Poplack, Shana. 2001. Variability, frequency and productivity in the irrealis domain of French. In Bybee, Joan & Hopper, Paul (eds.), Frequency and the Emergence of Linguistic Structure. Amsterdam: Benjamins. 405-428.

Variability, frequency, and productivity in the irrealis domain of French*

SHANA POPLACK

University of Ottawa

1. Introduction

In its view of grammar as anchored in concrete utterances, its quest for regularities in the repetition of such utterances in discourse, indeed in its characterization of grammar as ultimately "social" in nature (Hopper 1987), Emergent Grammar intersects in many largely unacknowledged ways with another seemingly very disparate framework for linguistic analysis—that of linguistic Variation Theory. Variation Theory (Labov 1969; Sankoff 1988a; Sankoff and Labov 1985), like Functional Linguistics, seeks to account for grammatical structure in discourse, paying particular attention to form-function asymmetry. The alternation of two or more variant forms in fulfilling a single function, so characteristic of discourse, is a major focus. The working hypothesis of Variation Theory is that within a given locus of variability, or *variable context*, each of two or more competing variants will occur at greater or lesser rates depending on the features that constitute the context. The expected proportion of each variant is the resultant of the combined contributions of the independent features defining its context.

The large corpora of spoken discourse which are the data of the variationist approach, coupled with its quantitative methodology, facilitate tests of alternative hypotheses as well. In this paper I make use of the data of natural conversation and the analytical tools of Variation Theory to shed light on the role of frequency in discourse, paying special attention to its relationship to productivity, ritualization (Haiman 1994) and the retention of archaic linguistic structures. The variationist apparatus is ideally suited to testing such developments. The multivariate analytical techniques of variable rule analysis (Sankoff 1988b) enable us to ascertain which are statistically significant and to disentangle their effects, if any, from those of the

other crosscutting linguistic and extra-linguistic factors simultaneously at play during variant production.

In what follows I investigate the extent to which usage data support the theory that type frequency is a major determinant of productivity and that token frequency may actually detract from productivity (Bybee 1985, 1995; Bybee and Thompson 1997; Langacker 1988). Illustrating with three sets of form-function asymmetries in the irrealis domain of spoken Canadian French, I show that the relationship between token frequency, type frequency and productivity are not as straightforward as a frequency-based approach would imply. In one of the contexts examined, the predictions of the model dovetail well with the facts of variable usage; in the second, the fit between data and theory is less good, but the patterning of variability suggests a possible explanation. In the third case, they fail to account for the data, raising, if not answering, questions about the properties of contexts hospitable to frequency effects.

2. The irrealis domain of French

By irrealis, I refer to the domain of imagined, projected, predicted or otherwise unreal situations or events, following, for example, Bybee (1998: 264). Most, if not all, such situations are conventionally assumed to be expressed in French by the invariant selection of one of the *subjunctive* (SUBJ) mood, as in (1), the inflected *future* (IF) "tense", as in (2), and *conditional* (COND) modality, as in (3).

(1) Elle a attendu que ses enfants ***seyent*** (SUBJ) assez grands pour aller travailler. (047/1939)[1]
'She waited until her children ***were*** old enough to go to work.'

(2) Dans bible ça dit, "Et les hommes ***auront*** (IF) la terre, ils ***feront*** (IF) la terre de [sic] quelque chose de bon". (001/622)
'In the Bible it says, "And man ***will have*** the earth, he ***will make*** of the earth something good".'

(3) Si c'***était*** (IMP) à mon choix, je les ***enlèverais*** (COND) de là. (025/657)
'If it ***were*** up to me, ***I'd get*** them out of there.'

In ordinary speech, however, the irrealis sector is host to considerable variability. This results in rampant form-function asymmetry, even in contexts in which a specific form is prescribed as obligatory. Thus both the indicative (IND) or the conditional may appear in contexts "requiring" the subjunctive, as in (4a and b), the periphrastic future (PF) has virtually replaced its inflected counterpart in all but a few future temporal reference contexts (5), and the conditional in *-rais* is ousting the

prescribed imperfect (IMP) in protases of hypothetical conditional complexes (6).

(4) a. Faut je lui ***dis*** (IND) c'est vrai. Faut je lui ***dise*** (SUBJ) c'est la vérité (064/356–69)
'I have to ***tell*** him it's true … I have to ***tell*** him it's the truth.'
b. Faut au moins que je ***serais*** (COND) bien obligée. (067/78)
'At I'***d*** least ***have to be*** really forced.'

(5) Ce soir, on ***va*** te ***ramener*** (PF) puis tu y ***alleras*** (IF) à soir à cinq heures. (071/584)
'Tonight, we'***re going to bring*** you ***back*** and you'***ll go*** there tonight at 5:00.'

(6) Si mon petit ***allait*** (IMP) à l'école là, s'il ***serait*** (COND) à l'école puis qu'il ***reviendrait*** (COND) puis qu'il ***dirait*** (COND), "Un professeur m'a tapé dans face là", il ***aurait*** affaire à moi. (037/437)
'If my kid ***went*** to school, if he ***would be*** at school, and he ***would come*** back, and he ***would say***, "a teacher slapped me across the face", he'***d have*** to deal with me.'

The facts illustrated in (1)–(6) furnish an interesting test of the relationship between frequency and productivity. The replacement of both the subjunctive by the indicative and the imperfect by the conditional are thoroughly non-standard, while the incursion of the periphrastic variant into the domain of the inflected future is generally considered colloquial. Moreover, historical research (LeBlanc 1999; LeBlanc and Poplack 1999a,b; Poplack 1992; Poplack and Turpin 1999) indicates that this variability is sufficiently longstanding and widespread as to have attracted the attention of the prescriptive and descriptive enterprise—each of these cases has been described, "explained" or denigrated by the French grammatical tradition from the 1600s through to the present. Indeed, the situation of the (Canadian) French irrealis domain qualifies as "emergent" par excellence, in the sense of Hopper (1987), insofar as it reflects centuries of prior, and as yet, unresolved variability.

The grammarian typically responds to such situations by attempting to factor out the variability, either by (1) ignoring it, (2) condemning the offending variant, or (3) attempting to redress the form-function asymmetry, typically by assigning to each form a preferred "reading" or function. In the latter effort they are abetted by the symbiotic relationship between members of the irrealis sector and the various domains of modality (especially epistemic modalities involving speaker commitment to the truth value of the proposition). This makes it possible to attribute the variability to such unobservables as speaker intent, and thereby explain it away. The abiding distaste of grammarians (and many linguists) for inherent variability, coupled with the important interpretive component they assign to speaker commitment and hearer inference, conspire in the observations—with which the literature is rife—that each

variant form fulfills a specific semantic task. Thus selection of the indicative in place of the subjunctive is often (and as we shall see, incorrectly) explained by the assertion that the speaker did not wish to commit herself to the reality, probability or truth value of the complement proposition (e.g., Grevisse 1986). Selection of the periphrastic future was associated by grammarians with proximity for centuries, but since the 1930s has been justified (again erroneously) by the opportunity it purportedly affords the speaker of envisaging the future eventuality in a more engaged, immediate, certain, committed and affective way than its inflected counterpart (e.g., Confais 1995; Deshaies and Laforge 1981; Fleischman 1982; Leeman-Bouix 1994). After an auspicious, though short-lived début (Maupas 1625), use of the conditional in protases has alternately been ignored or vilified, with remarkably little effect on its rapidly-increasing usage (Section 7).

