Project
MUSE®
Today's Research. Tomorrow's Inspiration.

THE PRESENT OF THE ENGLISH FUTURE: GRAMMATICAL VARIATION AND COLLOCATIONS IN DISCOURSE

Rena Torres Cacoullos

James A. Walker

The Pennsylvania State University

York University

We use the variationist method to elucidate the expression of future time in English, examining multiple grammaticalization in the same domain (*will* and *going to*). Usage patterns show that the choice of form is not determined by invariant semantic readings such as proximity, certainty, willingness, or intention. Rather, particular instances of each general construction occupy lexical, syntactic, and pragmatic niches. While putative differences in meaning are largely neutralized in discourse, grammaticalization paths are reflected in particular constructions of different degrees of lexical specificity, which bear different nuances of meaning or tenacious patterns of distribution inherited from once-meaningful associations. We conclude that collocations contribute to the shape of grammatical variation.*

Keywords: variation, future, English, collocation, construction, grammaticalization path

1. Introduction: form-function asymmetry in discourse. Contrary to idealizations of isomorphism in linguistic form-function relationships—'one form for one meaning, and one meaning for one form' (as Bolinger (1977:x) puts it)—language use is characterized by form-function asymmetry: that is, choices among different constructions serving generally similar discourse functions (Labov 1969, Sankoff & Thibault 1981). This heterogeneity is structured, as multivariate quantitative models of speaker choices have shown, conditioned by both linguistic and extralinguistic factors. For tense/aspect/mood constructions, a major source of the variation is grammaticalization, the process whereby grammatical constructions gradually develop out of discourse patterns (Bybee 2006:719–21, Sankoff & Brown 1976). Studies of grammaticalization, however, have tended to examine individual cases, leaving unanswered the question of what happens when multiple forms undergo grammaticalization in the same grammatical domain. Is there a linguistic explanation for continued variation between two or more constructions with the same grammatical function (also known as LAYERING; Hopper 1991:22–24)?[1] Can the distribution of forms be attributed to traces (RETENTION) of semantic nuances from their lexical sources (e.g. Bybee & Pagliuca 1987)?

Future temporal reference in English is a case in point: this is an area of grammar in which a number of morphosyntactic alternatives coexist, as illustrated in 1: the simple Present (1a), the Present Progressive (1b), the modal *will* + verb (1c), and the periphrastic *BE going to* + verb (1d).[2]

[1] As opposed to social explanations, which we do not consider in this article (see Weinreich et al. 1968: 188).

[2] Following Comrie (1976:10), we capitalize the first letter of language-specific forms, to distinguish them from crosslinguistic functions or meanings.

(1) a. I **finish** on the twenty-seventh of June and I **start** the summer camp on
         the eighth of July.   (Q009:405)[3]
    b. In fact, I**'m leaving** on September the first for Belfast for a couple of
         weeks.   (M019:223)
    c. And he**'ll** probably live 'til a hundred.   (Q029:1480)
    d. My doctor tells me I**'m going to** live 'til a hundred.   (Q029:341)

Since many languages use more than one form to express future meaning (Bybee et
al. 1994:243), this area of grammar is clearly characterized by what Hopper (1991:24)
calls a ''cluttering'' . . . with functionally similar constructions'.

The considerations that lead speakers to choose among these alternate forms have
received much attention, with different meaning(s) attributed to each. In practice,
though, forms are equally acceptable in many of the same contexts, which calls for a
method that allows us to investigate the conditions that influence speakers' choices
among forms in spontaneous performance.

In this article, we make use of the variationist method to elucidate the grammatical
expression of future time in a recently constituted corpus of spoken Canadian English.
In contrast with much of the literature on this topic, rather than assuming differences
in meaning based on unverifiable intuitions or ascribing speaker motivations on an
example-by-example basis, we seek to determine the quantitative patterning of forms
in different linguistic contexts. Our goal is to empirically test the claims made in the
literature on the syntactic, semantic, pragmatic, and discourse factors constraining the
choice of form.

The facts of distribution revealed in our study suggest that different future expressions
are not distinguished by an overarching invariant meaning, be it temporal (proximity),
modal (certainty), or pragmatic (willingness, intention). Instead, although purported
semantic differences are largely neutralized in discourse (Sankoff 1988a; cf. Croft
2009), each future construction occupies small niches defined by particular construc-
tions of differing degrees of lexical specificity, from fixed collocations, or 'convention-
alized word sequences' (Bybee 2006:713), to more productive formations. Within
general constructions such as *will* + verb, expressions lie along a continuum of lexical
specificity and productivity, from *I'll tell you* to *X'll never* + verb. These niches
represent retention, not only of meaning (Bybee & Pagliuca 1987), but also, we suggest,
of (once-meaningful) patterns of distribution, as collocations that follow from the origi-
nal uses of the source constructions become conventionalized (cf. Torres Cacoullos
2001).[4]

Usage leads to different degrees of conventionalization of discourse patterns, includ-
ing collocations ('reusable fragments' (Thompson 2002:141) or 'prefabs' (Erman &
Warren 2000)) and more schematic constructions. In usage-based grammar, there is no
discrete division between fixed and productive expressions (see Bybee 2006, Goldberg
2006, Tomasello 2003). Collocations constitute an important locus of grammatical
development, since they may lead in changes and constitute subclasses that contour
the grammaticalization of more general constructions (Bybee & Torres Cacoullos 2009).

---

[3] Examples are reproduced verbatim from the recorded interviews in the Quebec English Corpus (Poplack
et al. 2006) and are identified by the community (Q = Quebec City, M = Montreal), speaker number, and
line number in the transcription.

[4] For example, in Spanish progressive forms, *estar hablando* 'be talking', *ir creciendo* 'be (go) growing',
and *andar buscando* 'be (go around) looking for', the original lexical meaning of the emerging auxiliary is
compatible with the main verb (gerund) (Torres Cacoullos 2000:175).

The patterns of future forms revealed in this article also point to an interaction of conventionalized expressions with general, productive constructions. Thus, rather than dismissing frequent collocations as simply lexicalized and therefore marginal to grammar, we view them as particular instances of constructions which, while formulaic, interact with their associated general construction (cf. Torres Cacoullos & Walker 2009). We conclude that collocations contribute to the shape of grammatical variation.

2. GRAMMATICALIZATION AND THE FUTURE. Grammaticalization is a recognized fact of language, despite controversy about its status in linguistic theory. Some critics have argued that it is merely a relabeling of multiple processes of linguistic change that occur independently, such as phonetic reduction or reanalysis (e.g. Newmeyer 1998, papers in Campbell 2001). While we acknowledge that grammaticalization involves independent components, it seems useful to us to give a name to a set of changes that have been found to go together (Hopper 1991:19; cf. Bybee 2008). More importantly, grammaticalization makes predictions about the role of frequency and crosslinguistic paths of development that are not offered by strictly modular views of grammatical change (e.g. Bybee & Hopper 2001, Givón 1979, Heine & Kuteva 2002, Traugott & Heine 1991). We thus adopt a definition of grammaticalization as the set of gradual structural and semantic processes whereby existing constructions with particular lexical items gain in frequency and become new grammatical constructions, following crosslinguistic evolutionary paths.

Each of the grammatical constructions used to express future temporal reference in English entered at different points in the history of the language. The earliest option was the simple Present, as illustrated in 2a, with the development of the Progressive in Early Modern English making a second variant available. The modal construction with *will* developed out of a verb of desire (< Old English *willan* 'want') between Old and Middle English (2b), and the periphrastic construction with *going to* grammaticalized from a verb of motion in a purpose-clause construction beginning in the fifteenth to sixteenth centuries (2c) (Royster & Steadman 1923/1968, Danchev & Kytö 1994).

(2) a. Þas flotmenn **cumað** and þé cucenne **gebindað** . . .

                    (Hml. S. 32, 70, 77; Toller 1921:227)

      'These seamen (will) come and (will) bind the living . . . '

    b. Hwyder **wilt** þu gangan? Min Drihten, ic **wille** gangan to Rome.

                    (971; *Blickl. Hom.* 191; *OED*)

      'Whither wilt thou go? My Lord, I will go to Rome.'

    c. Thys onhappy sowle . . . was **goyng to** be broughte into helle for the synne and onleful lustys of her body.

                (1482; *Monk of Evesham* (Arb.) 43; *OED*)

The basic discourse function of futures is prediction by the speaker. Two future grammaticalization paths, one from constructions expressing agent-oriented modalities of desire or obligation and another from allative movement-verb constructions (the *go*-future), recur across languages independently of genetic relatedness (Bybee et al. 1994, Heine & Kuteva 2002). An intermediate stage of intention, especially of first-person subjects, is proposed for both desire and movement futures (Bybee et al. 1994:254–70).

Multivariate models of synchronic variation can help identify the crosslinguistic features of grammaticalization paths. For example, absence of cooccurring temporal adverbials probabilistically favors the *go*-future, whether in alternation with the modal (obligation)-origin future in French (Emirkanian & Sankoff 1985:194) and Spanish (Blas Arroyo 2008:94) or in alternation with the futurate Present in Portuguese

(Poplack & Malvar 2007:153). Interrogatives also favor choice of the *go*-future in Spanish (Blas Arroyo 2008:103) and Portuguese (Poplack & Malvar 2007:151–52).[5] If we find similar results in other languages, as in the English variety studied here, these effects could be interpreted as evidence for a crosslinguistic grammaticalization path for the *go*-future, rather than being a reflex of the trajectory of other future forms with which it alternates, or attributable to other language-particular considerations. By contrast, the obligation-origin (infinitive + *habere* 'have') future in Spanish (Blas Arroyo 2008:100) and Portuguese (Poplack & Malvar 2007:151) is subject to the favoring effect of nonspecific time adverbials such as *luego/logo* 'soon, later'. The magnitude of the adverbial effect has strengthened over time, reflecting perhaps the increasing epistemic use of the form and consequent reliance on temporal adverbials to ensure future reading or the tinge of uncertainty in such future use (Aaron 2006, Sedano 1994: 234, Silva-Corvalán & Terrell 1989:207). If also found in other languages with modal-origin futures, such as English *will*, the nonspecific adverbial effect could reflect the pathways of agent-oriented modalities, which may feed into epistemic moods expressing possibility and probability (Bybee et al. 1994:240).

Of increasing relevance to the empirical study of linguistic variation and change are two complementary principles. The principle of RETENTION (Bybee & Pagliuca 1987) or PERSISTENCE (Hopper 1991), on the one hand, states that the semantic content of grammatical expressions not only includes their current grammatical functions but also retains features or nuances of meaning derived from their source constructions (Bybee et al. 1994:138, 148).[6] Quantitative studies have demonstrated that the lexical history of a grammaticalizing construction may be manifested in its current patterns of distribution, as shown in the favoring effects of linguistic contexts associated with the original lexical source, at least in the initial stages (e.g. Poplack & Tagliamonte 1999:332–37, Tagliamonte 1997:52–59, Torres Cacoullos 1999:29–32; cf. Hopper & Traugott 2003: 3). For example, if proximity is an early feature of the *go*-future following from its original 'agent on a path toward a goal' meaning, its effects should weaken over time. Evidence that *go*-futures first enter in proximate contexts is the finding of proximity effects precisely in earlier varieties (eighteenth-century English (Roy 2007) and nineteenth-century Brazilian Portuguese (Poplack & Malvar 2007:147)).

The principles of SEMANTIC BLEACHING and SYNTACTIC GENERALIZATION, on the other hand, describe the progressive weakening of such effects as the contexts in which the construction is used are generalized. These are both gradual processes. For example, in the development of the Spanish *a pesar de X* 'in spite of X', semantic bleaching from the notion of opposition by another person to concessive meaning is evidenced in the increasing frequency of coreferentiality between the subject of the verb and the notional subject of the abstract NP adnominal (following *de* 'of'), while syntactic generalization is shown by progressively increasing use with propositions in infinitives and finite clauses (Torres Cacoullos & Schwenter 2005).

Variation can be stable, without change, for quite a long time (Labov 2001:85–92). Given the complementary principles of semantic retention and bleaching, are forms eventually bleached of all of the semantics of their lexical source or do they always retain traces? Variation between the volition-based *will*-future and the motion-based *go*-future allows us to examine the principle of retention in cases of multiple grammati-

---

[5] We infer an interrogative effect from the favoring effect of second person reported in Poplack & Malvar 2007:151 (see §7.1).

[6] 'Retention' could be viewed as a new name for the observation that older forms may survive with a specialized function, in Kuryłowicz's (1947) fourth 'law' of analogy (see Joseph 2001).

calization within the same grammatical domain. Do lingering meaning differences explain the continued variation between *will* and *going to*? In particular, if we accept the view that newer grammatical resources express more specific meanings than previously existing grammatical forms (Bybee et al. 1994:133), we would expect retention of source-construction meaning to be more evident in *going to*, the newer variant, than in *will* (e.g. Bybee & Pagliuca 1987:116). If retention continues to be reflected in the distribution of these two variants, is it in fact more relevant for *going to*?

Collocations (sometimes called prefabs (Bybee 2006:713, and references therein) or discourse formulas (Torres Cacoullos & Walker 2009)) constitute important loci of grammatical development in diachrony. Rather than proceeding uniformly, grammaticalization of a general construction is contoured by particular subclasses. As Bybee and Torres Cacoullos (2009) show, high-frequency instances of constructions with particular lexical types may lead in semantic-structural changes and serve as the centers of subclasses of the general construction, attracting more lexical types and thereby contributing to the productivity of the construction, as found for Middle English verbs of communicating and cognition with modal auxiliary *can* and Old Spanish verbs of communicating with the progressive *estar* 'be located' gerund periphrasis.[7]

By supplementing multivariate analyses with cross-tabulations and other detailed views of the distribution of data (cf. Labov 1969:735, 742–47, 2004:10–11, Sankoff 1988b), we argue that collocations interact with general patterns, contributing to the structure of variation in synchrony as in diachrony (Bybee & Torres Cacoullos 2009). In response to the questions posed above, we argue that retention may involve collocations following from source-construction meanings rather than source-construction meanings per se.

3. THE MEANING OF THE FUTURE IN ENGLISH. Opinion is divided on whether future forms are semantically equivalent. Leech, who otherwise states that different verb forms are 'far from interchangeable', nevertheless notes that *will* and *going to* 'can often be substituted for one another with little change of effect' (1971:51). Similarly, Binnick (1971:47) says that 'speakers have a great deal of trouble . . . distinguishing meanings beyond noting some slight and tantalizingly subtle differences which seem perhaps to be matters of style and not of semantics'. Haegeman (1989) goes so far as to argue that *going to* and *will*, at least, are truth-conditionally equivalent, and that any difference between them lies in pragmatic import.

