

Are Complex Tenses Really Real?

Martin Edwardes

Kings College London & City of London Academy (Southwark)

martin.edwardes@btopenworld.com

CONTENTS

ABSTRACT	1
DESCRIBING TENSES	1
THE FIRST STUDY	3
THE SECOND STUDY.....	4
THE THIRD STUDY	5
THE ISSUES.....	7
REFERENCES.....	8

ABSTRACT

The future perfect indicative tense, represented in English by the composite form *will have been going to [verb]*, expresses a complex temporal relationship. It is formed from two sub-tenses: the future perfect (*will have been*), which places the point from which the action is viewed into the future, and the action itself in the past relative to that future; and a variant of the future simple tense (*be going to [verb]*) which places the action into the future from a viewpoint of the present.

Halliday & Matthiessen (2004) list the tenses generated by *will have been going to* as: 24 (future of future in past, *will have been going to take*); 27 (future of past in future in past, *will have been going to have taken*); 33 (future of present in future in past, *will have been going to be taking*); and 36 (future of present in past in future in past, *will have been going to have been taking*). Clearly, *will have been going to* is a construct with complex semanticity, and therefore limited application; but is it a tense (or group of tenses) that is actually used?

This paper looks at three studies of this group of tenses, using the Internet as a corpus. It also looks at a range of other complex tenses listed by Halliday & Matthiessen. The results show that the range of usages of these complex forms is extremely limited. The study raises questions about the value of these constructs as real markers of tense, and about the teaching of the constructs to both first- and second-language users.

DESCRIBING TENSES

English, because of its analytic nature, can be highly expressive; and this is visible in the way verb tenses are constructed. For instance, "I will have been going to have been about to begin having done" only scratches the surface of what is possible. However, being able to build a baroque language structure does not imply that there is a use for it; and complexity of possibility does not, in the case of tenses, convert to semanticity. Do we really understand what the complex tense form above is expressing?

If we look at the expression of time in language from different theoretical positions, we can see that there is some dispute as to how the grammar or grammars of language divide up time. Hornstein (1990) takes a formalist approach, using Reichenbach's (2005 [1947]) three-component description of tense. In this model, tense is a relationship between the

constant present, S, the time at which the event being described happens, E, and the relative point in time adopted as a viewpoint for the event, R. The simple present is a product of all three times co-occurring in the present, S; simple future and past are produced by moving the event point, E, out of the present; and the four complex tenses (past in past, future in past, past in future and future in future) are produced by also moving the viewpoint of the event, R, out of the present. The Reichenbachian analysis does not go beyond this, although it does suggest that multiple viewpoints are possible. If a second viewpoint were to be included then it would generate a further eight tenses, as the system is based on a binary progression (1+2+4+8 ...).

This analysis has been extended in Edwardes (2010) to include tense effects such as continuity (the extension of the event point through time), imminence (the closeness of the Reichenbachian points to each other), and completion (whether an event is being viewed as completed or ongoing). Comrie (1985) also raises the issue of relative tense, in which the present tense is used to represent events that are actually in the past or future (as in *Daddy's taking us to the zoo tomorrow*). In the Edwardes model these are treated as proximate tenses, with a viewpoint and event point located together in the past or future (see table 1).

Tense	Point of Speech (S)	Point of Reference (R)	Point of Event (E)
Simple Present		Present	
Simple Past		Present	Past
Simple Future		Present	Future
Past of Past	Present	Past	Past of (R)
Future of Past	Present	Past	Future of (R)
Past of Future	Present	Future	Past of (R)
Future of Future	Present	Future	Future of (R)
Proximate Past	Present		Past
Proximate Future	Present		Future

TABLE 1: The extended Reichenbach tense set

In contrast, the Halliday & Matthiessen (2004) model allows for much more complex tense structures to accommodate the possibilities of English; it therefore consists of a much larger set than the formalist model – 36 tenses are listed, compared to the seven of the Reichenbach model (or nine of the Edwardes model). The functionalist range of tenses may not have the language-nonspecific logical structure of the formalist range, but it is based on the real possibilities that a specific language offers.