What is the current role of these variant forms in discourse and why have they coexisted for so long? Does each perform the semantic task(s) claimed for it, or are some simply historical residue of erstwhile distinctions? If the latter, what accounts for their retention? Are they used indifferently with all eligible verbs or is the survival of some due rather to frequent repetition and sedimentation in grammar? In what follows I assess which account best fits the facts of spontaneous usage.

3. The network model of usage-based grammar

Bybee's (1985, 1988, 1995) network model claims that two factors are central in determining productivity, defined as the ability of a pattern to apply to novel items. The first is type frequency, or the number of different lexical items to which a particular pattern or construction is applicable. The more such items there are, the greater the likelihood that the pattern will also apply to novel items. This prediction is illustrated by Guillaume's (1927/1973, cited in Bybee 1995) demonstration that the widespread tendency of French children to generalize "first conjugation" verb morphology to other verbs is due, not to the greater token frequency of *-er* verbs, as is commonly assumed, but rather to their elevated type frequency (Table 1).

Table 1. *Token and type frequency of French verbs according to conjugation class (adapted from Guillaume 1927/73, cited in Bybee 1995)*

	Conjugation class	Token frequency	Type frequency
(I)	*Chanter*	36%	**76%**
(II)	*Finir*	6%	6%
(III)	*Vendre*	**57%**	18%

The other determinant of productivity, according to this model, is "schema strength", which is also based on type frequency. Schemas are generalizations about sets of words with similar patterns of semantic and phonological connections. If the defining properties of the schema are highly restrictive, it will not apply to many new forms. Thus the past tense formation pattern *string/strung* is not fully productive, since most English verbs do not meet the phonological description of *string*. Only an "open" schema, such as that of the English past tense *-ed*, can attain full productivity, since there are no restrictions on the forms to which it can apply (Bybee 1995: 430).

High token frequency, on the other hand, is not consistent with productivity. Bybee (1995: 434) explains this apparent contradiction as follows: Frequent forms can be learned by rote, without undergoing internal analysis or participating in schemas. This results in increased *lexical strength* (Bybee 1985) or *entrenchment* (Langacker 1987). Highly frequent entrenched words or phrases, according to these authors, tend to be stored unanalyzed , and are accessed more rapidly than their lexically weaker counterparts. Such items are also said to resist analogical levelling, resulting in the conservation of archaic structures. This "Conserving Effect" explains why high-frequency sequences (phonological, morphological or syntactic) are able to resist change toward newer more productive patterns (Bybee and Thompson 2000: 381).

The data of the French irrealis sector offer an appropriate testing ground for these claims, since as we shall see in Sections 5, 6, and 7, the three variables constituting it can be contrasted according to token frequency, type frequency and schema strength, among other factors, and these can be related to their productivity in actual usage.

4. Method and data

4.1 Data

The data I report on were all extracted from the *Corpus du français parlé à Ottawa-Hull* (Poplack 1989), a massive compendium of the highly informal conversation of a representative sample of 120 francophones native to the national capital region of Canada. This corpus of natural speech contains thousands of repetitions of each of the grammatical structures of interest to us, but no judgements, opinions or replies to queries concerning them. As such it is ideally suited to examining language as it is used unreflectingly, and to situating these uses in linguistic, social and historical context. At approximately 3.5 million words, the corpus is large enough to enable meaningful study of the kinds of frequency and usage questions that are the

focus of this volume. In each of the areas under study the same lexical types recur, used in the same conversations by the same speakers, though not necessarily at the same frequency levels, nor in the same morphological categories. In particular, verbs with suppletive morphology in the subjunctive and future sectors have regular conjugations in the imperfect. These circumstances enable us to disentangle purely lexical effects from those due to frequency of occurrence and morphological irregularity, an important check since these factors tend to be so highly correlated (Bybee and Thompson 2000; Poplack 1992). For example, an effect due to lexical identity can be expected to manifest itself by like behavior of the lexical item across variables; if morphological irregularity is determinative, effects should differ from variable to variable.

4.2 Method

Recognizing that the same linguistic "function" may at times be realized by different "forms", variationists seek to explain why one is actually chosen in a given context over another. The selection process is construed as resulting from the complex contribution of environmental factors, linguistic and social (plus a degree of inherent variability), which may conspire or conflict in the production of the form. This process is modelled by operationalizing hypotheses about selection constraints as *factors* in a multivariate analysis. Making use of a program for variable rule analysis (Rand and Sankoff 1990; Sankoff 1988b), I ascertain which of these factors contribute statistically significant effects to variant choice when all are considered simultaneously, as well as their relative magnitude with respect to each other.

The hypotheses I consider here relate to the contributions of token frequency, type frequency, schema strength and semanticity to productivity. As is standard in variationist research, it is first necessary to define the *variable context*, or locus of variability. The alternation between subjunctive and indicative, for example, is only relevant to specifically "subjunctive-selecting" contexts, since the reverse situation (subjunctive supplied in indicative contexts) is vanishingly rare (Poplack 1992). *Token frequency* is a count of the number of times a variant occurred in running speech (sometimes normalized as a percentage of all tokens being analyzed in a context). *Lexical identity* distinguishes the lexical types with which it co-occurred. *Lexical strength* refers to the proportion a given lexical type represents out of all lexical types in its cohort (e.g., the proportion the verb *falloir* represents of all matrix verbs used in the corpus). Due to discrepancies between the traditional accounts of these phenomena and the facts of usage detailed here, I distinguish *prescribed* type frequency or lexical schema (the class of lexical items to which the phenomenon is prescribed to apply) from *observed* type frequency or schema strength, i.e., the items (or contexts) in which a variant form actually occurred.

Productivity is a more elusive notion to operationalize. In the network model, it is defined as the ability to apply to novel items. By far the most important source of novel lexical material in Ottawa-Hull French is that originating from English, a phenomenon that has been thoroughly quantitatively studied (Poplack 1985, 1988; Poplack and Meechan 1998; Poplack *et al.* 1988). This work documents the strong tendency towards integration of borrowed material into the morphological and syntactic structures of French, due to which novel verb forms behave indistinguishably from the remainder of the verb paradigm. Similarities between established and nonce forms extend to the constraint hierarchies governing variability. The criterion of applicability to novel items is therefore less relevant here. For the purposes of this exercise, then, a variant will be considered *productive* if it fulfills the weaker requirement of occurring at a substantial, and relatively *homogeneous*, rate across all lexical items and contexts forming its domain.