Quite apart from the modal or nonmodal origins of each of these forms, future temporal reference has been argued to be inherently modal since, unlike past reference, which involves states or events that have already occurred, future reference always constitutes a prediction by the speaker and therefore reflects their attitude toward the predicted situation (see Binnick 1971:40, Fries 1927 and references therein, Palmer 1990:140, Quirk et al. 1985:47, Royster & Steadman 1923/1968:395). Those who argue that the forms are not linguistically equivalent offer an assortment of explanations for the choice of each.

Will has been called an 'irrealis potential' (Klinge 1993:315) that refers to a predicted rather than verified state of affairs (Leech 1971:52), involving nuances of volition, willingness, and intention (Brisard 1997:277, Palmer 1965:138, 248), no doubt reflecting its desiderative origins (see Comrie 1985:45). The nuance of willingness is detectable in requests or indirect imperatives (such as *Will you call me a cab?*), and the

---

[7] Compare Malkiel's notion of 'influential leader verb' in sound change and morphological innovations (e.g. Malkiel 1968:27, 1973–74). We thank Brian Joseph for bringing this to our attention.

desiderative sense is instantiated in expressions such as *Do what you will . . .* (Bybee et al. 1994:16–17). By contrast, Quirk and colleagues (1985:47) call *will* 'the closest approximation to a colorless, neutral future'.

In contrast, *going to* is said to refer to a verified state of affairs (Klinge 1993), reflecting the speaker's certainty that some event will occur (Brisard 2001:252, Leech 1971:54, Palmer 1965:147, 174, Royster & Steadman 1923/1968:403). Another commonly asserted connotation of *going to* is the proximity or imminence of the future event (Hopper & Traugott 2003:89, Langacker 1987:219, 2002:332), following from the original purposive meaning of its allative-motion source construction and its inherent progressive aspect (Bybee et al. 1994:268). Comrie (1976:64–65) calls it a 'prospective aspect', in which a present state, the intention of the subject, is related to a future situation. In this sense, *going to* is said to be oriented to the present moment (Palmer 1990:144, Quirk et al. 1985:48) in a manner analogous to the Present Perfect (cf. Fleischman 1982:95–97). Interestingly, *going to* has also been argued to indicate more distal future reference (Gee & Savasir 1985:153) and to be more 'colorless' (Joos 1964:23, 1968:281) or 'neutral' (Brisard 1997:275) than *will*.

Some of these distinctions are supported by empirical studies of interaction. Gee and Savasir's (1985) study of children's 'activity-types' shows that *will* is used for 'undertaking' in cooperative activities to indicate offers and compliances, conveying willingness and temporal proximity, whereas *going to* is used for 'planning' ordered sequences of events, representing the speaker's projected plans from which the interlocutors are dissociated and which are both temporally and attitudinally distant. The interpersonal willingness meaning of *will* is also indicated by Myhill's (1992:81–87) finding that future events under the control of the subject favor *will*, as do future actions performed cooperatively, whereas unilateral and uncooperative controlled futures are associated with *going to*, as are uncontrolled future events that are evaluated as undesirable or for which there is a specific source making the prediction.

The two other major exponents of the future, the simple Present and the Present Progressive, have received less attention in the literature. Both are said to be associated with plans, arrangements, or schedules (Comrie 1985:47–48, Leech 1971:60, Palmer 1974:66, Smith 1997:190, Whyte 1949:35, cited in Visser 1970:1952). Smith says that both 'focus a preliminary situation which licenses a prediction about a [future] situation', though the use of each form differs depending on whether the situation is dynamic or stative (1997:191).

As we have seen, the more common accounts of semantic differences rely on notions like volition and intention, which refer to the speaker's intent or attitude toward the realization of the future eventuality. While our interpretations as hearers may involve attributing intentions to the speaker, we have no way of knowing whether these attributions are accurate (Labov 1994:549–50). Rather, empirical tests of semantic differences can only rely on clues in the linguistic context at the moment of utterance (Sankoff 1988a:154). A related consideration is that observed differences based on one independent variable may be an artifact of another if distributions are not uniform across the two; that is, the choice of future form may be conditioned not by a single factor but rather by a confluence of factors. These considerations call for multivariate analysis that considers the simultaneous effect of independent variables that test hypotheses about speaker choices. In this article, we report on such an analysis.

4. METHOD.

4.1. VARIATIONIST METHOD. In this research, we adopt the framework of variationist linguistics (e.g. Labov 1969, 1972a, 2004, Sankoff 1988a), which seeks to discover

patterns of usage in the relative frequency of cooccurrence of linguistic forms and elements of the linguistic context. The interpretative component of the variationist method lies in identifying similar discourse functions of different constructions, or variants (Sankoff 1988a:150, 154–55, Sankoff & Thibault 1981). We account for the selection of variants to fulfill a particular discourse function by exhaustively extracting each instance of that function in discourse and applying quantitative techniques to determine the influences of contextual factors on the choice of form.

4.2. DATA. The data on which our study is based were taken from recordings of sociolinguistic interviews (Labov 1984) with seventy-four native speakers of English, made in 2002–2003 in Quebec City and Montreal (see Poplack et al. 2006 for a discussion of the constitution of the corpus and the characteristics of the speakers). The corpus constitutes approximately 340 hours of recorded speech, almost three million words of transcription. Full and contracted forms (e.g. *going to* vs. *gonna*) were distinguished during the transcription process, which involved three passes by different members of the research team.

4.3. DEFINING THE VARIABLE CONTEXT. For the purposes of this study, we limit the variable context—the place in discourse where the speaker has a choice between forms—to events or states occurring after speech time. From the seventy-four interviews, we extracted every reference to future time. The four most robust variants are *will*, *going to*, the Present, and the Present Progressive.[8]

We did not count 'future' forms falling outside the variable context, that is, those used for functions other than referring to future time, such as gnomic (generic situations), habitual, and modal uses, as shown in 3a–c. We also excluded temporal clauses headed by *when*, *until*, and *before*,[9] as in 4, though they refer to future time, since these are near-categorically simple Present (cf. Comrie 1985:48).

    (3) a. If your pastry's good, it**'ll** be good.   (Q017:363)
        b. He**'ll** go out about ten-o-clock for a little walk over there.   (Q029:714)
        c. I **will** not have a kitchen with a small counter.   (Q021:2652)
    (4) She said, 'I'll s- I'll remember to spit **when I get there**.'   (M009:2315)

We excluded references to a time occurring after a reference point in the past with *would* (i.e. past *will*) (5a), *was/were going to* (5b), and past Progressives (5c). Although such 'backshifted' futures have been included in previous quantitative studies of the future (with present- vs. past-tense contexts as an independent conditioning factor; e.g. Poplack & Tagliamonte 1999:334, Roy 2007), separate analysis showed them to be subject to different constraints from their present-tense counterparts and their inclusion might confound the results.[10]

---

[8] We did not include the few tokens of *shall*.

[9] If futures constitute a prediction by the speaker (Bybee & Pagliuca 1987, Bybee et al. 1994:244), we do not expect futures in *when* clauses, since these are contingent rather than asserted (Bybee et al. 1994: 274).

[10] Backshifted futures differ from their present-tense counterparts not only in relative frequencies, with *was/were going to* by far the preferred variant, but also in their conditioning by language-internal factors. In a variable-rule analysis including *was/were going to* and the Past Progressive, choice of *would* is highly favored by negative polarity (i.e. *wouldn't*), unlike *will*, and the clause-type effect is reversed, with *would* disfavored in main clauses, perhaps because it tends to have modal meaning, whereas in indirect speech (*he said he would leave*) the main-clause verb establishes a temporal reference point (Comrie 1985:75). Another difference is the favoring of *was/were going to* but not Present-tense *be going to* by first-person singular.

(5) a. I said I thought I **would** work until Christmas.   (Q018:326)
    b. So one night they decided they **were going to** play a trick on him.

(M011:1556)

    c. We **were going** to Paris and we had nowhere to stay.   (M015:1058)

These protocols resulted in a dataset of 3,337 tokens.

	%	N
going to	42	1,418
will	42	1,389
Present Progressive	13	428
Simple Present	3	102
TOTAL *N*		3,337

TABLE 1. Distribution of variants of future temporal reference in Quebec English.

4.4. OVERALL DISTRIBUTION. Table 1 shows the overall distribution of forms. Note that *going to* (42%) and *will* (42%) are equally divided in these data, together occupying most (84%) of the future-temporal-reference sector. The remainder of this space is occupied by the Present Progressive and, to a lesser degree, the simple Present.

Similar rates for *going to* have been reported for other varieties of Canadian English (34% in Ottawa (including *was/were going to*, Poplack & Tagliamonte 1999:326) and 42% in Toronto (Tagliamonte 2006a:314)) and American English (47% in the Santa Barbara Corpus of Spoken American English (Szmrecsanyi 2003:302)).[11] *Going to* is said to be more frequent in some North American varieties, such as African American English (Labov et al. 1968:250, Winford 1998:113; cf. Poplack & Tagliamonte 2001: 219) and spoken American as opposed to British English (Berglund 2005:75, Szmrecsanyi 2003:302, Tottie 2002), though increasing rates in apparent time have been observed in York (Tagliamonte 2002:753–56). Studies also indicate the constraints of register: *going to* and the Present Progressive are more frequent in speech than in writing (Berglund 1997:16, Palmer 1990:142, Quirk et al. 1985:214, Wekker 1976, Whyte 1949, cited in Visser 1970:1952) and *going to* is more frequent in drama than in fiction (Krug 2000:171). Similarly, Tagliamonte and Denis (2008:20–22) report that teenagers use *going to* more frequently in spoken conversations (43%) than in instant messaging (35%) with the same interlocutors.

However, overall rates of use can be uninformative about the organization of the grammar, precisely since they may fluctuate according to situational factors, such as channel, genre, topic, or style. Moreover, variationist comparisons of genres and regional varieties for other variables have revealed similar grammatical patterning. For example, despite disparate rates in Spanish subject-pronoun expression, morphosyntactic and discourse constraints (verb class and tense, distance from and form of previous coreferential subject) are shared across regions (Cameron 1993) and genres (conversation and narrative) (Travis 2007).

Thus, we take as more revealing of grammatical structure the LINGUISTIC CONDITIONING of forms, that is, probabilistic statements about linguistic subcontexts that differ significantly in the relative frequencies of the variants (Poplack & Tagliamonte 2001: 92).[12]

---

[11] For the purposes of cross-variety comparison, we note that the rate of *going to* (excluding future-in-past) is 27% (*N* = 370) in Ottawa and 45% (*N* = 2,312) in Toronto (Sali Tagliamonte, p.c.).

[12] While increasing token and relative frequencies indicate advancing grammaticalization, this is accompanied by changes over time in the configuration of constraints contributing to variant choice (Poplack & Malvar 2007, Torres Cacoullos 2009).

5. Operationalizing hypotheses as factors. We coded each token for a series of factors, adapted from hypotheses and findings in the literature. As noted above, many proposed semantic distinctions rely on access to speaker motivations, which cannot be directly ascertained in a replicable way. Coding such distinctions on the basis of the form would be circular, since it is the formal expression of future marking that concerns us. We therefore test a number of hypotheses indirectly, operationalizing predictions by relying on clues in the discourse context (cf. Poplack & Malvar 2007:137–43, Poplack & Taglimonte 1999:321, Poplack & Turpin 1999:145–46). We divide the factor groups into those concerned with temporal proximity, those involving the lexical verb, those involving the type of subject, and those involving the grammatical/discourse context.

5.1. Temporal proximity. Proximity in time is a frequently cited consideration in discussions of the future, not only in English but also in other languages (e.g. Blas Arroyo 2008, Fleischman 1982, 1989, Poplack & Malvar 2007, Poplack & Turpin 1999, Roy 2007).[13] A first problem in coding this factor group is determining how close in the future an event has to be in order to be considered 'proximate'. A further problem is that a full 70% of tokens in our dataset contain no overt indication of temporal distance.[14] Where proximity could be reliably inferred from the discourse context, we began by making a fine-grained distinction, ranging from events taking place within a minute of speech time, as in 6a, to events in the far future, as in 6i.

(6) within a minute:
    a. Yeah, I**'m gonna** tell the story 'cause you n- you just screwed it up.

<div align="right">(Q076:978)</div>

    within an hour:
    b. He**'s gonna** go in the shower, so we have another hour.   (Q021:2258)
    within a day:
    c. My niece **is coming** in tonight.   (Q029:13)
    tomorrow:
    d. Now tomorrow I**'m going out** with the girls from the railway for lunch.

<div align="right">(M019:1147)</div>

    within a week:
    e. Monday afternoon there**'ll** be a bingo.   (Q006:1973)
    within a month:
    f. We**'re** still **going to** be here in two weeks, you know.   (Q009:1162)
    within a year:
    g. He and I are getting married. We**'re gonna** elope in January. (M096:733)
    far future:
    h. We're accepting it now, but they **won't** always accept that.   (Q013:659)
    i. I can almost picture myself in ten years, I **will** want to go back.

<div align="right">(Q068:1170)</div>

---

[13] In their comparison of African American English enclaves and British-origin varieties in Canada, where situations up to one month away favor *going to* in the latter but not the former, Poplack and Tagliamonte (1999:336) actually suggest that the proximity reading of *going to* is a later development. We interpret the lack of significance for proximity in (urban, mainstream) Ottawa as opposed to (rural, conservative) Guysborough Village in Nova Scotia as indicating an earlier reading of proximity (see Roy 2007). In studies of present-day varieties, within-the-day contexts favor the *go*-future strongly in Spanish (Blas Arroyo 2008) but not so clearly in French (Poplack & Turpin 1999:150–51).

[14] Unfortunately, other quantitative studies of the English future do not report the proportion of data for which temporal distance could be coded.
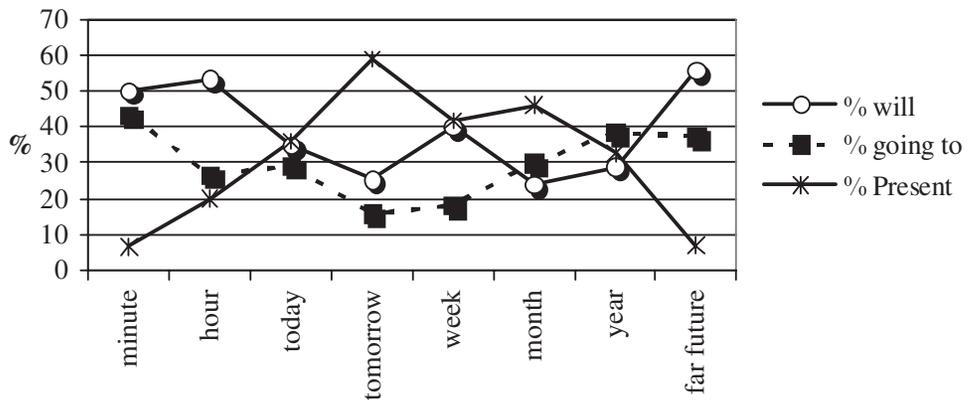
FIGURE 1. Distribution of future markers by temporal distance ($N = 1,034$).