If we compare some of the simpler tenses in the two models, we can see how they differ in their analysis. To aid comparison the sample verb of *take* will be used in the illustrations below. This is not necessarily the best verb to use as it is varied in meaning and therefore usage (*take umbrage, take a train, take possession, take time, take a cake, etc*), and it can be grammatically idiomatic (*I take it you already know, he takes and never gives*). It is, nonetheless, the sample verb used by Halliday and Matthiessen.

The basic tenses offered by the formalist and functionalist models differ mainly in their hierarchy. The formalist model offers a single base tense where the S, E and R timepoints are all in the present (*I take*), and two generated simple tenses of past and future where the E timepoint has been moved out of the present (*I took* and *I will take*). In the functionalist model, these form the three base tenses. At the next level, the formalist model offers four tenses, where the R timepoint has also been moved out of the present (*I had taken, I have been going to take, I will have taken, I will be going to take*). In contrast, the functionalist model offers nine tenses, as table 2 shows.

Tense	Tense no	Form	Reichenbachian equiv.
Past of past	4	I had taken	Past in past
Present of past	5	I have taken	Past (imminent)
Future of past	6	I will have taken	Past in future
Past of present	7	I was taking	Past (continuous)
Present of present	8	I am taking	Present (continuous)
Future of present	9	I will be taking	Future (continuous)
Past of future	10	I was going to take	Future in past (imminent)
Present of future	11	I am going to take	Future (imminent)
Future of future	12	I will be going to take	Future in future

TABLE 2: A comparison of Functionalism and Formalism tenses

As can be seen, there is considerable difference between the formalist and functionalist analyses about what some of the tenses signify in terms of temporality; and the formalist future in past tense (*I have been going to take*) seems to be missing from the list. This is given in the functionalist model as the present of future in past.

More fundamental differences occur with the remaining 24 functionalist English tenses, with the formalist analysis taking the view that they either express continuity, imminence or completion effects, or that they reflect an English-specific capacity; and, as has already been observed, there remains the question whether they are tenses with a real life, or just waxworks created by overzealous linguists.

In an attempt to review the usage question, I undertook three separate investigations, in 2005, 2008 and 2010. The search engine used for the first investigation was Alta Vista, so this was used for the other investigations, too. This was intended to give continuity to the studies; but, as both the search engine and the Internet have changed considerably over the period, this is unlikely by itself to have provided real continuity. Fortunately this does not cause serious problems in the diachronic aspect of this study. Any future Internet searches that I undertake, however, are likely to be conducted using Google, which is more user-friendly.

THE FIRST STUDY

The first data-mining exercise took place on 30 July 2005 (it is important to complete any one study within days – or, preferably, hours – to ensure that the constantly-changing Internet corpus has not changed too much). This exercise looked only at usages of the *will have been going to* form, which, being uninflected, was selected for ease of handling. A total of 73 cases were found: in 43 cases, *will have been* was the tense marker, with *going*

to as the verb form; in 26 cases *will have been going to* was the whole tense marker; and there was one case each of *will have been going to be* and *will have been going to have*, and two cases of *will have been going to have been* as tense markers. Even on the Internet, it would have been difficult at this time to produce a good statistical study of usage of these forms.

