

# Learning and teaching the subjunctive in French: the contribution of corpus data

**Angela CHAMBERS**

University of Limerick  
Plassey Technological Park, Limerick, Ireland  
angela.chambers@ul.ie

Depuis le milieu des années 1980, l'exploitation de corpus linguistiques dans le cadre de l'apprentissage des langues est en croissance constante. Cette étude contribue à la recherche consacrée à cette évolution en examinant l'apport potentiel des corpus de langue française accessibles sur Internet pour l'apprentissage et l'enseignement du subjonctif. Nous examinerons dans une première partie quelques publications scientifiques sur le rôle des corpus dans l'apprentissage et l'enseignement des langues, et nous nous pencherons ensuite sur trois exemples: l'emploi du subjonctif avec  *falloir*, l'emploi moins systématique du subjonctif pour exprimer une dimension subjective, et les temps plus rarement employés du subjonctif, l'imparfait et le plus-que-parfait. Le but de l'analyse est de savoir si les données consultées permettent d'enrichir le nombre limité d'exemples présentés dans les manuels et les livres de grammaire en fournissant un accès à des emplois concrets récurrents. L'étude soulève également des questions, telles que le rôle de la répétition dans l'apprentissage des langues, la présence ou l'absence de contexte, la façon dont les données peuvent être exploitées par les enseignants, et les défis posés par la variation et les erreurs.

**Mots-clés:** Corpus, enseignement des langues, apprentissage des langues, variation, le subjonctif en français

## 1. Introduction

Since the publication in the 1970s of one of the earliest corpus-based course books, *Les Orléanais ont la parole* (Biggs & Dalwood 1976), corpora have contributed to language learning and teaching in a number of ways. Dictionaries and grammars (in particular for English) are increasingly based on corpus data instead of invented examples. In addition, a substantial body of research publications are devoted to the use of corpora with learners, either by using concordances prepared by the teacher as language learning materials (see Kerr this volume), or by giving the learners direct access to corpora and providing training in the use of concordancing software. Boulton (2010) identified more than 70 studies in the context of learning outcomes, and in subsequent publications has studied other aspects of teachers' use of corpus data with learners (see, for example, Boulton 2012<sup>1</sup>). Despite developments such as these, researchers in the area of corpus consultation by learners are well aware that the potential of corpus data to enhance the language learning and teaching environment has yet to be realised.

---

<sup>1</sup> Alex Boulton has identified more than 116 relevant publications in all (personal communication, 10 October 2012). See <http://bit.ly/STZegS> ('autres utilisateurs' > 'connexion anonyme', 16.05.2013).

This article aims to contribute to this development by investigating the potential of a number of freely and easily available small corpora of French for learning and teaching the subjunctive. The hypothesis is that the data will provide a rich source of occurrences of naturally occurring discourse to complement the necessarily limited number of examples provided in course books and grammars. After a brief account of the research context in relation to theoretical and pedagogical investigations of the role of corpora in language learning and teaching, the paper focuses on three areas: the common and uncontroversial use of the subjunctive with *falloir*<sup>2</sup>; the less clear and less systematic use of the subjunctive to express a subjective dimension, in this case with *sembler*; and the more rarely used tenses of the subjunctive, namely the imperfect and pluperfect. Analysis of these data enables us to investigate the hypothesis above, and also raises a number of theoretical and pedagogical issues in areas such as the role of repetition in language learning, the presence or absence of context in these data, their appropriateness for different levels of learner, and the challenges which the presence of variation and errors poses for the teacher.