As is evident from Figure 1, which provides a graphic display of the distribution of the three variants (combining the simple and Progressive Present) in each of these contexts, there does not appear to be a straightforward linear correlation between degree of temporal proximity and any of the variants (cf. Poplack & Turpin 1999:150). The rate of *will* is higher than any other variant in the most distal context ('far future') but also in the most proximate contexts ('within a minute' and 'within an hour'); moreover, we observe a disjunction such that the proportion of *will* is much greater in these contexts than in later 'today' contexts. The rate of *going to* declines from 'within a minute' to 'tomorrow', which we would expect if it were associated with proximity, but then increases, presenting a U-shaped curve. The Present shows an upside-down U-shaped curve, peaking at 'tomorrow', with its highest rate at distances between 'tomorrow' and 'within a month'.[15]

However, proximity may be related to the speaker's PERCEPTION of proximity rather than an objective measurement of distance in time (cf. Dahl 1984, Fleischman 1982, Myhill 1994:11), consistent with the view of tense as a form of 'temporal deixis' that relates events or states to the speaker's 'now' (e.g. Comrie 1985:14–15, Dahl 1985: 2). Crosslinguistically, grammatical distinctions of temporal distance are more common for the past than for the future (Dahl 1984:113), but an 'immediate future', which may also convey aspectual and modal meaning, occurs in a number of languages (Bybee et al. 1994:244–47). In particular, Dahl (1984:112–13) cites a 'hodiernal/nonhodiernal' ('within the day' vs. 'beyond today') distinction as quite common (if not fundamental) crosslinguistically (though cf. Comrie 1985:88). Therefore, in subsequent analyses, we define as 'proximate' those contexts occurring within the same day, and all others as 'distal' (cf. Poplack & Malvar 2007:141, Poplack & Turpin 1999:150, Roy 2007). This distinction is supported by our data, in which all references within the day make up a large proportion (35% (358/1,034)) of tokens with an overt indication of temporal distance.[16]

---

[15] The only difference in temporal distance between the simple Present and the Progressive Present (not shown in Fig. 1) is that the highest rate of the former is in 'tomorrow' contexts while the latter peaks at 'tomorrow' and 'within a month'.

[16] Of the tokens for which temporal distance could be determined, those predicted to occur between 'tomorrow' and 'within a month' make up 16% (168/1,034), between one month and one year from speech time another 22% (232/1,034), and far future (over a year away) 27% (276/1,034).

5.2. LEXICAL VERB. We coded each individual lexical verb type to test for lexical effects (e.g. Poplack 1992:255), in addition to sorting lexical verbs into semantic or aspectual classes, to test for retention and proposed differences in meaning. With respect to MOTION verbs, a favoring effect on *going to* (7a) could be interpreted as an indication of retention, since combinations of *going to* and another motion verb would constitute the type of 'harmonic' constructions that appear in the early stages of grammaticalization (Aaron 2006:139–40, Torres Cacoullos 1999:34–36).[17] Furthermore, if the meaning of allative motion is retained, *going to* should be disfavored by STATIVES (7b), which are incompatible with motion (e.g. Hopper & Traugott 2003:68, Traugott 2003: 365). We also distinguished TELIC verbs (that is, verbs with an inherent end point) (7c), which have been found to favor the futurate Present Progressive (Myhill 1992:89, but cf. Palmer 1974:66). Since telic verbs are associated with higher transitivity (Hopper & Thompson 1980:271), their cooccurrence may be taken as a measure of volitional subjects. All other verbs were coded as dynamic, or verbs of ACTIVITY (7d) (cf. Comrie 1976:41–51).[18]

(7) a. I'm one of those scared people, 'I'm **gonna** *fall*.' (M096:2266)
 b. So they **are going to** *be* in Montreal at one point. (Q068:1240)
 c. And I thought, 'This **will** never *end*.' (M012:1510)
 d. I think they**'ll** *make* more out of the kids. (Q040:1848)

As another way to test for volition or control, we distinguished whether the verb was used TRANSITIVEly (8a) or INTRANSITIVEly (8b), since transitive predicates are more likely to be under the control of the subject (cf. Hopper & Thompson 1980:265). Although transitivity and verb class are correlated (statives tend to be intransitive (73%) while telic predicates tend to be transitive (63%)), they do not overlap completely.

(8) a. Take your soother, Mommy**'s gonna** *pick you up*. (Q048:698)
 b. There seems to be a- a small boom going on right now and I think it**'s** only **gonna** *get bigger*. (M050:513)

5.3. TYPE OF SUBJECT. To further operationalize ideas about modal meanings, we distinguished among grammatical persons (9a–d) of the subject and, in third person, between ANIMATE and INANIMATE subjects (9d–e). We predict that retention of desire or volition (from the origin of *will*) or 'agent on a path toward a goal' (from the origin of *going to*) should be manifested in animacy effects, since inanimate subjects are incompatible with volition or purposive motion (cf. Bybee & Pagliuca 1987: 113, Bybee et al. 1994:255, 268, Hopper & Traugott 2003:89). Intention is operationalized via grammatical person, since FIRST-PERSON SINGULAR subjects (9a) are most compatible with the inference of intention (e.g. Bybee & Pagliuca 1987:118, Bybee et al. 1994:264, Poplack & Turpin 1999:155, Sedano 1994:235). By contrast, nuances

---

[17] However, Poplack and Tagliamonte's (1999:336; cf. Poplack & Malvar 2007:138) finding that motion verbs disfavor *going to* in Early African American English, coupled with the lack of significance for verb class in Roy's (2007) study of eighteenth-century English texts, suggests that a favoring effect for motion verbs would indicate semantic bleaching rather than retention.

[18] Dynamic predicates involve change, whereas stative predicates describe situations that will continue unchanged unless something happens (e.g. *know* (*a language*) (Comrie 1976:48–50)). We coded stativity considering the lexical type, independently of contextual interpretation, for the sake of consistency (cf. Poplack & Tagliamonte 2001:128, 141, Walker 2000, 2001:17–19). The most frequent stative lexical types were *be* ($N = 445$), *have* (100), *know* (16), *live* (26), *need* (16), *see* (79), *stay* (34), and *try* (20). Less frequent statives include *cost*, *depend*, *enjoy*, *feel*, *forget*, *hear*, *hurt*, *keep*, *like*, *love*, *remember*, *sit*, *understand*, and *want*.

of willingness and cooperative activity should be reflected in a favoring effect of FIRST-PERSON PLURAL (9b).

(9) a. *I*'m never **going to** be on that end of it, never.   (M024:1563)

   b. I'll just lift up the stickers and *we*'ll slide out the boxes.   (M038:451)

   c. I think *you*'re **gonna** like it.   (Q025:547)

   d. They're so afraid of losing their culture but *they*'ll never lose their culture.
      (M038:2183)

   e. 'We're teaching you Irish dancing, and *it*'s **gonna** be in English.'
      (Q021:1865)

We also coded for the thematic role of the subject independently of verb transitivity, distinguishing between AGENTS (compare 10a and 10b) and all OTHER roles (10c–f) (e.g. Fillmore 1968, Gruber 1976, Jackendoff 1972). Since the development of *going to* involved a pragmatic shift from intention of the subject-experiencer to prediction by the speaker (Langacker 1990:23, Traugott 1995:36, 50, n. 4), this should be evinced by attenuation of agentivity in syntactic subjects (cf. Verhagen 1995:107–8). By contrast, subjects' greater control of *will* future events should be manifested in a favoring effect by agentive subjects.

(10) a. Wait I'll just put the dog out. (agent, transitive verb)   (Q023:1241)

   b. I'll play against them. (agent, intransitive verb)   (Q040:126)

   c. It's not **gonna** take away all the attention. (cause)   (Q060:253)

   d. Oh, it- it- it- **it will** change. (theme)   (M037:2068)

   e. That's- if **they're gonna** feel at home. (experiencer)   (Q040:1248)

   f. **I won't** get them 'til tonight, though. (recipient)   (M009:3056)

Although these factors represent operationalizations of degree of volition or control on the part of the subject, congruent with desire or willingness (*will*), purposive motion (*going to*), or intention (both) (cf. Bybee & Pagliuca 1987:118), the fact that the same types of subject are likely to exhibit willingness, purposive motion, and intention (Myhill 1992:80) raises the question of how any subject effects should be interpreted. Following from the expectation that retention applies more to *going to*, as the newer variant, than *will* (e.g. Bybee & Pagliuca 1987:116), the prediction is that first-person singular, animate, and agentive subjects should favor choice of *going to* over *will*. Thus, both sets of factor groups outlined above, those related to the verb and those related to the subject, test the twin grammaticalization principles of retention and bleaching.

5.4. GRAMMATICAL/DISCOURSE CONTEXT. The type of clause in which the future form occurred was coded in one of three ways. If one future expression conveys more uncertainty than the other, we hypothesize that it should be favored when the realization of the future event depends on the satisfaction of a previously expressed condition, as in apodoses (that is, the consequences of conditional *if*-clauses) (11a) (Poplack & Malvar 2007:142, Poplack & Turpin 1999:153, Silva-Corvalán & Terrell 1989:201). We also hypothesize that *going to*, as the newer variant, should be favored in MAIN clauses (11b), which are viewed as more innovative than OTHER clause types (11c) (Bybee 2001; cf. Heine 2003:586).

(11) a. If we can't have kids, then we'll adopt kids.   (M096:1732)

   b. I said, 'Oh, I'll tell you what happened.'   (Q013:945)

   c. I just don't think they're **going to** do the same thing again, you know.
      (M024:1289)

In addition, we coded for the presence of temporal adverbials in the same clause, distinguishing between 'definite' adverbials (including *now*, *tonight*, *tomorrow*, *this*

afternoon/week/summer/year etc., as well as clock times and calendar dates), which make reference to a specific time (12a), and 'indefinite' adverbials (including *soon*, *later*, *someday*, *eventually*, *always*, *ever*, and *never*), which have nonspecific temporal reference (12b).[19]

(12) a. Well, *tomorrow* me and a couple friends **are going to** be going over to my friend's house. (Q041:1159)

b. Well I- *someday* **I'll** be leaving. (Q051:726)

For the type of sentence in which the token occurred, we distinguished among declaratives (13a), negatives (13b), *yes/no* questions (13c), and WH-questions (13d).

(13) a. And we**'re going** out for our two-year anniversary uh- Saturday night. (Q056:353)

b. You had to be very aware that if you're English, you**'re** not **gonna** go far. (M009:2284)

c. **Will** that hurt if I pull this around? (Q047:865)

d. I said, 'Where in the heck **are** you **going to** travel to?' (M012:1195)

5.5. VARIABLE-RULE ANALYSIS. All of the factor groups were considered individually and together in multivariate analysis (in this case, variable-rule analysis) using Gold-Varb X (Sankoff et al. 2005), the goal of which is to discover the set of factor groups that JOINTLY account for the largest amount of variation in a statistically significant way (Paolillo 2002, Sankoff 1988b, Tagliamonte 2006b). In experimental studies, such as those done in psycholinguistics, goodness of fit is measured by the amount of variance accounted for by the statistical model. However, as Sankoff (1988b) points out, natural spoken data are not distributed evenly across cells, and factors are often not completely independent of each other. For this reason, maximum likelihood is a more meaningful measurement of goodness of fit.

In the tables, the 'input' indicates the overall likelihood that the variant will occur. The numbers in the first column represent the probability (or factor weight) that each factor contributes to the occurrence of the variant: the closer to 0, the less likely, the closer to 1, the more likely. The 'range' provides an indication of the relative strength of each group of factors within each analysis. The second column shows the percentages, and the third column shows the total number of tokens in each factor.

6. THE NICHE OF THE FUTURATE PRESENT: *go* AND MOTION VERBS. We first examine the conditioning of the futurate Present, combining both simple Present and Present Progressive into a single variant as opposed to *will* and *going to*.[20] Table 2 shows a multivariate analysis of the relative contribution of the factors selected as significant to the occurrence of the futurate Present.

---

[19] The temporal adverbial and proximity factor groups are independent. A greater proportion of definite than indefinite temporal adverbials cooccur with future situations within the same day as speech time (17% (52/299) vs. 4% (9/218)), but also with those predicted to occur 'tomorrow' or later (66% (197/299) vs. 23% (50/218)). Not surprisingly, most indefinite temporal adverbials cooccur with future situations whose distance from speech time is not indicated (73% (159/218)). Unlike in this study, Poplack and Turpin (1999: 151) found that 'temporal proximity [is] correlated with . . . type of adverbial specification, such that contexts coded as proximal happen to cooccur with non-specific ['indefinite'] adverbs, while those coded as distal tend to cooccur with specific ['definite'] adverbs' (in the present data, proximal and distal contexts cooccur with definite adverbials 18% (65/358) and 34% (227/676), with indefinite adverbials 3% (10/358) and 9% (61/676), respectively; 79% (283/358) of proximal and 57% (388/676) of distal lack cooccurrence with a temporal adverb of any kind).

[20] Comparison of analyses showed that this combination provides a significantly better fit to the data than considering them separately, suggesting that the overall distribution of simple Present with respect to Present Progressive reflects the more general variation between simple and Progressive forms across all tenses.

		Input:	.129	TOTAL *N*: 3,337	
				%	N
SEMANTIC CLASS OF VERB					
Motion			.83	48	626
Telic			.55	12	1,218
Activity			.36	7	548
Stative			.27	5	931
	Range:		56		
TEMPORAL ADVERBIAL					
Specific/definite			.71	37	360
No adverbial			.48	14	2,701
Nonspecific/indefinite			.42	10	276
	Range:		29		
PROXIMITY					
Distal			.58	26	676
Proximal (same day)			.35	16	358
	Range:		23		
SENTENCE TYPE					
WH-Question			.68	24	132
yes/no Question			.57	23	73
Declarative			.50	16	2,652
Negative			.46	13	470
	Range:		22		
TRANSITIVITY OF VERB + THEMATIC ROLE OF SUBJECT					
Intransitive + Agent			.63	34	1,028
Intransitive + Other			.44	8	894
Transitive + Agent			.45	9	1,077
Transitive + Other			.43	6	327
	Range:		20		
GRAMMATICAL PERSON AND ANIMACY OF SUBJECT					
2nd person			.62	22	433
1st person pl.			.55	19	371
1st person sg.			.48	16	1,232
3rd person, animate			.47	17	795
3rd person, inanimate			.45	8	506
	Range:		17		

Factors not selected: Clause type.