An additional measure of productivity is *semanticity*, or the extent to which variant choice is associated with the nuances typically ascribed to it. If a morphological form performs a semantic task, it should occur freely with any lexical item which predicates a proposition consistent with that task, and this occurrence should be unhampered by restrictions imposed by lexical type, frequency and/or other morphosyntactic considerations. Productivity, as already observed by Bybee and Thompson (1997), is of course gradient. I class as *fully productive* the occurrence of a variant, in fulfillment of a specific semantic task, at comparable rates across (1) all lexical items and (2) all eligible subcontexts within its variable context. In what follows I examine how these categories interact with the expression of irrealis in three areas of French grammar.

5. The subjunctive

Standard French requires subjunctive morphology on every verb embedded under the set of "subjunctive-selecting" matrices.[2] Since at least 1698, however, it has been noted that the indicative sometimes appears in such contexts (Templery 1698), resulting in the type of variability illustrated in (4). Notwithstanding, unambiguous subjunctive morphology is currently quantitatively robust, appearing in a full 77% of verbs embedded under subjunctive-selecting matrices. But closer inspection reveals that its use is actually highly restricted and largely lexically determined. This is because the lexical types with which the subjunctive co-occurs, whether matrix or embedded, are highly restricted. Table 2 shows that a single verb---*falloir* 'must' —accounts for nearly 2/3 of the subjunctive-selecting matrices, and with a rate of 89%, displays a preternaturally high association with subjunctive morphology. This imbalance is compounded by two other verbs, *vouloir* 'want' and *aimer* 'like',

which are also frequent and display equally strong associations with the subjunctive. *Falloir*, *vouloir* and *aimer* together account for nearly 3/4 of all the subjunctive governors in the corpus. All the other matrix verbs are as likely to co-occur with the subjunctive as not (Poplack 1992).

Table 2. *Distribution of frequent matrix verbs according to lexical strength and propensity to select subjunctive morphology*

	Token frequency (N)	Lexical strength (% of all matrix verbs)	% Subjunctive morphology
Falloir	1,669	62	89
Vouloir	273	10	91
Aimer	86	3	67
Total N	**2,694**		

Moreover, though all French verbs are theoretically eligible to take the subjunctive so long as they are embedded under a subjunctive-selecting matrix, resulting in a prescribed lexical schema which is wide open, only four do so with any regularity (Table 3).

Table 3. *Distribution of frequent embedded verbs according to lexical strength and propensity to select subjunctive morphology*

	Token frequency (N)	Lexical strength (% of all embedded verbs)	% Subjunctive morphology
Être	659	24	65
Aller	390	14	87
Avoir	386	14	66
Faire	358	13	86
Total N	**2,694**		

As with the matrix verbs, the individual token frequency of these four embedded verb types is again extremely high, as is their lexical strength: together they represent 65% of all embedded verbs in the corpus. The next most frequently occurring verbs only account for an additional 3%, and there are only three of them (*prendre* 'take', *venir* 'come', *mettre* 'put'). The vast majority of embedded verbs each occurred one or two times, with associated lexical strengths of well under 1%.[3]

Although such overwhelming effects of lexical type are difficult to reconcile with the selection of subjunctive morphology to express modal nuances of doubt, non-assertion and the like, a significant proportion of French grammars has endorsed this

position for centuries. The variable rule analysis in Table 4 examines the conditioning of variant choice according to a number of factors which could contribute to a non-factual reading of the utterance, and compares their contribution to choice of the subjunctive with that of factors of a morphosyntactic nature. I then relate these to measures of frequency and ritualization, some of which are analyzed independently. Fully productive use of the subjunctive should be relatively impervious to the dictates of processing, priming, distance or purely morphosyntactic considerations.

(7) Factors considered in the analysis of variant choice in subjunctive-selecting contexts

Modal:
- Indicators of non-factual modality
- Structure of matrix clause
- "Semantic" class of matrix verb

Syntactic:
- Overtness of complementizer *que*
- Distance between matrix and embedded verb

Measures of frequency and ritualization:
- Lexical identity
- Token frequency
- Type frequency
- Conjugation class
- Priming

Table 4 displays the factors selected by the variable rule analysis as statistically significant to the probability that subjunctive morphology will be selected under "subjunctive-selecting" matrices other than *falloir*. As these do not entertain the same overriding lexical associations with the subjunctive, they are more likely to feature productive use. Yet no factors relevant to the putative meaning of the subjunctive were retained as significant by the stepwise multiple regression procedure incorporated in the variable rule program, indicating that semantic considerations do not play a role. (Selection of the factor incorporating the traditional categorization of matrices into "semantic" classes is an artifact of the inclusion in each of some (relatively) high-frequency verbs with high (or, as in the case of negated verbs of opinion, low) rates of subjunctive). The operative *independent* influences involve priming—the tense of the matrix verb tends to be copied to the embedded verb, processing—presence of the (variably deleted) complementizer *que* favors selection of subjunctive morphology, and a combination of morphological suppletion and high token frequency of the embedded verb. These results are entirely consistent with the strong lexical effects described above.

Table 4. *Variable rule analysis of the contribution of factors selected as significant to the choice of subjunctive morphology in embedded clauses governed by verbs other than* falloir *(from Poplack 1992)*

Overall tendency:	.526
"Semantic" class	
Volitive	.77
Emotive	.66
Opinion	.09
Tense of matrix verb	
Imperfect	.65
Present	.51
Passé composé	.42
Periphrastic Future	.38
Conditional	.25
Presence of *que*	
Overt	.52
Absent	.47
Morphological form/frequency of embedded verb	
Suppletive/frequent	.56
Regular/rare	.36

Summarizing this section, although the token frequency of the standard subjunctive variant is elevated, virtually all its uses are concentrated among a handful of highly favoring matrix verbs collocated with a small cohort of frequent and irregular embedded verbs. Outside of these few contexts, in which its use has become ritualized, selection of the subjunctive is very rare. Thus despite an open prescribed schema (admitting all verbs embedded under the class of subjunctive-selecting matrices), the *observed* lexical strength of this variant is highly restricted. I have also ruled out semanticity as a contributor to variant choice. I therefore categorize it as low to nil in terms of productivity. The French subjunctive clearly exemplifies the Conserving Effect of high token frequency, coupled with very low type frequency. Bybee and Thompson (1997) have explained this by observing, correctly, in my opinion, that the "high-frequency expressions have maintained their traditional (subjunctive) form despite general changes which allow the construction of sentences with indicative forms in comparable, but less frequent, contexts".

6. The future

Consider next the expression of future temporal reference, for which three morphological variants have been competing since the thirteenth century: the periphrastic (PF), inflected (IF) and futurate present (P) forms, illustrated in (8a–c).