Form	Count	Korean website	Grammar related	Time trope	Legit. usage
Will have been [going to] [N]	43	11	15		17
Will have been going to [V]	26		18	5	3
Will have been going to be [V-ing]	1			1	0
Will have been going to have [V-ed]	1			1	0
Will have been going to have been [V-ing]	2			2	0

TABLE 3: Usages of *will have been going to*, 30 July 2005

However, some key features of usage did emerge. Of the 43 *will have been* cases, 11 occurred on a single Korean financial website, now gone, and seemed to be the product of a mistaken preference of a single ESL writer (the usages should all have been simple past, simple present or *have been X-ing*); and 15 occurred as grammatical examples on language learning websites. Of the 26 *will have been going to* cases, five were time tropes (using the tense to demonstrate temporal confusion time rather than to place an event in time); and 18 of the remaining 21 cases were used as examples on language learning websites (one of which was for Quenya Elven and one for "Timelord", presumably Gallifreyan). All of the four extended tenses were also time tropes. What was notable was the lack of unmarked legitimate usages, especially as the tenses became more complex.

THE SECOND STUDY

The second data mining exercise occurred over five days, from 26 to 30 December 2008. Once again, the exercise looked only at usages of *will have been going to*. A total of 207 cases were found, of which 147 were *will have been* cases, and 60 were *will have been going to* cases. Of the 60 *will have been going to* cases, three were *will have been going to be*, seven were *will have been going to have*, five were *will have been going to have been*, and one was *will have been going to be being*. This last was used on a grammar website and is a form not recognised by Halliday & Matthiessen. The writer of the web page clearly had problems explaining the meaning of this form, not recognising that he was probably describing a passive form.

Form	Count	Grammar related	Time trope	Legit. usage
Will have been [going to] [N]	147	38		109
Will have been going to [V]	44	17	7	20
Will have been going to be [V-ing]	3	1	2	0
Will have been going to have [V-ed]	7	2	5	0
Will have been going to have been [V-ing]	5	2	5	0
Will have been going to be being [V-ing]	1	1		0

TABLE 4: Usages of *will have been going to*, 26 to 30 December 2008

Table 4 gives a breakdown comparable to table 3. The Internet had grown over the three-and-a-half years, and the more complex forms had become more frequent. Unmarked legitimate usage had also increased, but had not yet invaded the complex forms. However, the presence of complex forms remains vanishingly small.

THE THIRD STUDY

The third data mining exercise, undertaken on 26/27 July 2010, was somewhat different to the others, in that it extended the study to include other complex tense forms from the Halliday & Matthiessen list: all the tense forms from 25 to 36 were reviewed. This became possible because the Internet had more than doubled in size in the 18 months between the second and third studies, finally creating a corpus in which rare forms can begin to be explored. Table 5 shows the estimated size of the Internet based on the relative frequency of some core words in the British National Corpus.

Word	Size of BNC	BNC Count	Alta Vista Count	Estimated A-V Corpus size
Against	100 million	55,164	2940 million	5,330 billion
Because		100,509	4740 million	4,716 billion
Between		90,638	4350 million	4,799 billion
However		59,690	2780 million	4,657 billion

TABLE 5: Estimates of the size of the Internet as at 27 July 2010 (www.natcorp.ox.ac.uk/)

Roughly, Alta Vista has a corpus size of 4.8 trillion English words. In 12pt Arial, this is a line of text stretching from here to the moon 125 times. More importantly, it is nearly 50,000 times the size of the BNC, indicating that the grammar forms reviewed here are unlikely to be statistically visible in the BNC.

The first exercise on the data from the third study was to produce an analysis which could be compared with the previous data sets. The much larger datasets have required some estimation – it was impossible to track every usage where counts exceeded 100. Instead, the first 100 occurrences of a form were sampled and used to estimate figures, based on the total count given for that form by Alta Vista (estimated figures are indicated by an entry in the *Est* column). This should not significantly affect the data for two reasons: first, most of the total counts are below 500, making 100 a strong sample; second, counts above 500 indicate forms which are less rare, so of less interest here.