TABLE 2. Factors contributing to the choice of futurate Present (simple Present and
Present Progressive vs. *will* and *going to*) in Quebec English.[21]

As indicated by its range (= 56), which is approximately double the next largest, the most important factor group is verb class.[22] Most highly favoring the futurate Present are motion verbs (.83), a result unanticipated in the literature, to our knowledge (but see Poplack & Malvar 2007:152–54). Moreover, motion verbs, as Table 3 shows, are overwhelmingly represented by two lexical types, *go* and *come*. These two verbs together account for 88% (553/626) of this class; *go* alone represents 70% (436/626). In addition, HALF OF ALL PRESENT TOKENS (277/530) are *go* and *come*, and such verbs are overwhelmingly associated with the Progressive rather than the simple Present (44% of motion verbs occur as Progressive, while just 4% are simple Present). Although such disparities might lead us to conclude that the strong effect of verb class merely

---

[21] Note that totals within each factor group may not add up to the total number of tokens in the entire analysis. Uncodeable tokens or tokens that could not be included in any of the factors were excluded.

[22] The disfavoring effect of statives (.27) applies to the Progressive, which has a rate with stative predicates of only 2% (compared to an overall rate of 13%); in comparison, the rate of the simple Present with statives is

reflects the lexical properties of two particular verbs, a separate analysis excluding *go* and *come* still shows verb class to have the greatest effect, with verbs of motion most favorable to the Present.[23] Our finding that futurate Present *go* and *come* have a high token frequency and make up a big proportion of futurate Present occurrences, coupled with the finding that motion verbs as a class favor choice of this future variant, provides evidence for the role of particular subclasses or instances of constructions in general patterns synchronically.[24]

	will	*going to*	Present	*N*
go	26%	24%	50%	436
come	27%	21%	51%	117
Other	32%	37%	32%	73
OVERALL:	27%	25%	48%	626

TABLE 3. Distribution of variants of future temporal reference in motion verbs.

The strong favoring effect of definite temporal adverbials lends credence to claims that the futurate Present involves considerations of planned events, as in 14. In other words, the interpretation of Present as future largely depends on the presence of a definite adverbial. Indeed, the proportion of tokens unaccompanied by an adverbial is significantly smaller for the Present (70%, 369/530) than for either *will* (80%, 1,113/1,389) or *going to* (86%, 1,219/1,418).[25] Furthermore, although only a quarter (138/530) of futurate Presents cooccur with a definite adverbial, futurity may be inferred by shared knowledge of the situation as planned. For example, marriage, giving birth, coursework, and graduation from school, moving to another location, and vacations, as in 15, are all usually scheduled or expected to occur at specific times in the future.

(14) a. Because they**'re closing** up the Holland Centre. *Next Wednesday*.

(Q029:816)

b. My family**'s having** a reunion (*coughs*) *in August*. (Q034:1973)

c. She**'s going** *on the day after Thanksgiving*, she**'s going** to Japan to teach for a year. (M056:105)

(15) a. I said, 'I'm very sorry, **I'm getting married**.' (Q006:204)

b. The second you hit grade eleven, like, everybody's like, 'Oh, **you're graduating**, you can do what you want.' (M048:437)

c. So, well **we're staying in youth hostels**, but it's still like twenty-five bucks a night or thirty dollars a night. (Q031:990)

---

identical to its overall rate (3%) (see Table 1). Among statives, *be* is the most frequent in simple Present (13/25) (i–ii), while *have* is the most frequent in Progressive (8/16) (iii).

(i) It**'s** tonight. (Q023:1058)

(ii) There**'s** a wedding next week. (Q061:418)

(iii) After the mass they**'re having** this thing in th- in the school. (Q029:2200)

[23] Variable-rule analysis of factors contributing to the choice of futurate present excluding *go* and *come*, $N = 2,784$, Input 0.111: Verb type—Motion 0.77, Telic .64, Activity .46, Stative .32 (Range = 45); Proximity—Distal .61, Proximal .31 (Range = 30); Transitivity + Thematic Role—Intransitive + Agent .63, Intransitive + Other .50, Transitive + Other .45, Transitive + Agent .45 (Range = 18); Grammatical Person and Animacy—2nd person .62, 1st person pl. .54, 1st person sg. .49, 3rd person animate .47, 3rd person inanimate .46 (Range = 16); Clause type—Main .54, Apodosis .50, Other .37, (Range = 17); not selected as significant: Sentence type.

[24] Diachronic study would show whether first to be used in the futurate Present Progressive were high-frequency *going* and *coming*, which then served as central members of a motion verb class, attracting other verbs with similar semantics.

[25] For Present vs. *will*, $\chi^2 = 24.08111687$, $p = 0.0000$; vs. *going to*, $\chi^2 = 68.40354568$, $p = 0.0000$.

Nevertheless, the inference of futurity is not NECESSARILY correlated with an adverbial and a scheduled event, as with *going shopping* or *going to the bar* in 16 (contra Bybee et al. 1994:251, Comrie 1985:118). In particular, Present tokens without adverbial modification largely occur with the verb *go*: 76% (166/218) of tokens of *go* occur in the absence of a temporal adverbial, compared to 65% (203/312) of all other verbs in the futurate Present. In other words, highly frequent *go* is less dependent on adverbial modification for future interpretation than all other verbs.

(16) a. She was um, on the phone to me this morning, you know, to say uh- you
        know, '**Are** you **going** shopping?' I said, 'No, I can't- I can't go this
        morning.'   (M009:1768)
     b. I**'m going** to the bar.   (Q059:007)
     c. But he- he liked Canada so much, he said, 'I**'m bringing** my family.'
                                                                       (Q013:189)

The proximity effect, with proximal (i.e. same-day) events DISFAVORING use of the Present, is unexpected, since we would expect the Present to be more closely tied to proximal future (e.g. Jespersen 1924:260–61).[26] We argue that the proximity effect is largely apparent and can be attributed to the effect of definite adverbials, which mostly (66%) occur in distal contexts (see n. 19).[27] Most favorable to the Present are expressions with proximate deictic demonstrative *this*, such as *this morning*, *this weekend*, or *this year* (in which context the rate of Present is 42% (19/45)), highlighting the importance of speaker viewpoint in defining proximity. Nevertheless, the finer temporal-distance distinctions in Fig. 1 indicate that the Present may be most frequent for events predicted to occur between tomorrow and one month from speech time.

We can attribute the favoring effect of questions (whether *yes/no* or WH) to the Present Progressive, the relative frequency of which is much greater than average in these contexts (22% (16/73) in *yes/no* questions, 21% (28/132) in WH-questions) and which may be generally favored in interrogatives and subordinate clauses (cf. Walker 2001: 25–26).[28] The effects of transitivity/thematic role and person/animacy are epiphenomenal. The favoring effect of intransitive verbs with agentive subjects largely follows from the favoring effect of motion verbs: while the rate of the Present with intransitives in general is 22% (419/1,922) (compared to 8% (111/1,404) with transitives), the rate with intransitive verbs of motion is more than double, 49% (301/626). Similarly, the favoring effect of second-person subjects likely reflects the favoring effect of questions (see §7.1). Finally, the (weak) disfavoring effect of inanimate subjects can also be related to that of motion verbs, most occurrences of which (91%, 570/626) have an animate subject.[29]

---

[26] Although the distal effect for the Present (Table 2) could merely reflect a favoring effect of proximal futures for one of the other variants (*will* or *going to*), neither of these variants is particularly associated with proximity (see Table 4).

[27] This interpretation is supported by a cross-tabulation of temporal adverbials and temporal distance, which shows that definite adverbials favor the Present just as much in distal (41%) as proximal (35%) contexts, while indefinite temporal adverbials disfavor as much in proximal (10%) as in distal (13%) contexts.

[28] This interpretation receives further support from separate analyses of the Progressive and the simple Present, in which sentence and clause type are selected as significant for the Progressive but not for the simple Present.

[29] Grammatical person + animacy interacts with transitivity + thematic role, in that agents are (near-) categorically (99%, 2,081/2,105) animate. Furthermore, subject animacy is not independent of verb class, in that stative verbs have a higher proportion of inanimate subjects (26%, 242/931) than telic (12%, 141/1,218) or atelic (13%, 69/548) dynamic verbs; the rate of inanimate subjects with motion verbs is just 9% (56/626).

In conclusion, the results confirm that futurate Present is associated with temporal adverbials indicating a definite time in the future, as anticipated by intuition-based analyses (e.g. Comrie 1985:47–48), though, contrary to assumptions (e.g. Comrie 1985: 118), reference to a scheduled event is not required. Further, the results reveal that speakers' choice of the futurate Present is most likely with going, coming, and other verbs of motion. The observed effects of proximity, transitivity/thematic role, and grammatical person/animacy follow from these two associations, while that of sentence type is likely related to general properties of the English Progressive. Thus, the futurate Present occupies a small but well-defined niche of the future-temporal-reference sector, that involving 'going' and planned events.

7. *will* vs. *going to*. In the remainder of the analysis, we set aside the futurate Present tokens and focus on the variation between the two most robust future forms, *will* and *going to*, neither of which predominates (Table 1). Besides featuring similar overall rates, the constructions appear to be functionally equivalent. Both can appear in near-identical contexts, for example upcoming birthdays (17a–b), offers to assist (17c–d), and schooling plans (17e–f), including from the same speaker (1c–d above, repeated in 18).

> (17) a. He**'ll** be two in August. (Q023:1247)
>   b. Same age as Emily, he**'s gonna** be six in- August. (M019:1470)
>   c. I**'m going to** help you. (Q006:1448)
>   d. But, sure. I**'ll help** you out. (M050:25)
>   e. I**'m gonna** send them to an English school. (Q026:2329)
>   f. Well since I did all my high school and elementary in English so I **will** send them to an English school. (Q075:839)
> (18) a. And he**'ll** probably live 'til a hundred. (Q029:1480)
>   b. My doctor tells me I**'m going to** live 'til a hundred. (Q029:341)

Whether the situation is currently a case of stable variation (cf. Berglund 2005:111, 132, Krug 2000:169) or of ongoing change in some communities (e.g. Mair 1997, Tagliamonte 2002:748, 2006a:313), the distribution of future variants in English is quite different from that of French, Brazilian Portuguese, and Spanish, where the *go*-future is clearly the default variant (Blas Arroyo 2008, Poplack & Malvar 2007, Poplack & Turpin 1999). Certainly, there is nothing like the rapid generalization of the *go*-future in Romance languages.[30] Thus, the ongoing robustness of *will* is apposite for testing the retention hypothesis. Do speakers choose between these two alternatives based on lingering semantic differences?

Table 4 shows the results of a multivariate analysis of the factors contributing to the occurrence of *will* (disfavoring effects are interpreted as favoring *going to*). Refuting temporal distance as one of the considerations in intuition-based analyses, the effect

---

[30] In another study of our data (Walker et al. 2004), we found a slight increase in *going to* in younger speakers, more so in Montreal than in Quebec City, suggesting change in apparent time. Royster and Steadman (1923/1968:396–97) report an increase in written texts since the seventeenth century and Krug (2000:169) a sharp rise in normalized frequency between 1850 and 1950. Although Mair (1997) suggests that *going to* increased in the latter half of the twentieth century, we note that this claim is based on absolute rather than relative frequencies (1997:1539). Unlike English, in Spanish the *go*-future is the default, even though it appears with appreciable frequency (4%) after the seventeenth century (Aaron 2006:135–36), about the same time as *going to* (Roy 2007, Royster & Steadman 1923/1968). This difference suggests that the *go*-future has grammaticalized faster in Spanish than in English.

of proximity is virtually nil.[31] Nor are subject agentivity-verb transitivity (a measure of retention of volitionality in *will*) and verb class (a measure of retention in *going to*) significant. Instead, a strong consideration, one not even mentioned in previous treatments, is sentence type, with a range (= 45) nearly double the next largest, that of clause type (= 23), together with temporal adverbials and grammatical person-animacy. Since this account of the variation has not been foreseen in previous analyses, we discuss each of these effects in turn. We find that contributing to these general effects are collocations or particular instances of each construction.

	Input:	.524	TOTAL *N*: 2,807	
			%	N
SENTENCE TYPE				
Declarative		.54	52	2,235
Negative		.47	49	410
yes/no Question		.31	29	56
WH-Question		.09	8	100
	Range:	45		
CLAUSE TYPE				
Apodosis		.59	56	166
Other main		.53	52	2,069
Other		.36	38	572
	Range:	23		
TEMPORAL ADVERBIAL				
Nonspecific/indefinite		.67	65	249
Specific/definite		.48	51	226
No adverbial		.48	48	2,332
	Range:	19		
GRAMMATICAL PERSON AND ANIMACY OF SUBJECT				
1st person sg.		.56	57	1,034
1st person pl.		.50	51	302
3rd person, inanimate		.49	46	468
3rd person, animate		.47	46	663
2nd person		.38	35	340
	Range:	18		
PROXIMITY				
Proximal (same day)		.49	55	302
Distal		.51	54	498
	Range:	2		

Factors not selected as significant: Semantic class of verb, verb transitivity, thematic role of subject.

TABLE 4. Factors contributing to the choice of future *will* (vs. *going to*) in Quebec English.

7.1. *I'll* COLLOCATIONS AND THE ASSOCIATION OF *will* WITH WILLINGNESS, PROXIMITY. Since the subject-type factors operationalize modal meanings, the relative weakness of their effects is striking. The absence of an animacy effect indicates that both variants have generalized beyond their original meanings (*going to* and physical movement toward a goal; *will* and desire/volition). The lack of significance for transitivity/thematic role and verb class similarly reflects the high degree of grammaticalization of both variants as simple expressions of prediction by the speaker, thus failing to provide

---

[31] In Table 4, proximity is selected as significant despite (near-)identical percentages because of its interaction with adverbials (i.e. there are very few indefinite adverbials in proximal contexts) once futurate Present tokens are excluded (cf. n. 19).

evidence for subject control (e.g. Myhill 1992:82) or verb telicity (e.g. Gee & Savasir 1985:155) as general considerations.[32]

The effects of grammatical person fail to support putative semantic distinctions. The significance of this factor group is largely due to the disfavoring of *will* by the second person, which we attribute at least in part to the disfavoring effect of questions. Not only are more than one-third of questions (37% (57/156)) second person (compared to 14% (56/410) of negative declaratives and 10% (223/2,235) of affirmative declaratives), but also 17% (57/336) of second-person subjects occur in questions (compared to only 4% of first- (50/1,335) and third-person (49/1,130) subjects). This strong association between questions and second-person subjects may help to explain person effects in multivariate analyses that include both interrogatives and second person as factors.