(8) a. Bien ***demain***, tu ***vas aller*** (PF) au Bingo, tu ***vas gagner*** (PF). (065/2301)
'***Tomorrow*** you'***re going*** to go to Bingo and you'***re going to win***.'
b. J'ai dit, "Laisse faire, on ***ira*** (IF) à messe ***demain*** matin". (070/686)
'I said, "Forget it, we'***ll go*** to Mass tomorrow morning".'
c. Il dit, "J'y ***vas*** (P) ***demain*** matin chez vous". (119/861)
'He says, "I'***m going*** to your house ***tomorrow***".'

As in the case of the subjunctive, variants are generally claimed to be chosen according to distinctions in the way the speaker envisages the future eventuality. Typically, however, there is little consensus as to what those distinctions are nor which variants are capable of expressing them. From 1753 to 1935 grammarians were virtually unanimous in ascribing to the periphrastic form a reading of proximity (belied by the usage examples in (7)); subsequently the epistemic readings cited in Section 2 above gain favor. Aside from these semantic associations, the prescribed lexical schema for the future variants is totally open: each is (theoretically) equally felicitous with any verb in the language. Nonetheless, Table 5 shows that variant distribution is once again highly skewed: The periphrastic form features the highest token frequency by far, accounting by itself for nearly 3/4 of all future temporal reference expression. The inflected form (traditionally considered the default variant) appears from Table 6 to occur no more than 20% of the time; as we shall see below, it is in fact a good deal less frequent than this, at least in productive uses.

Table 5. *Distribution of major variant expressions of future temporal reference*

	%	N
Periphrastic future	73	2,627
Inflected future	20	725
Futurate present	7	242
Total		3,594

I again investigate the conditioning of variant occurrence, using factors meant, as previously, to capture lexico-semantic and morphosyntactic properties of the contexts in which they appear. These are listed in (9).

(9) Factors considered in the analysis of variant choice in future temporal reference contexts

Modal:
- Contingency
- Temporal distance
- Imminence
- Polarity
- Stativity
- Person and number of subject

Discourse:
- Adverbial specification

Measures of frequency and ritualization:
- Lexical identity
- Token frequency
- Type frequency
- Conjugation class
- Priming

Table 6 displays the results of a variable rule analysis of the contribution of the above factors to the choice of the inflected future. Only three of these factors con-

Table 6. *Variable rule analysis of the contribution of factors selected as significant to the choice of inflected (IF) morphology in future temporal reference contexts* (from Poplack and Turpin 1999)

	IF
Overall tendency:	.145
Total N (/variant)	725
Type of adverbial specification	
Non-specific	.85
No adverbial	.47
Specific	.37
Grammatical person	
Formal *vous*	.81
Other	.49
Polarity	
Negative	.99
Affirmative	.36

tribute statistically significant effects, and as in the case of the subjunctive, none is relevant to the meanings generally ascribed to this variant.

These include a functional effect of adverbial specification, promoting the inflected variant in the context of a non-specific time adverbial, as in (10), thereby avoiding the habitual reading that would result from use of the futurate present.

(10) Tôt ou tard ils ***reviendront*** (IF). (023/659)
'Sooner or later they'***ll come back***.'

A second factor, meant to test the purported association of the periphrastic variant with the more "subjective" 1st person subjects, shows instead that the inflected future is favored with the (rarely used) formal pronoun of address *vous*. This is consistent with the strong association of this variant with frozen and formulaic expressions, such as those in (11).

(11) a. Dieu a toujours dit, 'Aide-toi et le ciel t'***aidera*** (IF).' (113/855)
'God has always said, "Heaven helps those who help themselves".'
b. C'est comme qu'ils disent, hein? 'Qui a bu ***boira*** (IF).' (101/1315)
'It's like they say, eh? "A leopard can't change its spots".'

But by far the greatest effect on variant choice is contributed by negation of the future eventuality. The inflected future is overwhelmingly preferred (and the other variants correspondingly eschewed) in negative contexts, as in (12).

(12) Dire que dans quatre cents ans d'ici il ***va avoir*** (PF) encore des Asselin, puis ils ***vont*** encore ***parler*** (PF) français. Qu'ils ***parleront*** (IF) *pas* l'anglais. (004/3611)
'To think that 400 years from now, there ***are*** still ***going to be*** Asselins, and they'***re*** still ***going to speak*** French. That they ***won't speak*** English.'

This spectacular contribution of negation to the selection of the inflected variant has been amply attested in empirical analyses of usage (e.g., Lesage 1991; Sundell 1991), especially oral (e.g., Chevalier 1994; Deshaies and Laforge 1981, Emirkanian and Sankoff 1985; Lorenz 1989; Zimmer 1994). It remains largely unacknowledged in other contemporary studies of French (with the notable exceptions of Franckel 1984 and Vet 1993). Nor was the negative effect noted in a single one of the 130 prescriptive and descriptive grammars of French dating from the 1600s to the present that we have consulted (Poplack *et al.* in preparation). It comes as no surprise that the retention of the inflected future in this context has as yet received no convincing explanation. Whatever the reason, negative contexts are now the only loci in which the inflected future variant is used productively in spoken Canadian French.

What of the lexical effect? Given that the same frequent, morphologically irregular verbs that retained the subjunctive also have irregular inflected futures (*serai* < *être, irai* < *aller*, *aurai* < *avoir; ferai* < *faire*), they should exert a conserving effect here as well. But when the effect of negation is factored out, as in the middle portion of Figure 1 (Section 8), no association between lexical form and variant choice can be detected. On the contrary, all verbs, when negated, are overwhelmingly conjugated with the inflected variant, regardless of lexical type. Verbs in the affirmative, on the other hand, occur with it only rarely, again regardless of token or type frequency, conjugation class or purely lexical considerations. The only exception involves the few uses that are formal in nature or entrenched in conventionalized expressions. When these are removed from the data, truly productive uses of the inflected form (as in 8b) fall to fewer than 6% of the data. Even these are slowly receding, since they are preferred by speakers over 70 (Poplack and Turpin 1999).

Summarizing, the prescribed lexical schema for the inflected future is wide open, since there are no restrictions on the lexical items with which it can co-occur. Examination of usage confirms that variant selection is indifferent to token frequency, lexical identity, lexical strength, conjugation class or any other property relating to the verb. Nonetheless, its observed schema is highly restricted—selection of the inflected variant is basically limited to negative and some formulaic contexts, which themselves account for only 10% of the future temporal reference data. This explains the very low overall token frequency of this variant. Its restricted productivity is corroborated by the results of the multivariate analysis (Table 6), which reveals that the choice between competing variants is not made to effect the semantic tasks usually ascribed to it. Retention of the inflected future does not appear to be motivated by operation of the frequency effects predicted by the network model.

