

Szmrecsanyi, B. (2003). BE GOING TO versus WILL/SHALL: Does Syntax Matter? *Journal of English Linguistics*, 41, 295-323.

Introduction:

-There are two syntactic options for marking future tense in English: BE GOING TO and WILL/SHALL

-Previous studies have dealt with the difference between the two markers as semantic or pragmatic. The research states that they vary stylistically, regionally, and sociolinguistically.

-It has been shown that in informal conversation, speakers tend to use contracted or cliticized future marker variants, and the paradigm BE GOING TO instead of WILL/SHALL.

-BE GOING TO has been shown to be more frequent in American English than British English.

-Regardless of the register and regional variety, research has shown that the paradigm WILL/SHALL is more frequent than BE GOING TO.

-A systematic investigation of how the syntactic environment can affect the choice of future marker is absent from the literature.

-The present study attempts to fill this gap in the literature by answering:

1. Are there differences in the frequencies of the future markers? Do they differ between American English and British English? Do they differ between formal and informal spoken English?

2. Is the paradigm GOING TO preferred in contexts of negation?

3. Are there differences in the frequencies of BE GOING TO and WILL/SHALL depending on whether they are in syntactically dependent or independent clauses?

4. Is there a restriction on future markers in IF-clauses?

-And, are there differences between registers and varieties in each of these syntactic environments?

5. Is there a correlation between sentence length and the likelihood of the occurrence of a given future marker?

Method and Data:

-BE GOING TO paradigm includes: *going to* and *gonna*

→included past tense forms, but excluded spatial uses

-WILL/SHALL paradigm includes: *will*, *won't* (negated), *'ll* (cliticized will), and *shall*

→excluded non-future usages of *will*

-Three computerized corpora of spoken English were used:

1. British National Corpus (BNC) which is subdivided into an informal section (DS) and a formal section (CG)
 - The two sections are treated as separate corpora.
2. Santa Barbara Corpus of Spoken American English (CSAE) used to represent informal American English
 - contains only 14 conversations with 51 speakers
3. Corpus of Spoken Professional American English used to represent formal American English
 - the contracted form *gonna* is not present in this corpus

Results and Discussion:

Overall Frequency:

	Informal Speech	Formal Speech
BE GOING TO	- <i>gonna</i> is more frequent than <i>going to</i> -paradigm is more frequent in informal corpora	- <i>going to</i> (in CG) is more frequent than <i>gonna</i>
WILL/SHALL	- <i>will</i> is less frequent than the contracted forms - <i>won't</i> is more frequent in informal corpora - <i>shall</i> seems to be more frequent in informal corpora	- <i>will</i> is more frequent than the contracted variants

-The WILL/SHALL paradigm outnumbered the BE GOING TO paradigm

	American English	British English
BE GOING TO	- <i>gonna</i> is most frequently found in informal American English -BE GOING TO is more frequent in American corpora (in both formal and informal)	
WILL/SHALL		-the cliticized 'll is more frequent in the British English corpora - <i>shall</i> is less marginal (but still marginal) - <i>won't</i> is more frequent in British English corpora

-Future marker choice is affected by both region and register.

→ The informal corpora contain lower percentages of non-contracted future marker forms than the formal corpora.

Contexts of Negation:

	American English	British English
BE GOING TO	-preferred paradigm for negation	
WILL/SHALL		- <i>won't</i> is the most frequent negated marker - <i>'ll not</i> is infrequent -preferred paradigm for negation

-Contexts of negation affect the distribution of the future markers

-Variety of English has an effect

Syntactically Independent/Dependent Environments:

-The difference between dependent and independent clauses is statistically significant in the British English corpora.

	Independent environments	Dependent environments
BE GOING TO		- <i>going to</i> and <i>gonna</i> are more frequent in dependent clauses
SHALL/WILL	-paradigm is more frequent in independent clauses	- <i>won't</i> and <i>'ll</i> are less frequent in dependent clauses - <i>will</i> is slightly more frequent in dependent clauses in all corpora

IF-Clauses:

	Main clause of an If-subclause	If-subclause
BE GOING TO	-infrequent in main clauses except in CSAE (formal American English)	-paradigm is more frequent in If-subclauses
WILL/SHALL	-preferred paradigm (except in CSAE)	

-The distribution of the future markers are sensitive to IF-clause environments.

Sentence Length:

-This was measured in words.

-A sentence is defined as the words occurring between two punctuation marks.

-Nonverbal material (um, uh, oh, er) was not counted.

-The future markers were excluded from the word count, as they differ in length.

-Data from the CG (formal) section from the BNC was not included in this section

-There is a consistent tendency for BE GOING TO to occur in longer sentences than WILL/SHALL. But, this is only significant in the CSPAE (formal American English).

-There is a correlation between sentence length and the future marker employed.

Conclusion:

-The choice between the two future marker paradigm is affect by:

→register

→variety of English

→syntactic environment

-The author concludes that the study suggests that long, subordinated, and syntactically complex environments lead speakers to use BE GOING TO instead of WILL/SHALL.

Critique:

-The author does not exclude non-future usages of *shall* due to the low frequency of *shall*. However, he does exclude the non-future usages of *will* and *be going to*. Even though *shall* has a low frequency, this could lead to an overestimation of the frequency of *shall* as a variant of the future marking WILL/SHALL paradigm.

-The author treats If-clauses separately from independent and dependent clauses. However, as IF-clauses are a type of independent clause, they should be treated as related to the dependent clauses. This could be achieved by examining them as a subtype of dependent clauses.

-The author defines 'sentence' as the words occurring between two punctuation marks (but not commas). That is, he is treating an utterance as a sentence. In syntax, a clause is a sentence. Therefore, in the cases where the author claims to have a complex sentence composed of two clauses, he really has two sentences.

-Moreover, in the case of a 'sentence' that contains two clauses, the author does not explain how the length of the sentence will be counted if there are two future markers. That is, if there is a future marker in each clause, how is the length of the given 'sentence' counted, and how is the length associated with each future marker. This leads to an issue with accountability and the ability to replicate the study. Without, an explanation of how these cases are handled, it is not possible to replicate the study.

-There is another problem with accountability and the ability to replicate the study. The author does not include the CG (formal) section from the BNC in the sentence length calculations. He does not explain why this is done.

Discussion Questions:

Szmrecsanyi notes that there are “minor” shortcomings in the American corpora used in this study. That is, the CSAE was a small, and the CSPAE lacked the variant *gonna*. Do you think that this poses a problem for the results of this study? How?

Non-future marker *shall* was not excluded from the data due to low frequency. Do you think this could have affected the results of the study? Why or why not?

Sentence length was used in this study to attempt to operationalize syntactic complexity. Do you think that it is a reliable measure? Why or why not? And, if not, what would you use instead?