EDITORIALS VS. FRONT PAGE NEWS: DIFFERENCES IN REALIS AND TENSE ACROSS NEWSPAPER GENRES

Mick O'Donnell

WagSoft Systems

1. Introduction[1]

While most papers in this volume address cross-language comparisons, the contrast between texts written in the same language is also a valid area of concern. Various areas of monolingual contrastive studies include: register studies (differences in text due to differences in context: field, tenor or mode); gender studies (differences in text due to gender); cross-temporal studies (differences in a text-type over time), etc.

This paper falls into the first of these categories: it offers a comparison between the registers of two newspaper text-types: Front Page News (FPN) and Editorials. In particular, it explores how these text-types differ in the way they use grammatical tense, and tries to explain these differences. Note, however, at another level, it is offering a methodology for contrastive studies in general.

Most contrastive register studies concern themselves with either lexical or grammatical differences between corpora. They attempt to explain formal choices directly in relation to context of use. This paper will argue that this approach is limited, proposing a methodology of register analysis which first relates context to the meanings expressed in the text, and then relating these meanings to the forms that express them. I argue that this three level approach allows a clearer view of the registerial patterns, patterns which are obscured if context is related directly to forms.

Traditional register analyses of tense attempt to correlate grammatical selection directly with the context. I will go beyond these studies by examining the selection of tense in the context of the semantics it expresses (the *realis* of the process), and how this mapping between realis and tense is conditioned by register. I will demonstrate that the major differences between the newspaper text-types is due to differences in the semantic selections (realis), and that grammatical preferences form only a secondary factor.

2. Traditional Register Analysis

Traditional register studies examine the relationship between a contextual variable and a grammatical or lexical pattern. For instance, Nesbitt & Plum (1987) explore differences in clause complexes across text-type, and Bernstein (1971) looks at the incidence of passives in relation to class. Others study lexical selection in relation to context.

In relation to grammatical tense, there have been several studies that have tried to explain patterns of grammatical selection directly in terms of contextual variables. Plum & Cowling (1987) studied the correlation of social class, gender, and age with tense selection. They found that, for instance, use of past-tense (primary tense) increases with both age and rising social class. Halliday & James (1993) also looks at the variation of tense selection over differing registers.

To demonstrate what a single-level study of tense looks like, I will exemplify using my corpus. Table 1 shows the distribution of tense selections in two newspaper article-types. Each row of the table shows the percentage occurrence of each of the eight tenses studied here. For instance, 19% of Editorial finite clauses use simple-past tense, as opposed to 56% of the FPN clauses.

Tense	Editorial	FPN	
N	312	246	
simple-past	19%	56%	+++
simple-present	39%	19%	+++
simple-future	8%	2%	+++
simple-modal	22%	13%	+++
past-perfect	2%	2%	
present-perfect	8%	6%	
future-perfect	-	-	
modal-perfect	1%	1%	

Table 1: Distribution of Tense Selections over Article Types

Throughout this paper, statistical significance of differences is shown using one or more '+' sign.

- '+' indicates significance at the 90% level (10% chance of error);
- '++' indicates significance at the 95% level (5% chance of error);
- '+++' indicates significance at the 98% level (2% chance of error).

The table thus shows four significant difference between Editorials and Front Page News as regards tense selection. Editorials use far less *simple-past* than FPN (19%+++ vs. 56%), made up for by a higher degree of *simple-present* (39%+++ vs. 19%), *simple-future* (8%+++ vs. 2%), and *present-perfect* (8%++ vs. 6%).

These results do not, however, explain much by themselves. It is up to the analyst to posit some explanation of these results, such that FPN tends to express events which have happened already, so *simple-past* is common, while editorials tend to express the consequences and background of these events, e.g., relational processes such as *The United States has friends and interests in the Gulf; The correct policy today is to shift more of the costs of collective security onto our more prosperous allies.*

However, this is the analyst intruding on the data -- the data does not tell us this -- all we know directly is the probabilities of particular grammatical choices. The analysis itself has not explained the data, just given the analyst a clearer idea of the patterns which need to be explained.

3. A Three-Layered Approach

Register studies which try to explain grammatical patterns directly in terms of contextual differences often show some result, but a result obscured. Writers do not make grammatical choices without taking semantics into account. A newspaper writer, for instance, does not say, this is an editorial, so I will use simple present tense for this sentence. Rather, they needs to take into account what kind of event or state they are writing about, whether that event/state happened in the past, is ongoing now, or is one predicted for the future.