7. The conditional

The protasis of hypothetical *si-* complexes represents another locus of long-term variability. The requirement that the standard imperfect (or pluperfect) indicative be employed in this context is often flouted in favor of the non-standard conditional, as in (6). This usage, first recorded by grammarians in 1625 (when the conditional was last considered to "go well" with *si* [Maupas 1625]), has continued to flourish: its token frequency now exceeds that of the standard imperfect in Ottawa-Hull French, as can be seen in Table 7.

What are the factors that encourage retention of the imperfect? Unlike the other two domains of the irrealis system, where form-function asymmetry has been

Table 7. *Distribution of variants in the protases of hypothetical si complexes*

	%	N
Conditional	53	766
Imperfect	47	687
Total		1,453

persistently—if controversially—ascribed to modal or semantic differences, no such explanation has ever been offered for the choice between imperfect and conditional in the protasis. Nonetheless, drawing on observations of actual usage, LeBlanc (1999) and LeBlanc and Poplack (1999a,b) examined the conditional complexes according to traditional (Latin) semantic classifications invoking the likelihood that the condition in the apodosis would be realized. We distinguished possible or potential conditions, as in (13), from impossible or counterfactual ones, as in (14). Variable rule analysis compares, as previously, their contributions to variant choice with those of factors of a syntactic nature, which can then be related to the effects of frequency and ritualization. These are listed in (15).

(13) Si votre père ***serait*** (COND) mort puis la petite vous ***dirait*** (COND) quelque chose comme ça, que c'est vous *feriez*? (018/486)
'If your father ***would be*** dead and the kid ***would tell*** you something like that, what ***would*** you ***do***?'

(14) Si j'***avais*** (IMP) des jeunes puis ***fallait*** (IMP) je travaille là, je pense que ça l'*arriverait*. (114/1334)
'If I ***had*** youngsters and I ***had*** to work, I think that that ***would happen***.'

Aside from a priming effect, whereby the tense/mood of one verb is copied to successive verbs in coordinate protases, exemplified in (6) above, Table 8 (p. 420) shows that *none* of the factors hypothesized to play a role in the retention of the imperfect was selected as significant by the stepwise multiple regression procedure. The one notable exception, in striking contrast to the behavior of the subjunctive and the inflected future variants, as well as to prescribed usage, involves the semantic value of the condition. If the realization of the condition is viewed as possible, as in (13), the conditional is favored in the protasis; if it is viewed as counterfactual, as in (14), the imperfect is more likely. This effect, a replica of that operating in Latin, has not been attested in prescriptive or descriptive grammars of French.

(15) Factors considered in the analysis of variant choice in the protases of conditional *si* complexes.

Modal:
- Semantic reading
- Polarity of protasis and apodosis
- Person of subject (protasis)

Syntactic:
- Embedding of conditional complex
- Linear order of protasis and apodosis
- Distance between protasis and apodosis

Measures of frequency and ritualization:
- Lexical identity
- Token frequency
- Type frequency
- Conjugation class
- Priming

Table 8. *Variable rule analysis of the contribution of factors selected as significant to the choice of IMPERFECT morphology in the protases of si complexes (from Leblanc 1999)*

Overall tendency:	.473	
Total N:	1,406	
Semantic reading:		N
Counterfactual	.55	379
Potential	.43	361

Most cases of linguistic variability result in neutralization of semantic distinctions in well-defined discourse contexts. Conditional complexes represent one of the comparatively rare cases where inherent variability *introduces* a semantic distinction (detailed in Poplack and LeBlanc 1999a,b). This extends speakers' highly productive usage of conditional morphology in its exponentially more frequent main-clause uses so as to align form with function in protases of conditional *si* complexes as well.

What of the lexical effect? The disparities in distribution of lexical types familiar from the other variables are equally operative in conditional complexes. For the most part they involve the same irregular verbs that we have seen to have played

such an important role in choice of the subjunctive, although the suppletive morphology characteristic of the subjunctive and future is not at issue with the imperfect, which derives from an infinitival base (e.g., *étais* < *être, allais* < *aller*, *avais* < *avoir, faisais* < *faire*). The two lexical types *avoir* 'have' and *être* 'be', for example, constitute a full third of verbs occurring in protases. As with the inflected future, however, neither these, nor a second tier of 10 somewhat less frequent verbs accounting for another third of the data, display any particular association with either of the variant forms. In particular, none of the frequent forms displays a greater propensity to preserve the archaic (in this context) imperfect than its less frequent counterparts. On the contrary, Figure 1 shows that all lexical types feature the same rate of imperfect usage, regardless of token frequency, type frequency or schema strength.

Summarizing this section, despite a restricted variable context involving only protases of conditional complexes, the prescribed lexical schema for the imperfect is again wide open in terms of permissible lexical hosts. And in contrast to the subjunctive and the inflected future, the imperfect does in fact occur freely not only across all lexical items, but also in all contexts constituting its domain; it is the preferred variant in counterfactual conditions. This despite the fact that it is rapidly losing ground in protases. Even its apparently moderate rate in Table 7 is artificially inflated by inclusion in the calculation of all 120 speakers constituting the *Corpus du français parlé à Ottawa-Hull*. When speakers are distinguished according to age (LeBlanc 1999; LeBlanc and Poplack 1999a,b), the variants are seen to be involved in vigorous change in progress, with the imperfect again retained mainly in the speech of those currently aged over 70.

8. Summary of effects

8.1 The effect of token frequency on variant choice

Bybee and Thompson (2000) have suggested that high token frequency leads to two types of outcome: 1) reductive changes including loss of internal structure and semantic bleaching, and 2) conservation of older forms in high-frequency contexts while other forms prevail in comparable, but less frequent contexts. According to these authors, this is because the more frequently a form is used, the more its representation is strengthened, making it easier to access in the same form the next time, and resulting in a "lexically arbitrary residue of formerly productive patterns" (Bybee and Thompson 2000: 384).

Figure 1 summarizes the effects of token frequency, or lexical strength, and lexical identity on choice of subjunctive, inflected future and imperfect morphology in

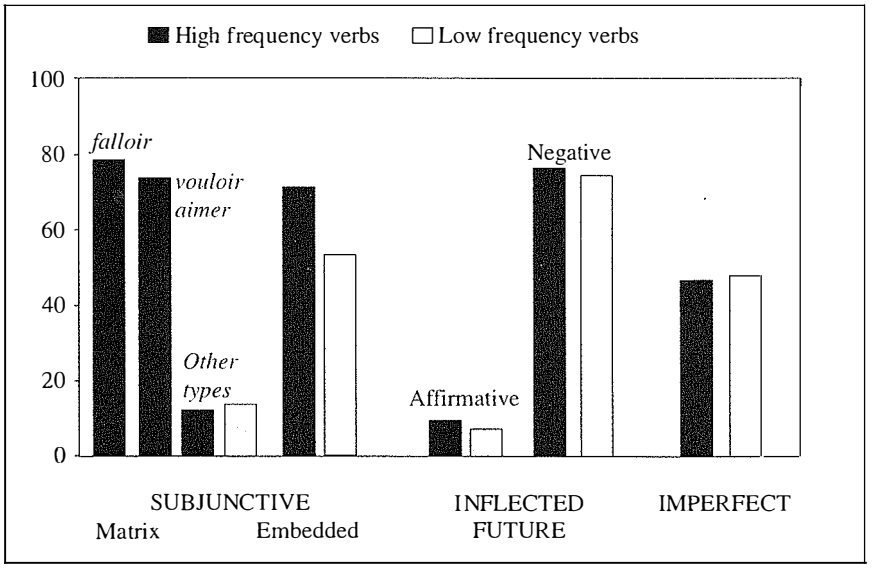

Figure 1. *The effect of token frequency on variant choice in three domains of French grammar. Vertical axis measures percentage of subjunctive, inflected future or imperfect variants, with frequency class indicated by shading*