What of the favoring of *will* in first-person singular contexts? This further argues for the bleaching of intention meaning in *going to*, contrary to the assumption that the newer construction may be expected to show more retention. But neither is this due to retention of a general willingness meaning in *will*. A close look at frequent first-person expressions reveals that a disproportionate number of the tokens coded as 'within a minute' occur in first-person singular (76% (171/224) of 'within a minute' is first person, while 48% (387/810) of all other tokens for which temporal distance could be determined is first person). This frequency of first-person, 'within a minute' *will* could be taken to indicate persistence of willingness in the use of *will* for offers (cf. Gee & Savasir 1985:148, 154).

However, first-person 'within a minute' tokens are largely collocations, consisting of first-person singular *will* (*'ll*) phrases (*I'll tell*, *I'll give*, and others listed in 19) that make up a greater proportion than average of their corresponding lexical type. For example, *I'll tell* tokens add up to more than half (58%) of all *tell* tokens (whereas first-person singular *will* makes up 21% of the data overall). These collocations together account for one-fifth (21% (123/589)) of first-person singular *will*.

(19)　I'll tell . . .　　　58% of *tell* (40/69)
　　　 I'll pay . . .　　　53% of *pay* (8/15)
　　　 I'll ask . . .　　　44% of *ask* (8/18)
　　　 I'll talk . . .　　　38% of *talk* (9/24)
　　　 I'll call . . .　　　35% of *call* (11/31)
　　　 I'll teach . . .　　 33% of *teach* (7/21)
　　　 I'll give . . .　　　33% of *give* (21/64)
　　　 I'll try . . .　　　 29% of *try* (6/21)
　　　 I'll speak . . .　　26% of *speak* (6/23)
　　　 I'll say . . .　　　 20% of *say* (7/35)

Furthermore, we note that 'within a minute' occurrences may have an interactional function, to introduce a topic, as in 20a, or to take or keep the floor, orient the listener, or otherwise manage the interaction, as in 20b. In fact, tokens of *I'll tell you* make up approximately 10% (24/224) of the 'within a minute' data.

---

[32] Considering attributions of attitudinal distance to *going to* (e.g. Gee & Savasir 1985), we distinguished third-person subjects with distal demonstrative *that* (i) and proximate *this* (ii) to test the prediction that the rate of *going to* will be higher with *that* subjects, but failed to find a significant difference (52% (24/46) with *that*, 74% (17/23) with *this*).

　(i)　Yeah. Oh, that's- *that***'s gonna** be hard.　　　　　　　　　　　(M008:1346)
　(ii)　*This* **is gonna** work out well.　　　　　　　　　　　　　　　　(M058:815)

(20) a. 'Oh my God, now she**'s gonna** tell this is when the cops came.'

<div align="right">(Q007:1667)</div>

  b. **I'll give** you an example.

<div align="right">(M096:2004)</div>

Rather than providing evidence for retention of willingness, the first-person singular effect is due to *'ll*. Independent analyses show that choice of the full form *will* is actually favored by NP subjects (i.e. third person), while contracted *'ll* is favored more by first person than second person, even excluding interrogatives.[33] While this result is consonant with the use of *'ll* for offers (*I*) or cooperative actions (*we*) (cf. Gee & Savasir 1985, Myhill 1992:81–87), such connotations do not represent a straightforward case of retention from desiderative *will*: if they did, they should appear in the full (presumably less bleached) rather than the reduced form. Instead, we must acknowledge the existence of particular collocations, such as *I'll give*, *I'll tell* (see 19 above), and *we'll see* (which alone makes up 11% (14/127) of the *we'll* data). Thus, we hypothesize that WHAT PERSISTS ARE COLLOCATIONS, which may have originated with meanings that followed from the source construction. We conclude that, while some offers may be made with *will*, the perception of proximity and willingness may have more to do with fixed discourse formulas that contribute to the effect of *I*.

7.2. TEMPORAL ADVERBIALS. With respect to temporal-adverbial cooccurrence, most favorable to *will* are indefinite adverbials. The distribution of future variants by adverbial context (Table 5) confirms that this effect concerns *will* and is not attributable to *going to*, which is not in fact favored by definite as opposed to indefinite adverbials. Rather, nonspecific time adverbials (other than *soon*), such as *eventually*, *later*, and *someday*, favor *will* (see 12b). Though we are unaware of any studies of the extent of epistemic use of English *will* (but see Gee & Savasir 1985:171), the favoring effect of *will* with indefinite adverbials (other than *soon*) may indicate the (partial) association of *will* with uncertainty about the prediction or, rather, about the timing of its realization. Nevertheless, it is important that the general indefinite adverbial effect in favor of *will* is due in part to one adverb, *never*, which would seem a rather certain prediction about nonrealization. Close to one-third (78/249) of the indefinite adverbial data consists of *never*, which mostly (72% (56/78)) occurs with *will*.[34] Furthermore, *will* + *never* occurs overwhelmingly (89% (50/56)) in the collocation *X will/'ll never*.

Some association of *will* with uncertainty about the predicted situation comes from other adverbial expressions. Although neither variant predominates with adverbials indicating certainty, *definitely* and *for sure* (47% (8/17)), *will* is preferred with *probably* and *maybe* (75% (63/88)), as in 21 (cf. Myhill 1992:87). However, while this cooccurrence pattern may contribute to perceptions of uncertainty in *will*, such explicit indications are quite rare (just 3% of the data (105/3,337)).

---

[33] Variable-rule analysis of factors contributing to the choice of *'ll* and *will* (within { }), each opposed to *going to*, N = 2,350 {1,737}, input .386 {.166}: Sentence type—Declarative .56 {.53}, Negative .14 {.35}; Temporal Adverbial—Indefinite .77 {.75}, Definite .47 {.51}, None .47 {.47}; Type of subject—NP .21 {**.76**}, Pronoun **.53** {.44}; Clause type—Apodosis .56 {.61}, Main .54 {.52}, Other .34 {.42}; Lexical verb—not significant. The favoring of the contracted form by first-person singular may also be interpreted as a frequency effect (e.g. Bybee 2006), since *I* is by far the most frequent subject type in the data. First-person plural and second person, however, show disparate effects despite near-identical token frequencies. In variable-rule analysis of factors contributing to choice of *'ll* in affirmative declarative contexts only, N = 1,965, input .449, in a factor group distinguishing grammatical persons, 1st singular .61 (N = 751), 1st plural .54 (N = 248), 2nd .42 (N = 209), 3rd inanimate .41 (N = 306), 3rd animate .40 (N = 451).

[34] 249 is the number of indefinite adverbials with *will* and *going to* (not including futurate Present variants).

	will	*going to*	Present	Progressive	TOTAL *N*
NO ADVERBIAL	41%	45%	2%	12%	2,701
DEFINITE ADVERBIAL					
this . . .	24%	33%	7%	36%	45
Other	33%	31%	12%	25%	315
INDEFINITE ADVERBIAL					
soon	38%	54%	0%	8%	24
never	72%	26%	0%	2%	78
Other	55%	32%	6%	8%	174

TABLE 5. Distribution of future variants with temporal adverbials.

(21) a. And who knows, some of the- some of them **will** come back, *maybe*.

(M009:2801)

  b. They**'ll** *probably* give him like a desk job. (Q072:143)

  c. I mean, they'll- they**'re** *probably* **gonna** be stuck here. (Q077:537)

The vast majority of tokens have no adverbial modification, the context least favorable for *will*. In fact, the proportion of *going to* tokens is higher than the proportion of *will* tokens without a temporal adverbial (86% (1,219/1,418) vs. 80% (1,113/1,389, χ^2 = 17.00249374, p = 0.000). This greater independence of *going to* from adverbial modification lends support to the view that temporal meaning was present as an inference from the spatial meaning of *going to* from the beginning (Bybee et al. 1994:269, Langacker 2002:330–33). Thus, rather than being a metaphoric extension of space into time (e.g. Langacker 1987:219–20), the grammaticalization of *go*-futures involves the bleaching of spatial meaning in tandem with the generalization to contexts in which subjects are not moving physically to fulfill the expressed intention.

7.3. CLAUSE TYPE. The second-strongest effect is clause type, with apodoses favoring *will* (cf. Szmrecsanyi 2003:309). This effect indicates a preference for *will* when the future event is contingent on another, as in 22 (or a 'potential' event, Nicolle 1998). This may indicate the uncertainty expressed by a modal-origin future as opposed to the certainty of a movement future (cf. Palmer 1983:242), or alternatively, the retention may have less to do with meaning than with a particular construction (i.e. 'if *p*, then . . . will *q*'), a question that can only be settled by diachronic study of distribution patterns.[35]

(22) a. I can tell you that *if I ever move*, you know, like, again, I **will** have them come in and do everything. (M028:1371)

  b. *If the place is ever raided*, I**'ll** be pulled in, but I'm doing it legal.

(M028:1424)

  c. And he says- he says, '*If I ever catch you back down here*,' Edward says, 'we**'re gonna** call the police and tell your parents.' (Q007:1689)

  d. You can contest this, but *if you try and contest it*, you**'re gonna** get a record. (Q066:1771)

Nevertheless, *going to* (and the Present) occurs in *if*-clauses (23) (cf. Szmrecsanyi 2003: 309), indicating that it has generalized to at least some kinds of clauses that do not make assertions about future time.[36] *Going to* still does not generally occur in temporal

---

[35] No crosslinguistic picture emerges so far of the effect of apodoses on future variation: this context only weakly favored the modal future (*Futur*) in French (Poplack & Turpin 1999:153) and favored the futurate Present in Portuguese (Poplack & Malvar 2007:152).

[36] In *if*-clauses, *going to* may add a sense of already-decided or -mentioned course of action (discourse-old information), which may follow from its original meaning, though this sense may well accompany *if*-clauses with future marking regardless of the future source, such as Greek *an* 'if' + *θa* (cf. Hedin 2000),

(*when*) clauses, though we did find the occasional token, as shown in 24. In contrast, *will* does not occur at all in our data as a future in either *if-* or temporal (*when*) clauses; in these contexts it may well have nonfuture interpretations (cf. Comrie 1985:48, 117–21, Nicolle 1998:240).[37]

    (23) And- she's so into her family, like, I love my family and everything, but *if* I**'m going to** do something, I- I gotta go.   (M080:971)

    (24) *When* I**'m gonna**, you know, give myself to someone, don't wanna like, wake up and see him in my house when I**'m going** to have breakfast or anything.   (Q021:735)

  Overall, clauses coded as Other disfavor *will*, or favor *going to* (cf. Szmrecsanyi 2003:306), contrary to our expectation that the newer variant should be favored in main clauses (cf. Bybee 2001).[38] For many clauses coded as Other, however, the main-clause antecedent is a frequent collocation of cognition verb and first-person subject (25a,b), such as *I think* (which alone makes up 11% (61/572) of Other clauses). Such collocations may function less as main-clause propositions and more as conventionalized epistemic or evidential phrases (Thompson & Mulac 1991; cf. Torres Cacoullos & Walker 2009). Clauses such as those beginning with *so* or *(be)cause* (25c, 26a) may also act more like main clauses than subordinates (cf. Cheshire 2005:90). In any case, if greater speaker certainty is part of the conventionalized meaning of *going to*, we might expect a higher frequency with main-clause predicates expressing certainty (*be sure*, *believe*, *fact that*, *find out*, *guarantee*, *know*, *realize*, *swear*, *think*) (67% (101/151)) than ones expressing uncertainty (*can't tell*, *doubt*, [*I*] *guess*, *hope*, *(not) imagine*, *not much chance*, *not likely*, *wonder*, negative *be sure*, *know*, *think*) (56% (57/101)) (25–26). However, the difference between certainty and uncertainty predicates in the rate of *going to* is not significant, belying an explanation for the clause effect based on attributions of nuances of (un)certainty.[39]

---

which derives from a desiderative rather than movement source construction (θelo 'I want' + subordinator *na*) (Joseph 1990 [1978]; cf. Pappas 2001).

[37] According to Comrie (1985:120), *will* is required when 'the time reference of the subordinate clause is subsequent to that of the main clause and there is a causal relation from the main clause to the subordinate clause' (e.g. *If you'll do the shopping for me, I'll give you some money*). If true, this generalization is compatible with retention of the sense of willingness in *will* (though Nicolle (1998:240) proposes that the volitional interpretation in such contexts is a pragmatic inference). The generalization of the *going to* but not *will* future to *if*-clauses and parallel developments in Spanish (Sedano 1994:229) may be because the evolution of *go*-constructions into futures is 'more direct' than that of modal-origin futures (Bybee et al. 1994:268).

[38] Sentence type and clause type are not entirely orthogonal (interrogatives are overwhelmingly (145/156) main clauses), but such overlap does not appear to cause interaction. In relative clauses, which generally encode background information (Goldberg 2006:130, but cf. Fox & Thompson 1990:306), the rate of *going to* is 50% (34/68). Thus, neither variant seems particularly associated with foreground information, 'the material which supplies the main points of the discourse' (Hopper & Thompson 1980:280; cf. Labov 1972b, Labov & Waletzky 1967).

    (i) And you**'ll** n- the best deal *you***'ll** *ever get* will be because somebody told you.   (M050:931)

    (ii) In fact, I have a- pictures *I***'m gonna** *show you* of Richard being with the Queen when the Queen Elizabeth came.   (Q006:396)

[39] Here we present alternative figures. There is a higher rate in the presence of *think* (65% (56/86)), *know* (80% (28/35)), and *be sure* (50% (8/16)) than *don't think* (42% (11/26)), *don't know* (69% (35/51)), and *not sure* (1/1), but the difference does not achieve significance (67% (92/137) affirmative vs. 60% (47/78) negative polarity). The *go*-future in Spanish, in contrast, is favored with verbs of 'certainty', such as *saber* 'know', as opposed to 'uncertainty', such as *no saber* 'not know' (Aaron 2006, table 5.44).