## 2. Corpora and language learning: the research and teaching context

### 2.1 *The corpus as a source of multiple genuine examples*

Given that several renowned researchers in applied linguistics have criticised the reliance on invented examples in language learning materials (Sweet [1899] 1964: 73; Firth 1957: 24; Sinclair 1988: 6; all cited in Cook 2001: 366-367. See also McCarthy & Carter 1995), it is not surprising that one of the attractions of corpora for language teachers and researchers lies in the fact that they are a source of multiple genuine examples. The term 'genuine' is used here in preference to 'authentic', in keeping with Widdowson's (1983: 30) use of the term, reserving 'authentic' to refer to the use to which texts are put in a classroom rather than as an inherent characteristic which they possess. Indeed, in several studies of corpus consultation by learners, they appreciate this access to examples which are not invented (Bernardini 2002: 179; Chambers 2005: 120; Cheng *et al.* 2003: 181; Yoon & Hirvela 2004: 275; Wen-Ming & Hsien-Chin 2008: 42). Cobb (1997: 303) stresses the importance of multiple examples, referring to this aspect of corpus consultation as "multicontextual learning". Ellis (2002: 177) describes the list of occurrences in a concordance as "mindful repetition in an engaging communicative context", seeing this new use of repetition as a modern reinvention of behaviourist repetition as a way of mastering aspects of language use. While the underlying reasons for the use of repetition are not at all the same, this

---

<sup>2</sup> *Vouloir* could also have been used here.

reappearance of repetition in the context of language learning and teaching in the late twentieth and in the twenty-first century is nonetheless of considerable interest and deserves more attention in research in corpora and language learning and teaching. In relation to Cobb's use of the term "multicontextual learning", it is important to note that context is not necessarily absent in this form of learning, as the learner can be presented with examples to complement a single occurrence of the aspect of the language in a text which is being studied in class. Furthermore, a concordancer such as Wordsmith Tools (Scott 2004) gives access to the whole text by a simple click of the mouse.

This is not to suggest that the inclusion of multiple genuine examples presents no problems for the learner and teacher. Learners who participated in empirical studies where they had the opportunity to search a corpus and analyse the results, while giving generally positive reactions, nevertheless described the process of sorting and analysing large numbers of concordance lines as time-consuming, laborious and tedious (Yoon & Hirvela 2004: 274; Chambers 2005: 120). In addition, teachers using concordance data can also encounter difficulties. As we shall see, the rarer uses of the subjunctive in this study provide examples of performance errors in an area where they are well known in French, the imperfect and pluperfect subjunctive, which presents a challenge to teachers of learners at an advanced level. Furthermore, an example of taboo language, which can be a problem of varying levels of difficulty for teachers in corpus data, also occurs in these data.

Finally, while genuine examples are now commonly used, and promoted in documents such as the *Common European Framework of Reference for Languages* (Council of Europe 2001), it is important to note that this is not universally accepted in language learning research. Cook (2001) makes the case for invented examples, claiming that they have four advantages: as a means of making the lesson more personal and spontaneous; illustration; promoting noticing; and as a mnemonic device (2001: 379-380). Without disagreeing with Cook, it can be argued that corpora also have the potential to fulfil some of these functions. For example, several researchers have pointed out that the multiple examples provided in a concordance promote noticing, in the sense in which Schmidt (1993) uses the term. As Schmidt (1995: 20) observes: "The 'noticing hypothesis', states that what learners notice in input is what becomes intake for learning". For Gabrielatos (2005), the sight of a concordance, with repeated occurrences of the search word surrounded by co-text, even encourages a new way of reading texts for learners, condensed reading. In conclusion, while corpora clearly make it possible to give learners access to multiple genuine examples to support their learning, there is still a need for specific investigations of what can be learned from consulting corpus data, of the extent to which learning actually occurs, and of the reactions of

learners and teachers to their use. This article deals with the first of these issues.

## 2.2 Finding corpora

Teachers wishing to integrate corpora into their teaching must first identify easily and, ideally, freely available resources. They may also wish to create a corpus, although this clearly requires much more time and effort (Timmis 2010). Cappeau and Seijido (2005) provide an online directory of corpora in French, and the French governmental body, the *Délégation générale de la langue française et des langues de France* (DGLFLF) hosts a site giving details of corpora which are available online<sup>3</sup>. In the teaching context, however, a large reference corpus is unlikely to be of immediate use to a teacher interested in producing teaching materials based on corpus data but lacking training in corpus consultation and analysis. Chambers *et al.* (2011) suggest a graded approach to this problem, beginning with the use of corpora which are available online and have built-in concordancers. A resource such as Lextutor<sup>4</sup>, for example, provides access to substantial corpora in French, as well as a concordancer which can be used with little or no prior training (cf. Kerr this volume). At the time of writing the French corpora, 2.8 million words in all, include a spoken corpus, as well as written texts in literary, journalistic and other genres. In addition, spoken language resources suitable for learning and teaching have been developed in several languages, including French, in two projects funded by the European Commission, SACODEYL and Backbone<sup>5</sup>. These corpora include interviews with native speakers of French, adolescents in the case of SACODEYL and adults in the case of Backbone, annotated in ways intended to be of practical help to language teachers. These, and other resources, are examples of pedagogic corpora, which can be defined as:

A large enough and representative example of the language, spoken and written, a learner has been or is likely to be exposed to via teaching material, either in the classroom or during self-study activities. (Meunier & Gouverneur 2009: 186)

While such resources make it easy for teachers to integrate corpus data in their teaching, it is important to note that easily and freely available resources are still few in number, particularly for languages other than English. Corpora such as SACODEYL and Backbone can be classified as small corpora. While Aston (1997: 54) defined small corpora as typically consisting of approximately 20,000-200,000 words, it has since become common to find corpora of one

<sup>3</sup> [www.dglflf.culture.gouv.fr/recherche/corpus\\_parole/corpus\\_en\\_ligne.htm](http://www.dglflf.culture.gouv.fr/recherche/corpus_parole/corpus_en_ligne.htm) (14.05.2013). At the time of writing this database appears to have been last updated in 2006.

<sup>4</sup> [www.lexutor.ca](http://www.lexutor.ca) (07.05.2012).

<sup>5</sup> [www.um.es/sacodeyl/](http://www.um.es/sacodeyl/) and [www.um.es/backbone/](http://www.um.es/backbone/) (14.05.2013).

million words created for language learning purposes (see, for example, Chambers & Rostand 2005; Chambers & Le Baron 2007).

A number of corpora are used for this study. The *Corpus Chambers-Rostand de français journalistique* (Chambers & Rostand 2005, henceforth referred to as CRFJ) consists of approximately one million words and includes 1,723 articles from the French newspapers *Le Monde*, *L'Humanité*, *La Dépêche du Midi*, in 2002 and 2003. The rubrics represented are editorials, culture, sport, national news, international news, and finance. As we shall see, this produces a substantial number of occurrences of the subjunctive with  *falloir*. As it includes fewer examples of *sembler*, Lextutor is also used. Finally, for the rarer uses of the imperfect and pluperfect tenses of the subjunctive, where the number of examples in CRFJ is very small, a one-million word corpus of research articles in French is also used (Chambers & Le Baron 2007, henceforth referred to as CLARF), containing 159 research articles in 20 peer-reviewed, freely available online journals. The ten disciplines featured in the corpus are: media / culture, literature, linguistics and language learning, social anthropology, law, economics, sociology and social sciences, philosophy, history, and communication studies. Small corpora such as these provide fewer examples than large reference corpora and two of the three are also limited in being genre-specific. They do, however, present a number of advantages for the teacher with little or no expertise in corpus linguistics, or for the learner with direct access to the corpora. As Aston (1997: 55) notes, they are easier to use and to become familiar with than their larger counterparts, and can serve as a useful starting point for students who may later wish to move on to the analysis of larger corpora in a research context. Aston (2001: 75-76) also points out that, while such corpora are useful for finding examples in the language-learning context, they cannot be used as the basis for generalisations about language use in the context of linguistic research.

### 3. Investigating the subjunctive in small corpora

#### 3.1 *Falloir*

Even a corpus of one million words will provide significant numbers of occurrences of common uses of the subjunctive. In a search for forms of the lemma *falloir*<sup>6</sup> in CRFJ, for example, the 1,054 occurrences are distributed as follows: *faut* (728), *faudra* (99), *va falloir* (12), *fallait* (92), *faudrait* (66), *fallu* (51), *fallut* (2), *faillie* (4). While the majority of these involve the infinitive or a noun rather than the subjunctive (which is in itself already quite an interesting finding), there are nonetheless enough examples to provide useful teaching and learning materials. A simple search for *faut* followed by *que* or *qu'*, for example, produces 35 occurrences of the subjunctive, while additional uses of

<sup>6</sup> Wordsmith Tools (Scott 2004) is the concordancer which has been used for this study.

