+	+	-	+	-
393.	first	and	foremost	AA	17	+	+	-	-	-
394.	hard	and	fast	(B)202	(B)78	+	-	-	-	-
395.	(with) intent	and	deliberation	23	6	+	+	+	-	-
396.	assault	and	robbery	22	9	+	-	+	-	-
397.	bar	and	restaurant	A	23	+	+	-	-	-
398.	brush	and	palette	A	1	+	+	-	-	-
399.	cuts	and	bruises	AA	17	+	+	+	-	-
400.	hope	and	pray	AA	A	+	-	+	-	-

Notes

1. The first, more theoretical part of this paper is based on a lecture given at the XIVth International Congress of Linguists, Berlin 1987 (Fenk-Oczlon 1987a). A report on the second, more empirically oriented part (Fenk-Oczlon 1988) was presented at the Symposium on Syntactic Iconicity, organized by Prof. Marge E. Landsberg at the 12th International Congress of Anthropological and Ethnological Sciences, Zagreb 1988. Correspondence address: Institut für Sprachwissenschaft, Universität Klagenfurt, Universitätsstrasse 65-67, A-9010 Klagenfurt, Austria.
2. The binomials numbered 1 to 63 in the Appendix come from Ross (1980). Numbers 63 to 229 and 229 to 353 were taken, with a few additions, from Malkiel (1959) and Cooper and Ross (1975), respectively. Of the remaining 47, the German ones are of my own finding, the English from Malkiel (1959) again.

References

- Abraham, R. D. (1950). Fixed order of coordinates. *Modern Language Journal* 34, 276-287.
- Allen, K. (1987). Hierarchies and the choice of left conjuncts (with particular attention to English). *Journal of Linguistics* 23, 51-77.
- Anttila, R. (1972). *An Introduction to Comparative and Historical Linguistics*. New York: Macmillan.
- Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure. *Journal of Experimental Psychology* 11 (4), 629-649.
- Bock, J. K. (1982). Toward a cognitive psychology of syntax: information processing contributions to sentence formulation. *Psychological Review* 89, 1-47.
- , and Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21, 47-67.
- Bradley, D. C. (1978). Computational distinctions of vocabulary type. Unpublished doctoral dissertation, MIT.
- Cooper, W. E., and Ross, J. R. (1975). World order. In *Papers from the Parasession on Functionalism*, R. E. Grossman, L. J. San, T. J. Vance, et al. (eds.), 63-111. Chicago: Chicago Linguistic Society.
- Cutler, A., and Cooper, W. E. (1978). Phoneme-monitoring in the context of different phonetic sequences. *Journal of Phonetics* 6, 221-225.
- Donegan, P. J., and Stampe, D. (1979). The study of natural phonology. In *Current Approaches to Phonological Theory*, D. Dinnsen (ed.), 126-173. Bloomington: Indiana University Press.
- Dressler, W. U. (1985). *Morphology: the Dynamics of Derivation*. Ann Arbor: Karoma.
- Edmondson, J. A. (1985). Biological foundations of language universals. In *Developmental Mechanisms of Language*, C. J. Bailey and R. Harris (eds.), 109-130. Oxford: Pergamon.
- Ertel, S. (1977). Where do the subjects of sentences come from? In *Sentence Production: Developments in Research and Theory*, 141-168, Hillsdale, N.J.: Erlbaum.
- Fenk, A. (1985). Is the reduction of subjective information (per unit of time) independent from the starting level? In *Cognition, Information Processing, and Motivation*, G. d'Ydewalle (ed.), 361-373. Amsterdam: North Holland.
- (1986). Informationale Beschränkungen der Wissenserweiterung? *Zeitschrift für experimentelle und angewandte Psychologie* 33, 208-253.
- , and Fenk, G. (1980). Konstanz im Kurzzeitgedächtnis — Konstanz im sprach-