A register study of grammatical phenomena thus needs to take into account the semantics of the text, and how different semantic patterns favour particular formal expressions. We can look at the problem in terms of two parts of the writing task[2]:

1. *Content selection*: choosing what event/state to write about;
2. *Content expression*: choosing how to express a given event/state.

---

[2] In the normal case, the writer chooses what to write first, and then how to express it. However, there is a level at which the expression controls what is said. For instance, in poetry, the rhythm and rhyme can be preset, and content selected to match this constraint. In normal writing also, often formal constraints take part in the decision as to whether or not to include content, for instance, a short paragraph may call for additional information as padding.

Some of the explanations of grammatical patterns lie in the first area -- what the writer chooses to talk about may favour particular grammatical expressions. For instance, writing about past events will favour the use of past tenses, while writing about the future will favour future tenses and modals.

However, grammatical patterns are not entirely explained by the choice of content -- the language system leaves possibilities for expressing any given event or state. For instance, an event can be expressed as a clause (e.g., *the US bombed Baghdad*) or as a noun phrase (e.g., *the bombing of Baghdad*). Similarly for tense, a past event can be expressed using present perfect ('*Baghdad has been bombed*') as well as one of the past tenses.

So, a more complete explanation of the effect of register on grammatical patterns needs to take into account both of these factors:

1. Does register influence what events/states are expressed?
2. Does register influence how particular types of events or states are expressed?

We are thus proposing to replace the typical two level register study (figure 1) with a three level study (figure 2) [3].

Figure 1: Traditional Register Studies | Figure 2: Three-Layered Register Study
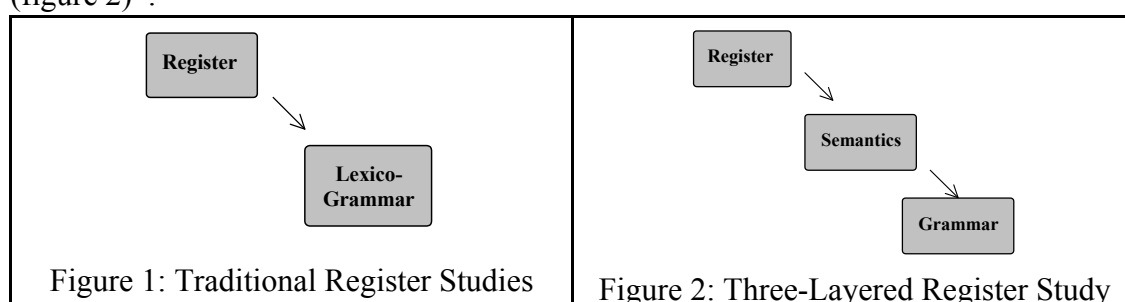
4. A Three-Layered Study of Tense

For this study, a corpus of newspaper articles was collected from 7 newspapers (6 US, 1 British), and segmented into finite clauses, resulting in a total of 558 tokens. Of these 312 were from editorials, and 246 from FPN. The corpus was drawn from financial and military reporting.

The data was coded using the *Systemic Coder*, software which facilitates the coding of text (O'Donnell 1995)[4]. All statistics in this paper were derived using the statistical functions of this program.

4.1. The Model

Each clause was coded on three levels:

1. *Text-Type*: editorial vs. FPN;
2. *Realis*: a representation of the temporal semantics of each clause;
3. *Grammatical Tense*: the grammatical tense of each clause.

These categories, apart from text-type, will be explored below.

4.1.1 Semantics: Realis

Realis concerns whether or not an event/state has actually taken place, or is still occurring, or has not in fact occurred. A *realis* process is one which has either actually occurred, or is still in the process of occurring. The two subtypes of realis are:

- *completed-process*: the process completed before the point of speaking;
- *ongoing-process*: the process is ongoing at the point of speaking.

---

[3] Caffarel (1993) also argues that tense needs to be explained by reference to an intermediating semantics.

[4] Available free from: http://www.wagsoft.com/Coder/index.html.

All other processes are labelled *irrealis* - processes which have not happened, and are not happening now. This includes expectations of the future (e.g., *he will run*), statements of obligation (*he should run*), denials (e.g., *he did not run*), etc.

Ongoing-process should not be confused with the grammatical notion of *progressive* aspect -- an *ongoing-process* is ongoing *at the time of speaking*, while an event reported in progressive aspect is ongoing at some reference point, which may be the point of speaking, but is often not. For instance, *He was running yesterday* reports a completed-process (assuming the running finished), but it is reported in *progressive* aspect.