*faut* with the subjunctive can be found including the negative or an inverted form (*encore faut-il que*), or where *faut* is followed by an adverbial or adverbial phrase such as *bien*, *bien entendu*, *d'abord*, or *néanmoins*. In the case of the other forms of the lemma *falloir* the occurrences of the subjunctive are as follows: *faudra* (10), *va falloir* (2), *fallait* (11), *faudrait* (15), *fallu* (6), *fallut* (0), *faillie* (0). The next step would involve finding expressions suitable for learners at the appropriate level. Examples<sup>7</sup> are provided below, which could be used with learners at intermediate or advanced level:

- Un contrôleur lui dit alors qu'avec une poussette ce n'est pas possible, qu'il **faut** qu'elle attende le suivant.
- Il **faut** que cette équipe aille aux Jeux, insiste Alain Weisz.
- Demain, il **faut** qu'on revienne dans cette putain de rue.
- j'ai fini par me dire qu'il y avait un problème. Il **fallait** que je prenne du recul.
- pour la saison prochaine tout est possible. Il **faudra** pour cela que je sois à 100 %, ce qui n'est pas encore le cas
- s'il veut être champion du monde, il **faudra** qu'il se batte, et je suis sûr qu'il le fera
- Il ne suffira donc pas à Sharon de gagner. Encore **faudra-t-il** qu'il puisse gouverner.
- On craint tout, reconnaît Eric Mariottat. Mais il **va falloir** que l'Etat craigne aussi, car bientôt, il n'y aura plus de restaurateurs
- le couple franco-allemand peut lui apporter cette colonne vertébrale. Encore **faudrait-il** qu'il fasse preuve de dynamisme.

As the above occurrences illustrate, the teacher can easily find straightforward examples at different levels of difficulty which, while not fully situated in context, nonetheless make it easy in some cases for the learners to relate them to the world in which they live. For teachers and learners with direct access to the corpus, Wordsmith Tools (Scott 2004) and other concordancers, such as the freely available AntConc<sup>8</sup> make it possible to access the full text with a simple click of the mouse, so that, in the last example below, for example, additional text could make it clear that *le couple franco-allemand* refers to the two governments. The unpredictability of concordance lines also presents the teacher with challenges which a course book will not normally include, such as the use of taboo language in the quotation in the third example above. Thus, rather than unquestioningly using the examples provided by the authors of course books, teachers are encouraged to reflect on the type of language use which they present to their learners and the advice which they give.

<sup>7</sup> Bold print is used in the examples to draw attention to items under discussion.

<sup>8</sup> <http://www.antlab.sci.waseda.ac.jp/software.html> (14.05.2013).

Paper or electronic resources created by a teacher could clearly include many more examples of the subjunctive with *falloir* than the small number listed here, thus encouraging the "noticing" (Schmidt 1993) and "condensed reading" (Gabrielatos 2005) referred to above. In addition, for less advanced learners, simple gap fill exercises could easily be created by providing the infinitive and asking the learners to include the verb in the subjunctive. At a more advanced level, learners could carry out the searches themselves, and study occurrences of more advanced uses, such as the uses of *falloir* where the subject and verb are inverted. It is also worth noting here that studying concordances can often result in what Johns (1988) terms serendipitous discoveries, i.e. uses which one notices but which were not the object of the search. It is clear from even a superficial examination of occurrences of *falloir* that there is also scope for producing examples of uses with numbers and temporal expressions, which could be of use to the teacher or learner in another context, or which could provide the learners with examples of incidental learning.

### 3.2 *The subjunctive with sembler*

A useful function of corpus data in the context of the subjunctive is the easy availability of examples of uses where native speakers will use the subjunctive with a particular verb, *sembler* for example, in certain situations but not in others. As we shall see, however, this can also present the teacher with a challenge. Grammars can provide a clear (though potentially questionable) explanation of such phenomena, but, understandably, they can only include a few examples. In the case of *sembler*, L'Huillier (1999: 165) provides the following explanation and examples:

*Il me semble que* means 'I think that', hence the indicative, while *il semble que* means 'it looks as if', i.e. there is a high degree of subjectivity, hence the subjunctive.