- lichen Informationsfluß? *Zeitschrift für experimentelle und angewandte Psychologie* 27, 400-414.
- , and Fenk-Oczlon, G. (1987). Forschungslogische und semiotische Anmerkungen zur Natürlichkeitstheorie. Paper presented at the XIVth Congress of Linguists. (Abstracts: 452, Berlin [DDR]: Editorial Committee.)
- Fenk-Oczlon, G. (1983a). Ist die SVO-Wortfolge die 'natürlichste'? *Papiere zur Linguistik* 29, 23-32.
- (1983b). *Bedeutungseinheiten und sprachliche Segmentierung, Eine sprachvergleichende Untersuchung über kognitive Determinanten der Kernsatzlänge*. Tübingen: Narr.
- (1986). Morphologische Natürlichkeit und Frequenz. Paper presented at the 19th Annual Meeting of Societas Linguistica Europaea, Ohrid. (Abstract in *Markedness in Synchrony and Diachrony*.)
- (1987a). Frequenz und Wortfolge. Am Beispiel von 'freezes'. Paper presented at the XIVth International Congress of Linguists. (Abstracts: 452, Berlin [DDR]: Editorial Committee.)
- (1987b). Konstruktioneller Ikonismus und Ökonomieprinzip. Am Beispiel des Russischen Aspektsystems. Paper presented at the Österreichische Linguistentagung, Innsbruck, 24-26 October.
- (1987/1988). Prototypentheorie und Frequenz. *Klagenfurter Beiträge zur Sprachwissenschaft* 13/14, 138-150.
- (1988). Token-frequency and word-order. An empirical study. Paper presented at the 12th International Congress of Anthropological and Ethnological Sciences, Zagreb.
- , and Fenk A. (1985). The mean length of propositions is seven plus minus two syllables — but the position of languages within this range is not accidental. In *Cognition, Information Processing, and Motivation*, G. d'Ydewalle (ed.), 355-359. Amsterdam: North Holland.
- Fidelholtz, J. E. (1975). Word frequency and vowel reduction in English. *Papers of the Chicago Linguistic Society* 11, 200-214.
- Gordon, B. (1985). Subjective frequency and the lexical latency function: implications for mechanisms of lexical access. *Journal of Memory and Language* 24, 631-645.
- Haiman, J. (1983). Iconic and economic motivation. *Language* 59, 781-819.
- (1985). *Natural Syntax*. Cambridge: Cambridge University Press.
- Hasher, I., and Chromiak, W. (1977). The processing of frequency information: an automatic mechanism? *Journal of Verbal Learning and Verbal Behavior* 16, 173-184.
- , and Zacks, R. T. (1984). Automatic processing of fundamental information: the case of frequency of occurrence. *American Psychologist* 39, 1372-1388.
- Hofland, K., and Stig, I. (1982). *Word Frequencies in British and American English*. Bergen: Norwegian Computing Centre for the Humanities.
- Holenstein, E. (1978). *Von der Hintergebarkeit der Sprache*. Frankfurt/Main: Suhrkamp.
- Jakobson, R. (1971). Signe zéro. In *Roman Jakobson, Selected Writings II*, 211-219. The Hague: Mouton.
- Jespersen, O. (1942). *A Modern English Grammar, Part VI, Morphology*. Copenhagen: Ejnar Munksgaard.
- Josselson, H. H. (1953). *The Russian Word Count*. Detroit: Wayne State University Press.
- Kausler, D. H., and Puckett, J. M. (1980). Frequency judgments and correlated cognitive abilities in young and elderly adults. *Journal of Gerontology* 35, 376-382.
- Keenan, E. L. (1978). The syntax of subject-final languages. In *The Great Underlying Ground-Plans*, W. P. Lehmann (ed.), 267-327. Austin: University of Texas Press.
- Kelly, M. H., Bock, K. J., and Keil, F. C. (1986). Prototypicality in a linguistic context: effects on sentence structure. *Journal of Memory and Language* 25, 59-74.
- Kuno, S. (1979). On the interaction between syntactic rules and discourse principles. In

- Explorations in Linguistics: Papers in Honor of Kazuko Inoue*, G. Bedell, E. Kobayash and M. Muraki (eds.), 279–304. Tokyo: Kenkyusha.
- Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Mass.: MIT Press.
- Malkiel, Y. (1959). Studies in irreversible binomials. *Lingua* 8, 113–160.
- Mayerthaler, W. (1977). *Studien zur theoretischen und zur französischen Morphologie*. Tübingen: Niemeyer.
- (1981). *Morphologische Natürlichkeit*. Wiesbaden: Akademische Verlagsgesellschaft Athenaion. (English translation 1985: *Morphological Naturalness*. Ann Arbor: Karoma.)
- (1987). System-independent morphological naturalness. In *Leitmotifs in Natural Morphology*, W. U. Dressler, W. Mayerthaler, O. Panagl, and U. Wurzel. Amsterdam: Benjamins.
- Meier, H. (1964). *Deutsche Sprachstatistik*. Hildesheim: Georg Olms.
- Oldfield, R. C., and Wingfield, A. (1965). Response latency in naming objects. *Quarterly Journal of Experimental Psychology* 17, 273–281.
- Phillips, B. S. (1984). Word frequency and the actuation of sound change. *Language* 60 (2) 320–342.
- Pinker, S., and Birdsong, D. (1979). Speakers' sensitivity to rules of frozen word order. *Journal of Verbal Learning and Verbal Behavior* 18, 497–508.
- Ross, J. R. (1980). Ikonismus in der Phraseologie. *Zeitschrift für Semiotik* 2, 39–56.
- Roth, E., and Shoben, E. J. (1983). The effect of context on the structure of categories. *Cognitive Psychology* 15, 346–478.
- Ruoff, A. (1981). *Häufigkeitwörterbuch gesprochener Sprache*. Tübingen: Niemeyer.
- Sapir, E. (1929). A study in phonetic symbolism. *Journal of Experimental Psychology* 12 225–239.
- Sweet, H. (ed.) (1871/1872). *King Alfred's West-Saxon Version of Gregory's Pastoral Care*. EETS 45, 50. London: Kegan Paul. (Reprinted 1958.)
- Thorndike, E. L., and Lorge, I. (1944). *The Teacher's Word Book of 30,000 Words*. New York: Columbia University.
- Underwood, B. J. (1971). Recognition memory. In *Essays in Neobehaviorism*, H. H. Kendle and J. T. Spence (eds.). New York: Appleton Century Crofts.
- Witkowsky, S. R., and Brown, C. H. (1983). Marking reversals and cultural importance. *Language* 59, 569–582.
- Zipf, G. K. (1929). Relative frequency as a determinant of phonetic change. *Harvard Studies in Classical Philology* 40, 1–95.