4.1.2 Grammar: Tense Selection

English provides six basic tenses, *simple-past*, *simple-present*, *simple-future*, *past-perfect*, *present-perfect* and *future-perfect*. However, we also need to take into account the possibility of modality -- *The US can bomb Baghdad.* Figure 3.1 shows the basic tense choices for English. The percentage occurrence of each tense in our corpus is also shown.

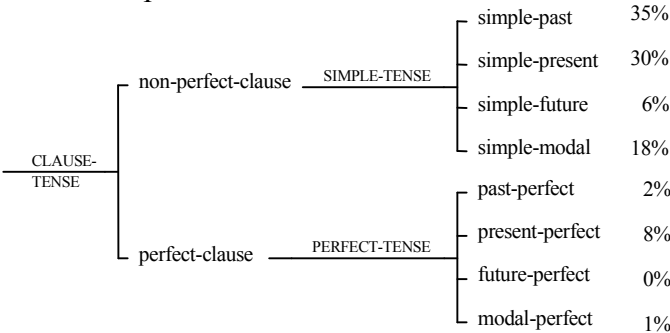|			
		simple-past	35%
	SIMPLE-TENSE	simple-present	30%
non-perfect-clause		simple-future	6%
		simple-modal	18%
CLAUSE-TENSE		past-perfect	2%
	PERFECT-TENSE	present-perfect	8%
perfect-clause		future-perfect	0%
		modal-perfect	1%

Figure 3: English Tense Systems

Each of these cases could also be expressed in the *progressive* aspect (e.g., *USA was bombing Baghdad*), although this somewhat changes the meaning. However, due to space limitations, I will not explore progressive aspect in this paper.

4.2. Results

We studied the relations between text-type, realis and tense in terms of three studies:

1. How do text-types vary in the realis they use?
2. How are realis types realised in terms of tense?
3. How do text-types vary in the way realis-types are typically realised?

By separating out these issues the patterns of tense selection are shown more clearly.
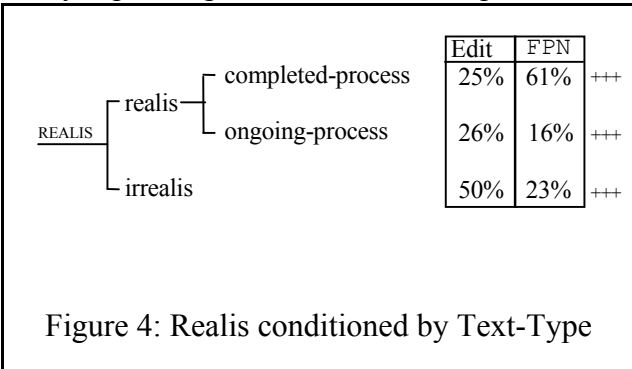
	Edit	FPN	
completed-process	25%	61%	+++
ongoing-process	26%	16%	+++
irrealis	50%	23%	+++

REALIS — realis — completed-process / ongoing-process; irrealis

Figure 4: Realis conditioned by Text-Type

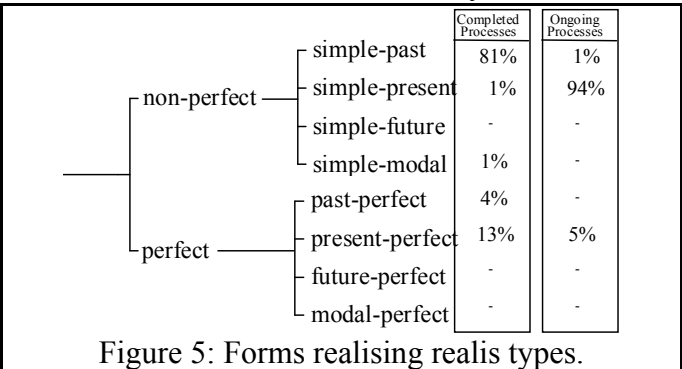|	Completed Processes	Ongoing Processes
simple-past	81%	1%
simple-present	1%	94%
simple-future	-	-
simple-modal	1%	-
past-perfect	4%	-
present-perfect	13%	5%
future-perfect	-	-
modal-perfect	-	-

Figure 5: Forms realising realis types.

4.2.1 Register & Realis

The distribution of realis types in each of the text-types is shown in figure 4. As can be expected, FPN reports more completed-process (the events which are being reported). Editorials on the other hand provide more of:

- *ongoing-processes* (26% vs. 16%): these are for the most part states rather than events, e.g., *We are ready for war,* or *The public finances are healthy*. Very few are ongoing actions, as in *He is counseling the government;*

- *irrealis* (50% vs. 23%): Editorials include more than double the number of irrealis clauses, which is also in keeping with the nature of the text-type - they need to talk about what may happen, what could have happened and did not, what people say could happen, etc.