Form	Est	Count	Grammar related	Time trope	Legit. usage
Will have been [going to] [N]	Est	375	40		335
Will have been going to [V]	Est	225	130	25	70
Will have been going to be [V-ing]		18	9	7	2
Will have been going to have [V-ed]		18	9	2	7
Will have been going to have been [V-ing]		20	15	5	0

TABLE 6: Usages of *will have been going to*, 26 to 27 July 2010

Table 6 is largely comparable with table 4, although four effects should be commented on. The first is the rise of legitimate usages, which now include some of the complex forms; the

second is the explosion in grammar-related usages of *will have been going to*, which is clearly the level at which most tense complexity is taught; and the third is the unexpected relative reduction in time tropes. A fourth effect, not obvious from the table, is that my abstract for this paper is now on the Internet, so the last four forms all include that abstract in the “grammar related” counts. This is an unanticipated consequence of the study which, because of the low counts, could be viewed as having a marked effect on the figures.

The third study also looked at all the complex tenses in the functional model – all the tenses which include four or more time points in their definition. These twelve tenses can be viewed as the “difficult” tenses, so their frequency and usages should have significance both for our understanding of tense as a meaning-sharing mechanism, and for our approach to teaching tenses to first- and second-language learners.

However, where *will have been going to* has a single inflection, five of the twelve complex tenses have more than one inflection, which complicates data collection. Interestingly, in two of the cases, *was/were going to have been* and *has/have been going to have*, the different inflections have significantly different distributions of usage. It seems that grammar sites have definite preferences for *was going to have been* and *have been going to have*; this is a little strange in that they share person only in first person singular, but the most commonly-used forms are third person singular and second person respectively. In one of the other cases (*Has/have been going to have been*) the numbers are too low to determine significance, and in the final two cases (*has/have been going to be* and *am/is/are going to have been*) there seem to be no significant differences between the inflections.

	FORM	Est	Count	Grammar related	Time trope	Porn Site	Legit. usage
24	Will have been going to [V]	Est	225	130	25		70
25	Had been going to have [V-ed]		12				12
26	Has been going to have [V-ed]	Est	130				130
26	Have been going to have [V-ed]		16	7			9
27	Will have been going to have [V-ed]		18	9	2		7
28	Was going to have been [V-ing]	Est	920	620			300
28	Were going to have been [V-ing]	Est	295				295
29	Am/is/are going to have been [V-ing]	Est	6,980	2,440			4,540
30	Will be going to have been [V-ing]		20	12	3		5
31	Had been going to be [V-ing]	Est	345				345
32	Has/have been going to be [V-ing]	Est	950	30	20		900
33	Will have been going to be [V-ing]		18	9	7		2
34	Had been going to have been [V-ing]		16	9	2	5	0
35	Has/have been going to have been [V-ing]		11	7	4		0
36	Will have been going to have been [V-ing]		20	15	5		0

TABLE 7: Usages of complex tenses, 26 to 27 July 2010

There is a large variation in frequency counts in table 7, a variation which does not correlate with either complexity of construct or length of form. Simply, some time relationships just seem to be more frequently expressed than others, which probably reflects their functional utility. The three most complex forms (34, 35, 36) have no

legitimate usages, so they are clearly linguistic toys rather than actual grammatical/semantic units; but even the less complex forms are relatively rare. If we conservatively say that the Internet 4.8 trillion word corpus includes 10 billion verb forms, then the approximately 10,000 cases above still constitutes a vanishingly small proportion of the corpus.

One final feature of table 7 should be remarked on. Since starting this exercise in 2005, the Internet has become a darker and less friendly place. One consequence of this is that “temptation” keywords are placed in site headers on porn and security scam web pages to encourage unsuspecting visitors to click on links. It appears that one phrase used on some of these sites is *had been going to have been*, which is why this form has generated an extra column, “Porn site”. Quite why this form is favoured for this purpose is unclear, but it does illustrate the point that using the Internet as a corpus needs to carry a health warning: don’t do it unless you have good virus protection and an effective firewall; understand the risks; and have a healthy mistrust of other Internet users.