each of their respective variable contexts. As noted earlier, high token frequency clearly plays a major role in the retention of the French subjunctive in the expected direction, as does lexical identity, since the subjunctive prevails specifically with three frequent matrix and four frequent embedded verbs; it is far rarer in the context of less frequent verbs. In neither case is it selected to perform any particular semantic work, consistent with the reductive bleaching effects also associated with high token frequency. In the case of the inflected future and imperfect variants, however, the very same verbs whose lexical strength contributed so much to the conserving and reduction effects operating on the subjunctive, play no role in variant choice. This is true whether they are considered individually, or aggregated according to frequency, as in Figure 1.

8.2 The relationship between type frequency and productivity

As noted above, type frequency should determine degree of productivity, since the more lexical items co-occur with a pattern or construction, the less likely it will be

associated with any one of them in particular (Bybee and Thompson 1997: 7). High type frequency is also said to ensure high token frequency, which in turn increases lexical strength, making the construction more accessible for further uses.

Table 9 depicts the relationship between token frequency, prescribed and observed type frequency and productivity in our data.

Despite restricted domains of application in the case of the subjunctive and the imperfect, every one of the morphological forms studied here has a high prescribed type frequency, or open schema, insofar as it is free to co-occur with any verb in the language. With neither the inflected future nor the conditional, however, is token frequency correspondingly high. Nor does observed type frequency necessarily correlate with prescribed type frequency, as illustrated by the behavior of the subjunctive. Neither the subjunctive nor the inflected future were found to convey, in usage, any of the meanings usually attributed to them. While the inflected future is virtually restricted to negative contexts, it carries no negative connotations outside of that context. The imperfect, on the other hand, does perform a semantic task in protases of hypothetical *si* complexes, albeit one which to my knowledge has never been reported. And while the low productivity of the subjunctive is perhaps predictable from its low type frequency, the same can certainly not be said of the other two variants. Thus, insofar as these data are concerned at least, the relationship among token frequency, type frequency and productivity is not straightforward.

Table 9. *The relationship between token frequency, type frequency and productivity in the French irrealis domain*

Morphological variant:	Subjunctive	Inflected Future	Imperfect
Variable context:	"Subjunctive-selecting matrices	All future temporal reference	Protasis of hypothetical *si*-clause
Prescribed:			
Type frequency	Very high	Very high	Very high
Observed:			
Frequency (token)	Very high	Very low	Low and decreasing
Frequency (type)	Very low	Very high	Very high
Semanticity	–	±	+
Productivity	Low and unchanging	Restricted and decreasing slowly	Restricted and decreasing rapidly

9. Discussion

In this paper I have brought a variationist perspective to bear on the tenets of frequency-based models that type frequency and schema strength result in increased productivity. I have argued that the variationist framework provides a particularly apt test of these claims, for a number of reasons. First the variationist focus on natural speech furnishes a representative data base on unreflecting usage, quantitatively important enough to allow meaningful analysis of a variety of frequency effects. Second, since the same verbs, uttered by the same speakers during the same interactions, figure in each of the analyses, we can control for the (highly correlated) relationship between lexical identity, morphological irregularity and frequency. Regardless of which predominates, its effects should be parallel across all variables.

Three areas of the irrealis domain in which two or more forms compete for the expression of a single function were selected for analysis. Inherent variability has been attested here for centuries, and in each case, token frequencies are currently highly skewed in favor of one of the variants, prompting us to seek the factors responsible for retention of the other(s). This involved operationalizing various proposals for variant choice as factors in a multivariate analysis, including the predictions of frequency-based models, and testing them in usage. The analytical tools of Variation Theory enable us not only to detect frequency and lexical effects, where operative, in a large corpus of conversational data, but also to distinguish their effects from the competing contributions of other (linguistic and social) factors.

In the case of the subjunctive, results show that a frequency-based analysis provides a good account of the facts (despite the existence of a small cohort of highly frequent matrix verbs which *disfavor* the subjunctive). For the other two variables, however, neither type frequency, token frequency nor schema strength is fully predictive of productivity, even using the weaker definition I have proposed here. In the case of the future, the network model would predict that some verbs—high frequency, morphologically irregular or lexically strong—should retain the older inflected variant, even in affirmative contexts. This prediction is not borne out. Instead we find a situation where variant usage has become highly differentiated according to polarity. However, this unexpected distribution, the motivation for which is still unclear, may well be implicated in the lack of frequency effect. The network model requires that the (putative) replacement form have the same meaning or function as the pre-existing form, as with the English past tense forms *-ung* and *-ed*. In the case of the French future, erstwhile variant expressions of a single function have now become functionally differentiated.[4] This entails selection of the form associated with the function, and the magnitude of this effect on the inflected and periphrastic futures may have obliterated any effect due to lexical identity or frequency.

What of the imperfect? As the clearly archaic and receding form in protases of conditional complexes, it too should have been retained on high-frequency verbs, according to the network model. This is not the case. Unlike the future, however, here the moderate semantic distinction fulfilled by the competing variants is compatible, within the variationist model, with the frequentist claim for lexically-based differentiation in rates.[5] Variable rule analysis could thus detect a frequency effect if one were operative. At this stage I can only speculate on why none appears. I noted earlier that the morphology of the imperfect is regular, in contrast to the other two cases, where frequent verbs are suppletive. This suggests that high type frequency may not operate independently of morphological (ir)regularity in the determination of productivity.

Another possibility is that frequency effects are most visible when the original variable context is lexically determined. We have seen that, as a function of the interaction of frequency with morphological effects, the list of subjunctive-selecting matrices is narrowing to a (much) smaller list. Similar results obtain with the ongoing regularization of the auxiliary *être* 'be' to *avoir* 'have' affecting the perfect tenses of the lexically-determined set of "*être*-verbs". Here as well, type frequency is clearly a major determinant of the uneven retention of the receding *être* (Willis 1999, 2000; see also Smith this volume). Of course a frequency-based explanation also accounts nicely for the verbs embedded under subjunctive-selecting matrices, which have no list basis. Nevertheless this is a promising line of study, which we are currently pursuing further. It is hoped that future research taking account of the kinds of complex patterning I have demonstrated to be operating here will help distinguish the preferred domains of application of frequency effects.