4.2.2 Realis & Tense

The next study explores the ranges of tenses which can realise each realis type. We have ignored irrealis, because this category is very complex, taking in lots of different forms (negations, conditionals, projections, modals, future, etc.), all of which behave differently. We thus look at how both completed-processes and ongoing processes are realised. Figure 5 shows the results.

- *Completed processes* can be realised by 5 tense combinations. The most common is, as expected, simple past (81%). The present-perfect is used in 13% of the cases. This tense is used when the writer wishes to indicate that a past event still has consequences for the present (Moens & Steedman 1977). For instance, someone usually says *I have spilt my coffee* while the consequence of the action (the spill) is still to be cleaned up. The past-perfect is used 4% of the time (indicating one past event had consequences for another past event).

  Two instances used simple-present to report on a past event: i) *Chirac intervenes in Merger row*: this is a headline, which follow different rules to running text. ii) *The decision follows Swissair's announcement Tuesday:* here two separate events, both completed, are related by an abstract process of sequence. This state is then reported in the present tense, as if it is a timeless truth. Note the alternative 'followed' would work equally as well.

  Two instances of a modal form were found expressing past events. These will be discussed below.

  In summary, the realis of an event partially constrains the tense options available. However, it still leaves a choice to be made by the writer.

- *Ongoing processes* can be realised by 3 tense choices. Simple-present is used in roughly 19 out of 20 cases. Present perfect was used in 1 case in 20, for instance: *[Sabena] has been refusing to honor Swissair tickets*. As this policy was continuing at the time of writing, it is coded as an ongoing-process. This tense is also used to express completed-processes: basically, the tense implies that some event/state which at least started in the past still has consequences now, whether or not it has completed.

  The third tense seems rather an unusual option for realising an ongoing processes – simple past. The one instance was: *[Sabena] was seeking protection from its creditors*. The writer expressed this as part of the context for an earlier event. However, Sabena was still seeking protection at the time of writing, which makes it an ongoing-process. The writer was chosen to focus only on the fact that the state was ongoing at the earlier point of time.

4.2.3 Register and Realis-Tense Mapping

The 3-layered approach makes possible a third study: Do different registers differ in how realis maps onto tense? Each realis option has a number of realisation possibilities, and we are interested to see if the way realis is expressed changes across registers. In other words: how does the expression of a given meaning type differ between registers? This relationship between one linguistic level being realised in terms of the relation between two others has been termed *metaredundancy* by Jay Lemke (as early as 1983, but see Lemke 1995).

Figure 6 presents the results of this study. In relation to the expression options of *completed-processes*, we can see that FPN more strongly prefers simple-past to express completed-processes (87% vs. 69%), while editorials use more of the *present-perfect* option (23% vs. 7%). Part of the job of an editorial is to show present significance of past events, and that is exactly what this tense
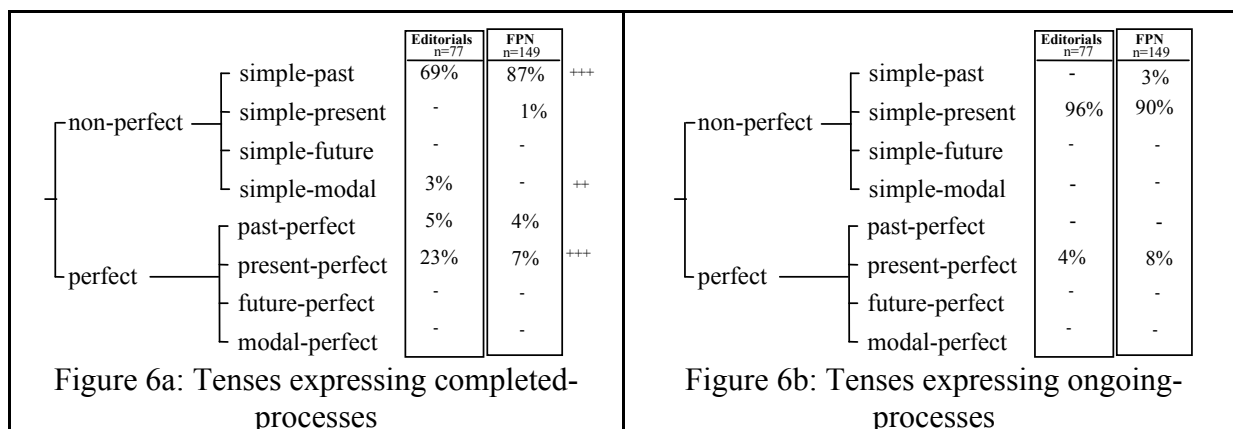
does. The purpose of FPN is more just to present events, without too much interpretation, so simple past is preferred.