(25) a. And *I think* she**'s gonna** need uh- you know, she'll need that extra support.
                                                                                        (Q048:865)
     b. *I'm sure* today there**'ll** be a lot of people at the movies.   (Q067:1310)
     c. This summer. *'Cause* I**'m going** to McGill next year.   (Q066:86)
(26) a. So *I don't know* what**'s going to** happen here.   (Q029:1982)
     b. She says, she says, '*I don't know* what they**'ll** do to us when they get
         in.'   (Q057:2319)

We could explain the clause-type effect by appealing to the origins of *going to* as part of the development of the Progressive (Danchev & Kytö 1994:72), which is more frequent in subordinate clauses (Walker 2001:25). That this effect is a retention from earlier patterns of distribution is supported by its diminished strength in mainstream as compared to (conservative) enclave varieties of English (Poplack & Tagliamonte 1999:334).[40] The initial spread and continued favoring of *going to* in the context of complement-taking predicates expressing speaker viewpoint (*I think*, *I don't know*) lends support to the view that the grammaticalization of the *go*-future involves a shift from motion by the subject to prediction by the speaker (Traugott 1995:36; cf. Langacker 1987:219). This is an instance of subjectification, the diachronic process whereby meanings grounded in external objective reference change toward meanings based in the speaker's internal belief or attitude (Traugott 1989, Traugott & Dasher 2002).

In summary, speakers tend to choose *will* over *going to* in apodoses of conditionals, which may be interpreted as a retention of nuances of (un)certainty in the two forms or as entrenchment of the older variant in a particular construction. Nevertheless, if earlier *going to* was associated with certainty about the already-in-motion future event, the use of *going to* in *if*-clauses indicates its generalization to nonassertive contexts. Furthermore, since *going to* is preferred with complement-taking predicates regardless of the degree of certainty, the clause-type effect may reflect the early generalization of *go*-futures via subjectification and present-day retention, not so much of the original meaning, but of early patterns of distribution (cf. Torres Cacoullos 2001).

7.4. SENTENCE TYPE. By far the strongest effect observed in Table 4 is that of sentence type. Interrogatives strongly disfavor *will* (*yes/no* (.31) and especially WH- (.09)) or, in other words, strongly favor *going to*. This result is somewhat unexpected, since questions are rarely if ever invoked in the abundant literature on the English future. If there is retention of source-construction meanings, questions with *going to* retain a nuance of intention, as in 27, and questions with *will* concern desire or willingness, as in 28, which functions as a request rather than an inquiry (cf. Ervin-Tripp 1976:33). That is, following from their 'agent on a path toward a goal' origins, *go*-futures can be used to inquire about already-decided intentions or projected plans (cf. Gee & Savasir 1985, Hopper & Traugott 2003:3), which may be a frequent function of future questions (or at least more frequent than inquiring about desires or making a request by appealing to the subject's willingness).

(27) Oh you**'re going to** go into your Ph.D. too?   (M019:981)
(28) So, I was like- he's like, '**Will** you always wear that perfume when you
     come?'   (M123:1008)

The interaction of sentence type and grammatical person (Table 6) is important, since meanings other than prediction are more likely in second-person questions (Nicolle

---

[40] In Spanish, the *go*-future is also favored by subordinate clauses more strongly in earlier than in present-day varieties (Aaron 2006:§5.3.5).

1998:238–39). While the difference between first person and second person is negligible (combining *yes-no* and WH-, 12% (6/50) vs. 9% (5/57)), the rate of *will* is significantly lower in second-person than in third-person interrogatives with a human subject, as in 29c (29% (7/24), combining *yes/no* and WH-questions).[41] This person effect (i.e. the favoring of *going to* particularly in second-person questions) suggests that speakers more often question interlocutors about their intentions or plans (with *going to*) than about their willingness or wants (with *will*).

Nevertheless, a retention explanation of the general question effect does not fully account for the data, since not all second-person interrogatives are patently queries about either intentions or willingness (for example, 29a seems rhetorical). Furthermore, the rate of *will* in questions with inanimate subjects (24% (6/25)), as in 29b, to which willingness is not applicable, is not significantly lower than that in questions with third-person human subjects, just as there is no general animacy effect (Table 4). And what of the one-third (50/156) of first-person questions, which also show a high rate of *going to*? Here a retention explanation does not work, since speakers are unlikely to ask about their own intentions (any more than their own wishes; cf. Bybee & Pagliuca 1987: 116).[42]

(29) a. They tell me, 'What **are** you **gonna** do? You **gonna** be a waitress for
         the rest of your life?'   (M096:2150)
      b. What **will** happen when they hit thirty?   (Q041:2526)
      c. So I'm here, but how long will I be here, I don't know. And then where
         will they be, I don't know.   (Q048:886)

| | AFFIRMATIVE | | NEGATIVE | | *yes/no* QUESTION | | WH-QUESTION | |
	N	% will	N	% will	N	% will	N	% will
1st	1,103	60	182	43	13	23	37	8
2nd	223	41	56	39	27	19	30	0
3rd	909	45	172	58	16	50	33	15
animate	526	43	112	62	5	80	19	16
inanimate	383	47	60	50	11	36	14	14

TABLE 6. Cross-tabulation of sentence type with grammatical person of subject.

A closer examination of the data again reveals collocations. Interrogatives are on a continuum according to the type of information or response requested, from factual questions to expressions of speaker stance (Freed 1994). First-person questions often serve as rhetorical devices. For example, the expression *What am I going to do?* alone makes up about a quarter (12/50) of first-person questions. (Testimony to the formulaic status of this expression is the lone occurrence of *What'll I do?*.) The expressions *Is there gonna be . . . ?* and *What's gonna happen?* also make up a substantial portion

---

[41] $\chi^2 = 5.566433066$, $p = 0.0183$.

[42] A structural explanation of the general interrogative effect may be a higher likelihood for auxiliary *be* (as opposed to *will*) to invert, since *going to* originated in the Progressive and we found the futurate Present Progressive to be also favored by questions. By contrast, the contracted form *'ll* does not invert in *yes/no* questions, being virtually restricted to affirmative declarative contexts in the present data. Nevertheless, we note that noninverted *yes/no* questions with *going to* and the Present Progressive are quite common, whereas all direct *yes/no* questions with noncontracted *will* are inverted in these data. Nor would a structural explanation tell us why WH-questions in particular disfavor *will*.

   (i) You**'re** all **gonna** be in one bedroom?   (Q023:1347)
   (ii) I said, 'You**'re getting** married?'   (M019:1502)
   (iii) I said, 'Simon, you **serving** this morning?'   (Q057:1651)

(18% (9/49)) of third-person questions. These distributions suggest that collocations also contribute to the question effect.

The effect of negation, by contrast, is slight, which indicates that both *going to* and *will* are highly grammaticalized and still viable, since neither is the newer *going to* favored in declarative contexts nor is the older *will* restricted to negative-polarity contexts, which may be conservative (cf. Givón 1979, Pappas 2001).[43] We again note a person effect, however, in that the rate of *will* is significantly higher in third-person negatives (58% (99/172)) than in first- and second-person negatives (42% (101/238)). Though cooccurrence with *want* (compare 30a and 30b) indicates bleaching of the original desire meaning, instances of *won't* may convey refusal in addition to futurity (31) (cf. Bybee & Pagliuca 1987:115). Thus, as with interrogatives, modal nuances may persist more in first and second person than in third person.

    (30) a. 'Cause their jobs are here, and they **won't want** to move.   (Q066:2094)

          b. I mean, they**'re not gonna want** to come to work in Quebec City there.

                                                                 (Q061:2009)

    (31) Well I **won't** be going, I **won't** be going on no picnic if it's raining.

                                                                 (Q006:008)

To summarize, the interaction of sentence type with person suggests that nuances of intended, already-decided course of action may persist in *going to* (and the Present Progressive), especially in second-person questions. Similarly, the disfavoring of *will* in first- and second-person negatives but not in third person may be due to retention of willingness. The lack of a strong general effect for negation, however, indicates that both variants are considerably advanced in their grammaticalization. Importantly, fixed rhetorical or discourse formulas (e.g. *What am I gonna do?*) contribute to a strong effect of questions in favoring *going to*.

7.5. CHOOSING BETWEEN *will* AND *going to*. What features of the linguistic context contribute to speakers' choice of *will* or *going to*? We have seen that three independent factors constrain the variation: sentence type, clause type, and temporal adverbial.

Before summarizing the results, we justify our decision to analyze reduced and nonreduced realizations together. We note that *gonna* is more frequent than *going to* (with a relative frequency of 32%, almost twice that of *going to*). In independent multivariate analyses of factors contributing to the choice of *going to* and *gonna* (each opposed to forms of *will*), one difference is that *going to*, unlike *gonna*, is favored by definite adverbials, though *gonna* too is favored more by definite than indefinite adverbials. Another is that *going to* is disfavored by inanimate subjects.[44] However, both the reduced and nonreduced form pattern similarly with respect to *will* in the lack of a favoring effect of first singular, and more importantly, the favoring effect of interrogative contexts and subordinate clauses.[45]

---

[43] Givón (1979:122) gives examples from Bantu languages of the lack of extension of new tense-aspect markers to negative clauses; Pappas (2001) shows that negation favors the conservative forms of the Greek future. In his eighteenth-century data, Roy (2007) found no cases of *going to* in seventy tokens with negative polarity. Negation has an inconsistent effect on the (older) modal-origin future in Romance (in alternation with the (newer) *go*-future), strongly favoring it in French (Emirkanian & Sankoff 1985:195, Poplack & Turpin 1999:155), but disfavoring it in Spanish (Aaron 2006, Blas Arroyo 2008:103).

[44] The definite adverbial effect (if it is an early constraint indicating certainty) and the absence of an animacy effect indicate that *gonna* is more grammaticalized (cf. Krug 2001), as anticipated by the finding that formal reduction proceeds in parallel with semantic reduction (Bybee et al. 1994:106–24, 279).

[45] Variable-rule analysis of factors contributing to the choice of *gonna* and *going to* (within { }), each opposed to variants of *will*, N = 2,291 {1,905}, Input .316 {.176}: Sentence type—WH-Question .76 {.68}, *yes/no* Question .65 {.57}, Negative .54 {.51}, Declarative .48 {.49}; Clause type—Other .57 {.62}, Main .49

As with *gonna*, the contracted *'ll* is more frequent than the full form (with a relative frequency of 35%, three times that of *will* (excluding *won't*), at 12%). Independent analyses of factors contributing to the choice of *will* and *'ll* (opposed to forms of *going to*; see n. 33) show that both are disfavored by negatives, though negative declarative *'ll* is categorically *X'll never* . . . , suggesting that this construction has become conventionalized. Both *will* and *'ll* are favored by indefinite adverbials and main clauses including apodoses. The major difference, with a reversal in the constraint hierarchy, is structural: *'ll* is favored by pronominal subjects, while *will* is favored by full NP subjects, exactly as we would expect for a contracted form (e.g. McElhinney 1993, Walker 2000).[46] Multivariate analysis of affirmative declarative *'ll* confirms that *'ll* and *will* are structurally determined allomorphs, with the same constraint hierarchy for clause type and temporal adverbial (cf. Gee & Savasir 1985:143, n.1, who find no differences between the two forms).[47]

The favoring effect of questions on the *go*-future, although absent from the intuition-based literature, has been hinted at in some empirical studies of English and Romance language varieties (Aaron 2006, Blas Arroyo 2008, Gee & Savasir 1985:153, Poplack & Malvar 2007). The interaction of sentence type with person suggests the persisting meaning of a planned, intended course of action in questions with the *go*-future, especially with second-person subjects. With respect to negative contexts, persistence of nuances of (un)willingness in first- and second-person *will* (*won't*) may account for the favoring of *going to* in first- and second- but not third-person contexts. At the same time, we find that frequently occurring rhetorical or discourse formulas (e.g. *What am I gonna do?*, *What's gonna happen?*) contribute to the general sentence-mode effect.

The lingering preference for *going to* with complement-taking predicates or epistemic phrases such as *I think* and *I don't know* may reflect the early generalization of the *go*-future via subjectification (Traugott 1995). The clause-type effect is doubtless a retention of earlier distribution patterns rather than of differences in meaning, since we find no difference between complement-taking predicates based on the speaker's degree of certainty about the proposition.

Nevertheless, contingency or certainty about the timing of the future situation may still contribute to speaker choice in some contexts, as indicated by the favoring of *will* in *if*-clause apodoses and by indefinite temporal adverbials. The nuance of (un)certainty is attributable to retention in both variants. On the one hand, certainty is consonant with the original progressive aspect of 'agent on a path toward a goal' in *going to*, and its retention would relegate *will* to uncertain situations. On the other hand, nuances of desire or willingness may persist in *will*, which can tinge the future situation with uncertainty. Such nuances may even reflect later grammaticalization of *will*, since future constructions that originate in agent-oriented modalities may develop into epistemic markers of present probability (Bybee et al. 1994:176–212), which may affect their

---

{.47}, Apodosis .42 {.48}; Temporal Adverbial—None .52 {.51}, Definite .45 {.58}, Indefinite .38 {.36}; Subject—3rd animate .55 {.49}, 2nd .54 {.61}, 3rd inanimate .54 {.45}, 1st plural .52 {.49}, 1st singular .43 {.50}; Lexical verb: not significant for *gonna*.

[46] Although *'ll* does cliticize to full NPs, as in (i)–(ii), it is disfavored.

   (i) *Derek***'ll** tell you more.   (Q035:1336)
   (ii) The *dust***'ll** be there after I'm gone.   (M028:1838)

[47] An analysis opposing *'ll* to *will* (excluding forms of *going to*, *won't*, and interrogatives) shows that the contracted form is distinguished by polarity (Declarative .62, Negative .06), as negated contracted *'ll* is restricted to *never*, temporal adverbial (Indefinite .75, No adverbial .47, Definite .43) (also largely due to *never*), and subject type (Pronoun .57, NP .08).

uses as futures (Aaron 2006, Silva-Corvalán & Terrell 1989:207). The only way to determine whether (un)certainty (here operationalized as clause type and adverbial specification) is attributable to the grammaticalization path of one or the other variant is by a diachronic study of variation between *will* and *going to* and further crosslinguistic comparisons. Furthermore, it is important to note that *never* alone makes up one-third of the indefinite adverbial data.

The advanced grammaticalization of both variants is indicated by the absence of animacy and verb-class effects. Further evidence for the generalization of *going to*, the younger variant, is its occurrence in *if*-clauses, which are not assertive contexts. At the same time, both variants are viable, as indicated by the lack of a general polarity effect (unlike French, in which the *go*-future is clearly the default and the synthetic *Futur* is restricted to negative contexts). Note that neither variant is strongly favored in the most frequent contexts, declaratives (.54), main clauses (.53), and absence of adverbial modification (.48).

Thus, we may conclude that *will* and *going to* are not distinguished by an overriding semantic difference along a single dimension of temporal distance, certainty, or interpersonal relations (i.e. willingness vs. plannedness). Instead, each has particular but small niches. Particular collocations are responsible at least in part for these distribution patterns. Thus, fixed formulas contribute to the interrogative effect, *I'll* collocations boost the proportion of *will* in first-person singular, especially immediate within-the-minute, contexts, and the *X'll never* collocation adds to the indefinite adverbial effect.