Notes

* The work reported here forms part of the project on Variation, Prescription and Praxis: Contact and Evolution of Grammatical Systems, generously funded by the Social Science and Humanities Research Council of Canada. Their support is gratefully acknowledged here. Carmen LeBlanc and Lauren Willis collaborated in constituting the *Historical French Grammars Resource* (Poplack *et al.* in preparation) from which the historical information reported here is drawn.

1. Codes refer to speaker number and line number in the *Corpus du français parlé à Ottawa-Hull* (Poplack 1989). Examples are reproduced verbatim from speaker utterances.
2. Here I report on the set of (approximately 67) verbal matrices which co-occurred at least once with the subjunctive, considering embedded verbs displaying unambiguous subjunctive morphology only. For discussion of subjunctive usage with non-verbal matrices, see Poplack (1997).
3. The same skewed distribution is observed with non-verbal subjunctive-selecting matrices, such as *avant que* 'before', *pour que* 'so that' (Poplack 1997).
4. Whether they are also semantically distinct is debatable; at this stage the inflected future conveys no nuance of negation except when collocated with a negative particle.

5. The contribution of the lexical item combined with the contribution of the meaning would yield the probability of variant usage with each lexical item.

References

Bybee, J. 1985. *Morphology: A Study of the Relation Between Meaning and Form*. Amsterdam and Philadelphia: John Benjamins.

Bybee, J. 1988. "Morphology as lexical organization". In *Theoretical Morphology*. M. Hammond and M. Noonan (eds.). San Diego, CA: Academic Press.

Bybee, J. 1995. "Regular morphology and the lexicon". *Language and Cognitive Processes* 10(5): 425–55.

Bybee, J. 1998. "Irrealis as a grammatical category". *Anthropological Linguistics* 40(2): 257–71.

Bybee, J., and Thompson, S. 2000. "Three frequency effects in syntax". *Berkeley Linguistics Society* 23: 378–88.

Chevalier, G. 1994. "L'emploi des formes du futur dans le parler acadien du Sud-Est du Nouveau-Brunswick". Paper read at Colloque Les Acadiens et leurs langues, at Université de Moncton.

Confais, J.-P. 1995. *Temps Mode Aspect: Les Approches des Morphèmes Verbaux et Leurs Problèmes à L'Exemple du Français et de L'Allemand*. Toulouse: Presses Universitaires du Mirail.

Deshaies, D., and Laforge, E. 1981. "Le futur simple et le futur proche dans le français parlé dans la ville de Québec". *Langues et Linguistique* 7: 23–37.

Emirkanian, L., and Sankoff, D. 1985. "Le futur simple et le futur périphrastique dans le français parlé". In *Les Tendances Dynamiques du Français Parlé à Montréal*, M. Lemieux and H. Cedergren (eds.), 189–204. Québec: Gouvernement du Québec.

Fleischman, S. 1982. *The Future in Thought and Language: Diachronic Evidence from Romance*. Cambridge: Cambridge University Press.

Franckel, J.-J. 1984. "Futur 'simple' et futur 'proche' ". *Le Français dans le Monde* 182: 65–70.

Grevisse, M. 1986. *Le Bon Usage*. 12th edition ed. Paris: Duculot.

Guillaume, P. 1927/1973. "The development of formal elements in the child's speech". In *Studies of Child Language Development*, C.A. Ferguson and D.J. Slobin (eds.), 240–51. New York, NY: Holt, Rienhart and Winston.

Haiman, J. 1994. "Ritualization and the development of language". In *Perspectives on Grammaticalization*, W. Pagliuca (ed.), 3–28. Amsterdam and Philadelphia: John Benjamins.

Hopper, P.J. 1987. "Emergent grammar". *Berkeley Linguistics Society* 13: 139–57.

Labov, W. 1969. "Contraction, deletion, and inherent variability of the English copula". *Language* 45(4): 715–62.

Langacker, R.W. 1987. *Foundations of Cognitive Grammar, Volume 1: Theoretical Prerequisites*. Stanford, CA: Stanford University Press.

Langacker, R.W. 1988. "A usage-based model". In *Topics in Cognitive Linguistics*, B. Rudzka-Ostyn (ed.), 127–61. Amsterdam and Philadelphia: John Benjamins.

LeBlanc, C. 1999. *Du Conditionnel dans les Propositions Hypothéques en si: Cet Intrus*. M.A. thesis, University of Ottawa.

LeBlanc, C.L., and Poplack, S. 1999a. "Conditions sur le conditionnel". Paper read at LSRL, at University of Michigan.

LeBlanc, C.L., and Poplack, S. 1999b. "Prescription vs. praxis: Conditional usage in French hypothetical si-clauses". Paper read at NWAVE 28, at Toronto.

Leeman-Bouix, D. 1994. *Grammaire du Verbe Français: Des Formes au Sens: Modes, Aspects, Temps, Auxiliaires*. Paris: Nathan.

Lesage, R. 1991. "Notes sur l'emploi du présent à valeur de futur dans les quotidiens Québécois". *Revue Québécoise de Linguistique Théorique et Appliquée* 10 (3): 117–31.

Lorenz, B. 1989. *Die Konkurrenz Zwischen dem* Futur Simple *und dem* Futur Périphrastique *im Gesprochenen Französisch der Gegenwart* Vol. 2. Münster: Kleinheinrich.

Maupas, C. 1625. *Grammaire et Syntaxe Françcoise*. Rouen: J. Cailloué.

Poplack, S. 1985. "Contrasting patterns of code-switching in two communities". In *Methods V: Papers from the Fifth International Conference on Methods in Dialectology*, H. Warkentyne (ed.), 363–86. Victoria: University of Victoria Press.

Poplack, S. 1988. "Language status and language accommodation along a linguistic border". In *Language Spread and Language Policy: Issues, Implications and Case Studies, GURT 87*, P. Lowenberg (ed.), 90–118. Washington, D.C.: Georgetown University Press.

Poplack, S. 1989. "The care and handling of a megacorpus: The Ottawa-Hull French Project". In *Language Change and Variation*, R. Fasold and D. Schiffrin (eds.), 411–51. Amsterdam and Philadelphia: John Benjamins.

Poplack, S. 1992. "The inherent variability of the French subjunctive". In *Theoretical Studies in Romance Linguistics*, C. Lauefer and T.A. Morgan (eds.), 235–63. Amsterdam and Philadelphia: John Benjamins.

Poplack, S. 1997. "The sociolinguistic dynamics of apparent convergence". In *Towards a Social Science of Language, Volume 2: Social Interaction and Discourse Structures*, G. Guy, C. Feagin, D. Schiffrin and J. Baugh (eds.), 285–309. Amsterdam and Philadelphia: John Benjamins.

Poplack, S., LeBlanc, C., and Willis, L. in preparation. The Historical French Grammars Resource. Manuscript, University of Ottawa.