THE ISSUES

The three studies discussed here raise important issues for pedagogy, especially in teaching English as a foreign language. The complex verb forms are seldom taught in British first-language English lessons at secondary level – not least because, at this level, English as an academic subject is mainly about encouraging reading comprehension and writing style; explicitly teaching grammar is a marginal activity. In Britain, secondary English is part of a long-established tradition of literature-qualified teachers teaching literature. It works as a self-generating system, with literature-qualified teachers generating the next generation of literature-qualified teachers. Skill in teaching grammatical complexity is, however, not a standard part of the literature-qualified teacher’s toolkit, nor is it part of the curriculum.

English as a foreign language is another matter, however. Explicit grammar teaching constitutes an important component of most courses; and, if the Internet grammar-related websites are representative, complex tenses are taught as a matter of routine. This raises the difficult question of how much of a language is needed to be a proficient user of that language. If first language users are not using a particular construct, no matter how valid it may appear, should we be bothering to teach it to second language users? An argument in favour of teaching complexity is that any form which is a logical possibility in a language represents something particular about the language; and knowing how that complex form works could unlock deeper knowledge of how the language works. An argument against is that teaching forms not actually used by native speakers does not help a student trying to achieve native-like proficiency. At the very least, therefore, we should make sure that we are not teaching these rare forms as active constructs.

At a simpler level, the question of what forms should be taught has to be matched to the level of language that the learner wants to achieve. If the learner wants to become competent in conversational English then the functionalist tenses 25 to 36 are irrelevant – if their use in writing is vanishingly small, they are likely to be even rarer in speech. Indeed, this argument can be extended to first-language learners as well, and it seems likely that failure or unwillingness to learn these tense outliers could be indicative of other personal

choices over first-language grammatical knowledge. People do not learn the totality of their language's grammar, either implicitly or explicitly; and often they rely on personal rule variants in their day-to-day language transactions. To be mutually comprehensible it is not necessary for two people to have the same grammar, even if they have the same official first language – just as it is unnecessary for them to have the same phonology or lexis.

Learning a language, whether a first or a second language, is a matter of becoming communication-competent; and this raises important issues in pedagogy about the difference between I-language (an individual's language system), E-language (the negotiation to common meaning in a linguistic exchange), and what could be called T-language (the totality of acceptable language forms existing within a community of minds). As long as the participants in a dialogue share a useful common subset of the T-language they can be viewed as mutually communication-competent. Their shared E-language works and their individual I-languages are irrelevant.

Perhaps we need to be more aware of the incredibly compressed semantic shorthand that any single-word name for a language constitutes; and, remembering that even first-language users do not necessarily fully share a common grammar, perhaps we should limit our teaching, whether for first or second languages, to forms which are statistically frequent in actual use. This paper does not address what those statistically frequent forms should be, and there remains much work to be done to generate a pedagogical grammar that reflects actual usage rather than what is possible; but it hopefully shows that, in the case of English tenses, we need to teach less than we can teach.

REFERENCES

- Robert de Beaugrande** (2001). Text Grammar Revisited. In *Logos and Languages – special issue*. At www.beaugrande.com/Revisited.htm (accessed 26 Jul 2010).
- Bernard Comrie** (1985). *Tense*. Cambridge University Press: Cambridge, UK.
- Martin Edwardes** (2010). What Time Tells Us about Being Human. In *The Origins of Grammar: an anthropological perspective*. Continuum: London, UK.
- Michael A K Halliday & Christian M I M Matthiessen** (2004). *An Introduction to Functional Grammar, third edition*. Arnold: London, UK. pp340-342.
- Norbert Hornstein** (1990). *As Time Goes By: tense and Universal Grammar*. MIT Press: Cambridge, Mass, USA.
- Hans Reichenbach** (2005 [1947]). The Tenses of Verbs. In Inderjeet Mani, James Pustejovsky, and Robert Gaizauskas (eds), *The Language of Time: a reader*. Oxford University Press: Oxford, UK.