8. CONCLUSION. Our multivariate analyses of variation in the expression of future temporal reference in English reveal a contoured picture that cannot be reduced to a simple, monotonic axis, such as 'temporal distance' or 'certainty'. However, although many of the putative differences in meaning are neutralized in discourse (cf. Sankoff 1988a), we have nevertheless been able to discern lexical, syntactic, or pragmatic niches for each of the constructions.

The niche occupied by the futurate Present (including both the more frequent Progressive and the quite rare simple Present) is small but well defined: *going*, *coming*, and other motion verbs and scheduled events, as shown in the magnitude of the verb class and temporal adverbial effects, which are independent of each other. The lexical verb *go*, which makes up 70% of the motion verb class, and temporal adverbials indicating a definite time in the future, such as *this* phrases (*this afternoon/week/summer/year*), both highly favor the choice of futurate Present.

Unlike the Romance languages, in which the *go*-future is taking over, English *will* and *going to* are equally viable. Importantly, neither form is strongly favored in the most frequent contexts: declaratives, main clauses, and absence of adverbial specification. Thus, putative invariant meanings are not supported by these facts, since most of the time neither form is 'expected' to appear. Nevertheless, vestiges of earlier source constructions persist, as revealed by distribution patterns.

The niches of *going to* include interrogatives and clauses with complement-taking predicates such as *I think* and *I don't know*. The niche occupied by *will* includes the space of indefinite adverbials and apodoses of *if*-clauses. The interaction of sentence type with person suggests that nuances of a planned or intended course of action may persist in questions with *going to*, especially with second-person subjects, as do nuances of (un)willingness in first- and second-person negatives with *will* (*won't*). The lingering preference for *going to* with epistemic phrases is likely a retention of earlier distribution patterns (i.e. the early generalization of the *go*-future via subjectification) rather than

differences of meaning, since we find no differences among complement-taking predi-
cates based on the speaker's degree of certainty about the proposition. However, (un)cer-
tainty about the future may be expressed by the different constructions in certain
contexts, as indicated by the effects of *if*-clause apodoses and indefinite adverbials. We
can attribute this nuance to the retention both of the original progressive aspect of
'agent on a path toward a goal' in *going to* and modal nuances of desire or willingness
in *will*.

Does retention apply more to the newer variant *going to* than the older *will*, as
predicted by grammaticalization theory (e.g. Bybee & Pagliuca 1987:116)? The analysis
of variation between *will* and *going to* shows that both forms are highly grammati-
calized. Effects associated with early stages of grammaticalization (in particular, an-
imacy or agency of the subject and the semantic class of the lexical verb) have largely
dissipated. *Going to* has (incipiently) generalized to contingent (nonasserted) *if* and
when clauses. Thus, we have shown that the newer variant is not necessarily the least
developed (more semantic retention, less bleaching). Instead, as both constructions
undergo grammaticalization in the same grammatical domain, each has held on to
particular constructions. Such niches help explain continued variation between two or
more constructions with the same grammatical function.

An important finding of this study is the extent to which particular collocations
contribute to the observed distribution. These findings add to an accumulating body of
work that points to a greater role for the lexicon in grammatical variation than is
generally acknowledged. For example, Poplack's (1992) work on variation in Canadian
French shows that one verb, *falloir* 'to have to', exerts an enormous effect on the choice
of verbal morphology in subjunctive contexts. Similarly, work on African American
English shows that the verb *come* accounts in large part for unmarked past tense in
irregular verbs (Poplack & Tagliamonte 2001:134–36, van Herk & Poplack 2003:
256–59). If we extend the notion of the lexicon to phrasal configurations (cf. Jackendoff
1997) or 'prefabricated' chunks (Bybee 2002), we find similar effects. Our work on
other variables shows that a great deal of variable *that*-deletion can be attributed to
frequent subject-verb collocations such as *I think* and *I guess* (Torres Cacoullos &
Walker 2009) and that a large portion of variable subject-verb agreement in existentials
is accounted for by the construction *there's* (Walker 2007).

In the findings presented in this article, we argue for a model of grammar in which
speakers have available a number of stored, or lexicalized, collocations and construc-
tions of differing levels of abstraction and productivity—including fixed expressions
(*What's gonna happen?*) as well as ones with open slots (*X'll never* . . .). These colloca-
tions interact with general patterns, contributing to the structure of variation.[48] In other
words, grammatical domains are contoured by lexical material, where stored material
includes collocations. We can think of grammar as a system shaped by, and interacting
with, a patchwork quilt of particular constructions, or in Bolinger's (1976:1) terms, as
'a jerry-built structure' that may be 'heterogeneous **but** tightly organized' (emphasis
in original).

A usage-based approach to grammar predicts that while grammatical constructions,
such as those used to convey the future, arise via crosslinguistic grammaticalization
paths, some distributions may well be language-particular (cf. Bybee 2008, Slobin
1997). This question awaits studies of variation that examine collocations in other

---

[48] Smirniotopoulos and Joseph (1998:486) argue that forms that may be viewed as the output of a productive
morphological process may still be stored, in a system with 'a dynamically evolving border' between rules
in syntax and those in the lexicon.

language varieties, based on variable-rule analysis, cross-tabulation, and consideration of lexical effects (cf. Poplack & Tagliamonte 2001:141).

So how may we explain the crosslinguistically common 'cluttering' of the grammar with more than one way of expressing future? The inherent modality of the future, which constitutes a prediction by the speaker, may be intrinsically conducive to the development of associations between different future constructions and different speaker attitudes. But our findings further suggest a generalization for layers of grammaticalization in particular languages.

When more than one construction develops within the same grammatical domain, there remain niches in which one or the other is more likely to be conventionalized, regardless of its degree of grammaticalization. Continued variation between different general tense/aspect/mood constructions (*will* vs. *going to* vs. futurate Present) is largely sustained and shaped by particular constructions. Our results show that, rather than conveying a general invariant meaning difference between the variant forms, particular constructions with one or the other variant may bear different nuances of meaning or define certain distribution patterns, even though the general construction is highly grammaticalized. Thus, we conclude that grammaticalization paths may be shaped by interaction with collocations or particular constructions. The partial retention of original territory represents not only 'semantic survival in extremis' (Bybee et al. 1994:16), but also, we submit, tenacious patterns of distribution inherited from once-meaningful associations.

REFERENCES

AARON, JESSI E. 2006. *Variation and change in Spanish future temporal expression*. Albuquerque: University of New Mexico dissertation.

BERGLUND, YLVA. 1997. Future in Present-day English: Corpus-based evidence on the rivalry of expressions. *ICAME Journal* 21.7–19.

BERGLUND, YLVA. 2005. *Expressions of future in present-day English: A corpus-based approach*. Uppsala: Acta Universitatis Upsaliensis.

BINNICK, RONALD. 1971. *Will* and *be going to*. *Chicago Linguistic Society* 7.40–53.

BLAS ARROYO, JOSÉ LUIS. 2008. The variable expression of future tense in Peninsular Spanish: The present (and future) of inflectional forms in the Spanish spoken in a bilingual region. *Language Variation and Change* 20.85–126.

BOLINGER, DWIGHT. 1976. Meaning and memory. *Forum Linguisticum* 1.1–14.

BOLINGER, DWIGHT. 1977. *Meaning and form*. London: Longman.

BRISARD, FRANK. 1997. The English tense-system as an epistemic category: The case of futurity. *Lexical and syntactical constructions and the construction of meaning: Proceedings of the bi-annual ICLA meeting in Albuquerque, July 1996*, ed. by Eve Sweetser, Kidong Lee, and Marolijn Verspoor, 271–85. Amsterdam: John Benjamins.

BRISARD, FRANK. 2001. *Be going to*: An exercise in grounding. *Journal of Linguistics* 37.251–85.

BYBEE, JOAN. 2001. Main clauses are innovative, subordinate clauses are conservative: Consequences for the nature of constructions. *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*, ed. by Joan Bybee and Michael Noonan, 1–17. Amsterdam: John Benjamins.

BYBEE, JOAN. 2002. Sequentiality as the basis of constituent structure. *The evolution of language from pre-language*, ed. by Talmy Givón and Bertram Malle, 109–32. Amsterdam: John Benjamins.

BYBEE, JOAN. 2006. From usage to grammar: The mind's response to repetition. *Language* 82.711–33.

BYBEE, JOAN. 2008. Grammaticization: Implications for a theory of language. *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*, ed. by Jiansheng Guo, Elena Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura, and Seyda Özçaliskan, 345–56. London: Psychology Press.

BYBEE, JOAN, and PAUL J. HOPPER (eds.) 2001. *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.

BYBEE, JOAN, and WILLIAM PAGLIUCA. 1987. The evolution of future meaning. *Papers from the seventh International Conference on Historical Linguistics*, ed. by Anna Giacalone Ramat, Onofrio Carruba, and Giuliano Bernini, 109–22. Amsterdam: John Benjamins.

BYBEE, JOAN; REVERE PERKINS; and WILLIAM PAGLIUCA. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

BYBEE, JOAN, and RENA TORRES CACOULLOS. 2009. The role of prefabs in grammaticization: How the particular and the general interact in language change. *Formulaic language, vol. 1: Distribution and historical change*, ed. by Roberta L. Corrigan, Edith A. Moravcsik, Hamid Ouali, and Kathleen Wheatley, 187–217. Amsterdam: John Benjamins.

CAMERON, RICHARD. 1993. Ambiguous agreement, functional compensation, and nonspecific *tú* in the Spanish of San Juan, Puerto Rico, and Madrid, Spain. *Language Variation and Change* 5.305–34.

CAMPBELL, LYLE (ed.) 2001. *Grammaticalization: A critical assessment*. (Special issue of *Language Sciences* 23.)

CHESHIRE, JENNY. 2005. Syntactic variation and spoken language. *Syntax and variation: Reconciling the biological and the social*, ed. by Leonie Cornips and Karen P. Corrigan, 81–106. Amsterdam: John Benjamins.

COMRIE, BERNARD. 1976. *Aspect*. Cambridge: Cambridge University Press.

COMRIE, BERNARD. 1985. *Tense*. Cambridge: Cambridge University Press.

CROFT, WILLIAM. 2009. The origins of grammaticalization in the verbalization of experience. *Linguistics*, to appear.

DAHL, ÖSTEN. 1984. Temporal distance: Remoteness distinctions in tense-aspect systems. *Explanations for language universals*, ed. by Brian Butterworth, Bernard Comrie, and Östen Dahl, 105–22. Berlin: Mouton.

DAHL, ÖSTEN. 1985. *Tense and aspect systems*. Oxford: Blackwell.

DANCHEV, ANDREI, and MERJA KYTÖ. 1994. The construction *be going to* + infinitive in Early Modern English. *Studies in Early Modern English*, ed. by Dieter Kastovsky, 59–77. Berlin: Mouton de Gruyter.

EMIRKANIAN, LOUISETTE, and DAVID SANKOFF. 1985. Le futur simple et le futur périphrastique dans le français parlé. *Les tendances dynamiques du français parlé à Montréal*, ed. by Monique Lemieux and Henrietta Cedergren, 189–204. Québec: Gouvernement du Québec.

ERMAN, BRITT, and BEATRICE WARREN. 2000. The idiom principle and the open choice principle. *Text* 20.29–62.

ERVIN-TRIPP, SUSAN. 1976. Is Sybil there? The structure of some American English directives. *Language in Society* 5.25–66.

FILLMORE, CHARLES. 1968. The case for Case. *Universals in linguistic theory*, ed. by Emmon Bach and Robert Harms, 1–88. New York: Holt, Rinehart and Winston.

FLEISCHMAN, SUZANNE. 1982. *The future in thought and language: Diachronic evidence from Romance*. Cambridge: Cambridge University Press.

FLEISCHMAN, SUZANNE. 1989. Temporal distance: A basic linguistic metaphor. *Studies in Language* 13.1–50.

FOX, BARBARA A., and SANDRA A. THOMPSON. 1990. A discourse explanation of the grammar of relative clauses in English conversation. *Language* 66.297–316.

FREED, ALICE. 1994. The form and function of questions in informal dyadic conversation. *Journal of Pragmatics* 21.621–44.

FRIES, CHARLES C. 1927. The expression of the future. *Language* 3.87–95.

GEE, JULIE, and ISKENDER SAVASIR. 1985. On the use of *will* and *gonna*: Toward a description of activity-types for child language. *Discourse Processes* 8.143–75.

GIVÓN, TALMY. 1979. *On understanding grammar*. New York: Academic Press.

GOLDBERG, ADELE E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

GRUBER, JEFFREY. 1976. *Lexical structure in syntax and semantics*. New York: North Holland.

HAEGEMAN, LILIANE. 1989. *Be going to* and *will*: A pragmatic account. *Journal of Linguistics* 25.291–317.

HEDIN, EVA. 2000. Future marking in conditional and temporal clauses in Greek. *Tense and*

aspect in the languages of Europe, ed. by Östen Dahl, 329–49. Berlin: Mouton de Gruyter.

HEINE, BERND. 2003. Grammaticalization. *The handbook of historical linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 575–601. Malden, MA: Blackwell.

HEINE, BERND, and TANIA KUTEVA. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

HOPPER, PAUL J. 1991. On some principles of grammaticalization. *Approaches to grammaticalization*, ed. by Elisabeth C. Traugott and Bernd Heine, 17–35. Amsterdam: John Benjamins.

HOPPER, PAUL J., and SANDRA A. THOMPSON. 1980. Transitivity in grammar and discourse. *Language* 56.251–99.

HOPPER, PAUL J., and ELISABETH C. TRAUGOTT. 2003. *Grammaticalization*. 2nd edn. Cambridge: Cambridge University Press.

JACKENDOFF, RAY. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.

JACKENDOFF, RAY. 1997. *The architecture of the language faculty*. Cambridge, MA: MIT Press.

JESPERSEN, OTTO. 1924. *The philosophy of grammar*. London: Allen & Unwin.

JOOS, MARTIN. 1964. *The English verb: Form and meanings*. Madison: University of Wisconsin Press.

JOOS, MARTIN. 1968. *The English verb: Form and meanings*. 2nd edn. Madison: University of Wisconsin Press.

JOSEPH, BRIAN D. 1990 [1978]. *Morphology and universals in syntactic change: Evidence from medieval and Modern Greek*. New York: Garland.

JOSEPH, BRIAN D. 2001. Historical linguistics. *The handbook of linguistics*, ed. by Mark Aronoff and Janie Rees-Miller, 105–29. Malden, MA: Blackwell.

KLINGE, ALEX. 1993. The English modal auxiliaries: From lexical semantics to utterance interpretation. *Journal of Linguistics* 29.315–57.