Poplack, S., and Meechan, M., (eds.) 1998. *Instant Loans, Easy Conditions: The Productivity of Bilingual Borrowing*; *Special Issue, International Journal of Bilingualism*. Vol. 2 (2). London: Kingston Press.

Poplack, S., Sankoff, D., and Miller, C. 1988. "The social correlates and linguistic processes of lexical borrowing and assimilation". *Linguistics* 26(1): 47–104.

Poplack, S., and Turpin, D. 1999. "Does the FUTUR have a future in (Canadian) French?". *Probus* 11: 133–64.

Rand, D. and Sankoff, D. 1990. GoldVarb. A variable rule application for the Macintosh. Centre de recherches mathématiques, Université de Montréal, Montreal, Canada.

Sankoff, D. 1988a. "Sociolinguistics and syntactic variation". In *Linguistics: The Cambridge Survey*, F.J. Newmeyer (ed.), 140–61. Cambridge: Cambridge University Press.

Sankoff, D. 1988b. "Variable rules". In *Sociolinguistics: An International Handbook of the Science of Language and Society*, U. Ammon, N. Dittmar and K.J. Mattheier (eds.), 140–61. Berlin: Walter de Gruyter.

Sankoff, G., and Labov, W. 1985. "Variation theory". Paper presented at NWAVE 14, at Georgetown University.

Sundell. L.-G. 1991. *Le Temps Futur en Français Moderne*. Uppsala: Textgruppen i Uppsala AB.

Templery, J.D. 1698. *Remarques sur la Langue Françoise*. Paris: Martin et George Jouyenei.

Vet, C. 1993. "Conditions d'emploi et interprétation des temps futurs du français". *Verbum* 4: 71–84.

Willis, L. 1999. "Être ou ne plus être: Auxiliary alternation in Ottawa-Hull French". Paper read at NWAVE 28, at Toronto.

Willis, L. 2000. "Être ou ne plus être: Auxiliary alternation in Ottawa-Hull French". M.A. Thesis, University of Ottawa.

Zimmer, D. 1994. "'Ça va tu marcher, ça marchera tu pas, je le sais pas' (71: 15): Le futur simple et le futur périphrastique dans le français parlé à Montréal". *Langues et Linguistique* 20: 213–26.

In the series TYPOLOGICAL STUDIES IN LANGUAGE (TSL) the following titles have been published thus far:

18. HAIMAN, John & Sandra A. THOMPSON (eds): *Clause Combining in Grammar and Discourse.* 1988.
19. TRAUGOTT, Elizabeth C. and Bernd HEINE (eds): *Approaches to Grammaticalization, 2 volumes (set)* 1991
20. CROFT, William, Suzanne KEMMER and Keith DENNING (eds): *Studies in Typology and Diachrony. Papers presented to Joseph H. Greenberg on his 75th birthday.* 1990.
21. DOWNING, Pamela, Susan D. LIMA and Michael NOONAN (eds): *The Linguistics of Literacy.* 1992.
22. PAYNE, Doris (ed.): *Pragmatics of Word Order Flexibility.* 1992.
23. KEMMER, Suzanne: *The Middle Voice.* 1993.
24. PERKINS, Revere D.: *Deixis, Grammar, and Culture.* 1992.
25. SVOROU, Soteria: *The Grammar of Space.* 1994.
26. LORD, Carol: *Historical Change in Serial Verb Constructions.* 1993.
27. FOX, Barbara and Paul J. Hopper (eds): *Voice: Form and Function.* 1994.
28. GIVÓN, T. (ed.) : *Voice and Inversion.* 1994.
29. KAHREL, Peter and René van den BERG (eds): *Typological Studies in Negation.* 1994.
30. DOWNING, Pamela and Michael NOONAN: *Word Order in Discourse.* 1995.
31. GERNSBACHER, M. A. and T. GIVÓN (eds): *Coherence in Spontaneous Text.* 1995.
32. BYBEE, Joan and Suzanne FLEISCHMAN (eds): *Modality in Grammar and Discourse.* 1995.
33. FOX, Barbara (ed.): *Studies in Anaphora.* 1996.
34. GIVÓN, T. (ed.): *Conversation. Cognitive, communicative and social perspectives.* 1997.
35. GIVÓN, T. (ed.): *Grammatical Relations. A functionalist perspective.* 1997.
36. NEWMAN, John (ed.): *The Linguistics of Giving.* 1998.
37. RAMAT, Anna Giacalone and Paul J. HOPPER (eds): *The Limits of Grammaticalization.* 1998.
38. SIEWIERSKA, Anna and Jae Jung SONG (eds): *Case, Typology and Grammar. In honor of Barry J. Blake.* 1998.
39. PAYNE, Doris L. and Immanuel BARSHI (eds.): *External Possession.* 1999.
40. FRAJZYNGIER, Zygmunt and Traci S. CURL (eds.): *Reflexives. Forms and functions.* 2000.
41. FRAJZYNGIER, Zygmunt and Traci S. CURL (eds): *Reciprocals. Forms and functions.* 2000.
42. DIESSEL, Holger: *Demonstratives. Form, function and grammaticalization.* 1999.
43. GILDEA, Spike (ed.): *Reconstructing Grammar. Comparative Linguistics and Grammaticalization.* n.y.p.
44. VOELTZ, F.K. Erhard and Christa KILLIAN-HATZ (eds.): *Ideophones.* n.y.p.
45. BYBEE, Joan and Paul HOPPER (eds.): *Frequency and the Emergence of Linguistic Structure.* 2001.
46. AIKHENVALD, Alexandra Y., R.M.W. DIXON and Masayuki ONISHI (eds.): *Non-canonical Marking of Subjects and Objects.* n.y.p.
47. BARON, Irene, Michael HERSLUND and Finn SORENSEN (eds.): *Dimensions of Possession.* n.y.p.

A full list of titles published in this series is available from the publisher.

TYPOLOGICAL STUDIES IN LANGUAGE 45

FREQUENCY AND THE EMERGENCE OF LINGUISTIC STRUCTURE

Edited by

JOAN BYBEE

PAUL HOPPER

OFFPRINT

JOHN BENJAMINS PUBLISHING COMPANY

This is an offprint from:

Joan Bybee and Paul Hopper (eds)
Frequency and the Emergence of Linguistic Structure.
John Benjamins Publishing Company
Amsterdam/Philadelphia
2001
(Published as Vol. 45 of the series
TYPOLOGICAL STUDIES IN LANGUAGE, ISSN 0167-7373)

ISBN 90 272 2947 3 (Eur.) / 1 58811 027 3 (U.S.) (Hb.)
ISBN 90 272 2948 1 (Eur.) / 1 58811 028 1 (U.S.) (Pb.)
© 2001 – John Benjamins B.V.

No part of this book may be reproduced in any form,
byprint, photoprint, microfilm or any other means, without written
permission from the publisher.