Another option which exists for editorials, but not for FPN, is the use of a *simple-modal* to express a past event. There were 2 instances:

> *[Saddam did not care] that an alliance <u>would</u> come together.*

> *[Saddam has devoted] what he <u>could</u> borrow [… into constructing a military machine].*

The first expresses a past expectation about future events (which eventually took place, and are thus realis). The second expresses a past ability, which also was realised, and is thus realis. This kind of sentences occur in editorials because it is the function of editorials to draw messages from the past, to apply to the present, to influence the future.

	Editorials n=77	FPN n=149	
non-perfect — simple-past	69%	87%	+++
simple-present	-	1%	
simple-future	-	-	
simple-modal	3%	-	++
perfect — past-perfect	5%	4%	
present-perfect	23%	7%	+++
future-perfect	-	-	
modal-perfect	-	-	

Figure 6a: Tenses expressing completed-processes

	Editorials n=77	FPN n=149
non-perfect — simple-past	-	3%
simple-present	96%	90%
simple-future	-	-
simple-modal	-	-
perfect — past-perfect	-	-
present-perfect	4%	8%
future-perfect	-	-
modal-perfect	-	-

Figure 6b: Tenses expressing ongoing-processes

In regards to ongoing-processes, it seems that there are no significant differences in the way these resources are used in FPN and editorials. Possibly a larger data set would reveal differences.

5. Summary

We have argued against analysing registerial distribution of grammatical choices in isolation of the semantics they realise. Without taking the semantic context into account, we may reach conclusions about grammatical preferences which are not true grammatical preferences, but rather a reflection of skewed content selection.

This approach allows three sub-studies:

a)   What semantic patterns are included in each text-type?

b)   How are these patterns expressed lexico-grammatically?

c)   How does the formal expression of meanings differ between text-types?

Our studies have shown that the major difference between the text-types is due to differences in realis. Subsidiary differences in expression of these realis cases also exist, and only become clear when separated from issues of content-selection.

This methodology can be extended to explore other areas of grammatical selection, for instance, separating out issues of speech function and grammatical Mood (interrogative/declarative/imperative).

In conclusion, it needs to be stated that text-type only partially explains tense choice. A full explanation needs to take into account the textual factors, such as the need for relating the time of an event to some previously mentioned event. For more details on this aspect of tense selection, see, for example, Schiffrin (1981) or Moens & Steedman (1988).

BIBLIOGRAPY/REFERENCES

Bernstein, B.: *Class, codes, and control. Vol. 1: theoretical studies towards a sociology of language*, Routledge & Kegan Paul: London, 1971.

Caffarel, Alice: "Interacting between a generalized tense semantics and register-specific semantic tense systems: a bi-stratal exploration of the semantics of French tense". *Language Sciences* 14.4: 385-418, 1992.

Halliday, M.A.K. & Z. L. James: "A quantitative study of polarity and primary tense in the English finite clause". In: John M. Sinclair, Michael Hoey and Gwyneth Fox (Eds.), *Techniques of description: spoken and written discourse*, Routledge, 1993.

Lemke, J: *Textual Politics: discourse and social dynamics*, Taylor & Francis: London, 1995.

Moens, M. & M. Steedman:. "Temporal ontology and temporal reference". *Computational Linguistics*, 14(2):15 - 28, 1988.

Nesbitt, Christopher & G.A. Plum: "Probabilities in a Systemic Grammar: The clause complex in English". In: R.P. Fawcett & D. Young (Eds.) *New Developments in Systemic Linguistics* (Vol. 2), Frances Pinter: London, 1988.

O'Donnell, Michael: "From Corpus to Codings: Semi-Automating the Acquisition of Linguistic Features". *Proceedings of the AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation*, Stanford University, California, March 27 – 29 1995.

Plum, G.A. & Ann Cowling: "Social constraints on grammatical variables: tense choice in English". In: R. Steele & T. Threadgold (Eds.) *Language Topics* (Vol. 2), John Benjamins: Amsterdam, 1987.

Schiffrin, D.: "Tense variation in narrative". *Language*, 57, 45-62, 1981.