KRUG, MANFRED G. 2000. *Emerging English modals: A corpus-based study of grammaticalization*. Berlin: Mouton de Gruyter.

KRUG, MANFRED G. 2001. Frequency, iconicity, categorization: Evidence from emerging modals. In Bybee & Hopper, 309–35.

KURYŁOWICZ, JERZY. 1947. La nature des procès dites analogiques. *Acta Linguistica* 5.15–37.

LABOV, WILLIAM. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45.715–62.

LABOV, WILLIAM. 1972a. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

LABOV, WILLIAM. 1972b. The transformation of experience in narrative syntax. *Language in the inner city*, 354–405. Philadelphia: University of Pennsylvania Press.

LABOV, WILLIAM. 1984. Field methods of the project on linguistic change and variation. *Language in use: Readings in sociolinguistics*, ed. by John Baugh and Joel Sherzer, 28–53. Englewood Cliffs, NJ: Prentice Hall.

LABOV, WILLIAM. 1994. *Principles of linguistic change, vol. 1: Internal factors*. Oxford: Blackwell.

LABOV, WILLIAM. 2001. *Principles of linguistic change, vol. 2: Social factors*. Malden, MA: Blackwell.

LABOV, WILLIAM. 2004. Quantitative reasoning in linguistics. *Sociolinguistics/Soziolinguistik: An international handbook of the science of language and society*, vol. 1, 2nd edn., ed. by Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier, and Peter Trudgill, 6–22. Berlin: Mouton de Gruyter.

LABOV, WILLIAM; PAUL COHEN; CLARENCE ROBINS; and JOHN LEWIS. 1968. *A study of the non-standard English of Negro and Puerto Rican speakers in New York City*. (Cooperative research report 3288, vol. 1.) Philadelphia: US Regional Survey.

LABOV, WILLIAM, and JOSHUA WALETZKY. 1967. Narrative analysis: Oral versions of personal experience. *Essays on the verbal and visual arts*, ed. by June Helm, 12–44. Seattle: University of Washington Press.

LANGACKER, RONALD W. 1987. *Foundations of cognitive grammar*. Stanford, CA: Stanford University Press.

LANGACKER, RONALD W. 1990. Subjectification. *Cognitive Linguistics* 1.5–38.

LANGACKER, RONALD W. 2002. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: Mouton de Gruyter.

LEECH, GEOFFREY. 1971. *Meaning and the English verb*. London: Longman.

MAIR, CHRISTIAN. 1997. The spread of the *going-to* future in written English: A corpus-based investigation into language change in progress. *Language history and linguistic modeling*, ed. by Raymond Hickey and Stanislaw Puppel, 1537–43. Berlin: Mouton de Gruyter.

MALKIEL, YAKOV. 1968. The inflectional paradigm as an occasional determinant of sound change. *Directions for historical linguistics: A symposium*, ed. by Winfred P. Lehmann and Yakov Malkiel, 21–64. Austin: University of Texas Press.

MALKIEL, YAKOV. 1973–74. New problems in Romance interfixation (1): The velar insert in the present tense (with an excursus on *-zer/-zir* verbs). *Romance Philology* 27.304–55.

McELHINNY, BONNIE S. 1993. Copula and auxiliary contraction in the speech of White Americans. *American Speech* 68.371–99.

MYHILL, JOHN. 1992. *Typological discourse analysis: Quantitative approaches to the study of linguistic function*. Oxford: Blackwell.

MYHILL, JOHN. 1994. Intuitive meanings and technical definitions: An analysis of *will* and *gonna*. Ann Arbor: University of Michigan, MS.

NEWMEYER, FREDERICK J. 1998. *Language form and language function*. Cambridge, MA: MIT Press.

NICOLLE, STEVE. 1998. *Be going to* and *will*: A monosemous account. *English Language and Linguistics* 2.223–43.

PALMER, FRANK R. 1965. *A linguistic study of the English verb*. London: Longman.

PALMER, FRANK R. 1974. *The English verb*. London: Longman.

PALMER, FRANK R. 1983. Future time reference in the conditional protasis: A comment on Comrie. *Australian Journal of Linguistics* 3.241–43.

PALMER, FRANK R. 1990. *Modality and the English modals*. London: Longman.

PAOLILLO, JOHN C. 2002. *Analyzing linguistic variation: Statistical models and methods*. Stanford, CA: CSLI Publications.

PAPPAS, PANAYIOTIS A. 2001. The microcosm of a morphological change: Variation in *thelō* + infinitive futures and *ēthela* + infinitive counterfactuals in Early Modern Greek. *Diachronica* 18.59–92.

POPLACK, SHANA. 1992. The inherent variability of the French subjunctive. *Theoretical analyses in Romance linguistics*, ed. by Christiane Laeufer and Terrell A. Morgan, 235–63. Amsterdam: John Benjamins.

POPLACK, SHANA, and ELISABETE MALVAR. 2007. Elucidating the transition period in linguistic change: The expression of the future in Brazilian Portuguese. *Probus* 19.121–69.

POPLACK, SHANA, and SALI TAGLIAMONTE. 1999. The grammaticalization of *going to* in (African American) English. *Language Variation and Change* 11.315–42.

POPLACK, SHANA, and SALI TAGLIAMONTE. 2001. *African American English in the diaspora: Tense and aspect*. Oxford: Blackwell.

POPLACK, SHANA, and DANIELLE TURPIN. 1999. Does the FUTUR have a future in (Canadian) French? *Probus* 11.133–64.

POPLACK, SHANA; JAMES A. WALKER; and REBECCA MALCOLMSON. 2006. An English 'like no other'? Language contact and change in Quebec. *Canadian Journal of Linguistics* 51.185–213.

QUIRK, RANDOLPH; SIDNEY GREENBAUM; GEOFFREY LEECH; and JAN SVARTVIK. 1985. *A comprehensive grammar of contemporary English*. New York: Longman.

ROY, JOSEPH. 2007. Peering into the future: The emergence of *going to* in the 18th century. Paper presented at New Ways of Analyzing Variation 36, University of Pennsylvania, Philadelphia.

ROYSTER, JANE F., and JOHN M. STEADMAN. 1923/1968. The 'going to' future. *Manly anniversary studies in languages and literature*, 394–403. Freeport, NY: Books for Library Press.

SANKOFF, DAVID. 1988a. Sociolinguistics and syntactic variation. *Linguistics: The Cambridge survey*, ed. by Frederick J. Newmeyer, 140–61. Cambridge: Cambridge University Press.

SANKOFF, DAVID. 1988b. Variable rules. *Sociolinguistics/Soziolinguistik: An international handbook of the science of language and society*, ed. by Ulrich Ammon, Norbert Dittmar, and Klaus J. Mattheier, 984–97. Berlin: Walter de Gruyter.

SANKOFF, DAVID; SALI TAGLIAMONTE; and ERIC SMITH. 2005. *GoldVarb X: A variable rule application for Macintosh and Windows*. Ottawa: Department of Mathematics, University of Ottawa, and Toronto: Department of Linguistics, University of Toronto.

SANKOFF, DAVID, and PIERRETTE THIBAULT. 1981. Weak complementarity: Tense and aspect in Monteal French. *Syntactic change: Natural language studies*, ed. by Brenda B. Johns and David R. Strong, 205–16. Ann Arbor: University of Michigan.

SANKOFF, GILLIAN, and PENELOPE BROWN. 1976. The origins of syntax in discourse: A case study of Tok Pisin relatives. *Language* 52.631–66.

SEDANO, MERCEDES. 1994. El futuro morfológico y la expresión *ir a* + infinitivo en el español hablado de Venezuela. *Verba: Anuario galego de filoloxia* 21.225–40.

SILVA-CORVALÁN, CARMEN, and TRACY TERRELL. 1989. Notas sobre la expresión de futuridad en el español del Caribe. *Hispanic Linguistics* 2.191–208.

SLOBIN, DAN ISAAC. 1997. The origins of grammaticizable notions: Beyond the individual mind. *The cross-linguistic study of language acquisition, vol. 5: Expanding the contexts*, ed. by Dan Isaac Slobin, 1–39. Mahwah, NJ: Lawrence Erlbaum.

SMIRNIOTOPOULOS, JANE C., and BRIAN D. JOSEPH. 1998. Syntax versus the lexicon: Incorporation and compounding in Modern Greek. *Journal of Linguistics* 34.447–88.

SMITH, CARLOTA. 1997. *The parameter of aspect*. 2nd edn. Dordrecht: Kluwer.

SZMRECSANYI, BENEDIKT. 2003. *Be going to* versus *will/shall*: Does syntax matter? *Journal of English Linguistics* 31.295–323.

TAGLIAMONTE, SALI A. 1997. Obsolescence in the English perfect? Evidence from Samaná English. *American Speech* 72.33–68.

TAGLIAMONTE, SALI A. 2002. Comparative sociolinguistics. *Handbook of language variation and change*, ed. by J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes, 729–63. Oxford: Blackwell.

TAGLIAMONTE, SALI A. 2006a. 'So cool, right?': Canadian English entering the 21st century. *Canadian English in a global context* (Special issue of *Canadian Journal of Linguistics* 51), ed. by Peter Avery, Alexandra D'Arcy, Elaine Gold, and Keren Rice, 309–31.

TAGLIAMONTE, SALI A. 2006b. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press.

TAGLIAMONTE, SALI A., and DEREK DENIS. 2008. Linguistic ruin? LOL! Instant messaging, teen language and linguistic change. *American Speech* 83.3–34.

THOMPSON, SANDRA A. 2002. 'Object complements' and conversation. *Studies in Language* 26.125–64.

THOMPSON, SANDRA A., and ANTHONY MULAC. 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. *Approaches to grammaticalization*, vol. 2, ed. by Bernd Heine and Elizabeth C. Traugott, 313–29. Amsterdam: John Benjamins.

TOLLER, THOMAS NORTHCOTE. 1921. *An Anglo-Saxon dictionary*. Supplement. Oxford: Oxford University Press.

TOMASELLO, MICHAEL. 2003. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

TORRES CACOULLOS, RENA. 1999. Variation and grammaticalization in progressives: Spanish *-ndo* constructions. *Studies in Language* 23.25–59.

TORRES CACOULLOS, RENA. 2000. *Grammaticization, synchronic variation, and language contact: A study of Spanish progressive* -ndo *constructions*. Amsterdam: John Benjamins.

TORRES CACOULLOS, RENA. 2001. From lexical to grammatical to social meaning. *Language in Society* 30.443–78.

TORRES CACOULLOS, RENA. 2009. Variation and grammaticisation: The emergence of an aspectual opposition. *Selected papers from the 4th International Conference on Language Variation in Europe (ICLaVE 4), University of Cyprus, 17–19 June 2007*, ed. by Stavroula Tsiplakou, Marilena Karyolemou, and Pavlos Pavlou, 215–24. Amsterdam: John Benjamins.

TORRES CACOULLOS, RENA, and SCOTT A. SCHWENTER. 2005. Towards an operational notion of subjectification. *Berkeley Linguistics Society* 31.47–58.

TORRES CACOULLOS, RENA, and JAMES A. WALKER. 2009. On the persistence of grammar in discourse: A variationist study of *that*. *Linguistics* 47.1–43.

TOTTIE, GUNNEL. 2002. Non-categorical differences between American and British English: Some corpus evidence. *Studies of mid-Atlantic English*, ed. by Marko Modiano, 37–58. Gävle: University of Gävle Press.

TRAUGOTT, ELIZABETH CLOSS. 1989. On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65.31–55.

TRAUGOTT, ELIZABETH CLOSS. 1995. Subjectification in grammaticalization. *Subjectivity and subjectivisation*, ed. by Dieter Stein and Susan Wright, 31–54. Cambridge: Cambridge University Press.

TRAUGOTT, ELIZABETH CLOSS. 2003. Constructions in grammaticalization. *The handbook of historical linguistics*, ed. by Brian D. Joseph and Richard D. Janda, 624–47. Oxford: Blackwell.

TRAUGOTT, ELIZABETH CLOSS, and RICHARD DASHER. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.

TRAUGOTT, ELIZABETH CLOSS, and BERND HEINE. 1991. *Approaches to grammaticalization*. Amsterdam: John Benjamins.

TRAVIS, CATHERINE E. 2007. Genre effects on subject expression in Spanish: Priming in narrative and conversation. *Language Variation and Change* 19.101–35.

VAN HERK, GERARD, and SHANA POPLACK. 2003. Rewriting the past: Bare verbs in the Ottawa Repository of Early African American Correspondence. *Journal of Pidgin and Creole Languages* 18.231–66.

VERHAGEN, ARIE. 1995. Subjectification, syntax, and communication. *Subjectivity and subjectivisation: Linguistic perspectives*, ed. by Dieter Stein and Susan Wright, 103–28. Cambridge: Cambridge University Press.

VISSER, FREDERIC THEODOR. 1970. *An historical syntax of the English language, part 1: Syntactical units with one verb*. Leiden: E. J. Brill.

WALKER, JAMES A. 2000. Prosodic optimality and variability in English auxiliaries. *McGill Working Papers in Linguistics* 15.105–19.

WALKER, JAMES A. 2001. Using the past to explain the present: Tense and temporal reference in Early African American English. *Language Variation and Change* 13.1–35.

WALKER, JAMES A. 2007. 'There's bears back there': Plural existentials and vernacular universals in (Quebec) English. *English World-Wide* 28.147–66.

WALKER, JAMES A.; SHANA POPLACK; and RENA TORRES CACOULLOS. 2004. Looking into the future in English and French. The role of minority status in language contact and change: The past and the future in Canada. Paper presented at Sociolinguistics Symposium 15, University of Newcastle, UK.

WEINREICH, URIEL; WILLIAM LABOV; and MARVIN HERZOG. 1968. Empirical foundations for a theory of language change. *Directions for historical linguistics*, ed. by Winfred P. Lehmann and Yakov Malkiel, 95–188. Austin: University of Texas Press.

WEKKER, HERMAN. 1976. *The expression of future time in contemporary British English*. Amsterdam: North Holland.

WHYTE, JOHN. 1949. *A personal adventure into the German and English future*. NeuZs I.

WINFORD, DONALD. 1998. On the origins of African American Vernacular English—A creolist perspective, part 2: Linguistic features. *Diachronica* 15.99–154.

Torres Cacoullos   [Received 16 November 2007;
Department of Spanish, Italian and Portuguese   revision invited 3 August 2008;
The Pennsylvania State University   revision received 1 November 2008;
231 Burrowes Building   accepted 4 January 2009]
University Park, PA 16802
[rena@psu.edu]

Walker
Department of Languages, Literatures and Linguistics
York University
4700 Keele St.
Toronto, Ontario
M3J 1P3
Canada
[jamesw@yorku.ca]