*Ex: Il me semble que vous avez changé d'avis*

It seems to me that you have changed your mind

But:

*Il semble que vous ayez changé d'avis*

It looks as if you have changed your mind.

The very small number of examples given in the corpus-based Oxford-Hachette French Dictionary follow this rule. A teacher searching for expressions with *il semble que* and *il me semble que*, will easily find a significant number of occurrences in CRFJ and in the French corpora in Lextutor, 59 in all for *il semble que* and 10 for *il me semble que* and *il nous semble que*. In expressions with *il me semble que* and *il nous semble que*, only the indicative is used. In the case of *il semble que* it quickly becomes evident, however, that the subjunctive is used in the majority of cases but not in all. The distribution in the two corpora is given in Table 1. "Same form"

indicates that it is not possible to decide if the indicative or subjunctive mood is used, as they have the same form.

Expression	CRFJ	Lextutor corpora
<i>Il semble</i> + subjunctive	9	21
<i>Il semble</i> + indicative	6	10
<i>Il semble</i> + same form	5	8
<i>Il me semble</i> + indicative	4	4
<i>Il me semble</i> + subjunctive	0	0
<i>Il me semble</i> + same form	1	0
<i>Il nous semble</i> + same form	1	0
Total	<b>26</b>	<b>43</b>

Table 1: The indicative and subjunctive moods with *sembler*

As Table 1 shows, carrying out simple corpus searches has the advantage of providing the teacher and learner with many genuine occurrences of *il semble que*, but the variation from the generally accepted rule presents a challenge.

As the number of uses of *il semble que* with the indicative (16), while fewer than the subjunctive (30), is not insignificant, it is tempting to examine the context to see if a pattern can be observed. Given the small size of the corpora (3.8 million words in all) and the limited number of occurrences, this could not produce a reliable finding, but perhaps an indication that a hypothesis could be tested in a much larger sample. At a first glance a teacher or learner might get the impression that the indicative is used when the degree of subjectivity is relatively low, as in the examples below (the source is CRFJ unless otherwise stated):

- Certains bureaux sont même restés fermés. "**Il semble qu'**il y a beaucoup de gens en Géorgie qui ne pourront pas voter [...]"
- le président uruguayen implore l'aide extérieure. **Il semble que**, pas plus que les grands organismes prêteurs, FMI, Banque Mondiale, BIB, les gouvernements argentins et uruguayens ne peuvent admettre l'échec de l'ultralibéralisme
- c'est la difficulté qu'il y a en Afrique à voir nos films. De plus en plus, **il semble que** nous les faisons pour un public européen.
- Comment les Agenais sont-ils parvenus à vaincre Toulouse? En étant tout simplement plus stupides que nous sur tous ces points de fixation pourris. **Il semble que** c'était une des solutions cette année pour nous battre.
- "Affaires de police! s'écria le roi, relevant les paroles de M. de Tréville affaires de police! et qu'en savez-vous, Monsieur? Mêlez-vous de vos mousquetaires, et ne me rompez pas la tête. **Il semble**, à vous entendre, **que**, si par malheur on arrête un mousquetaire, la France est en danger. Eh! que de bruit pour un mousquetaire! (Lextutor: Dumas, les Trois Mousquetaires)

In the last example the number of parenthetical comments which separate *il semble que* from *en danger* might derail all but advanced learners, but it is

nevertheless interesting to show that this use of *il semble que* with the indicative is not a new phenomenon. Other examples of Dumas's use of it are also present in searches of *Les trois Mousquetaires*. Consulting the occurrences of *sembler* with the subjunctive, however, one also finds cases where the level of subjectivity is even lower, and where the writer adds an adverb after the verb to strengthen the point being made:

- Mais **il semble** bien **que** Jacques Chirac soit enclin à laisser moins d'autonomie à Jean-Pierre Raffarin qu'il n'en avait accordée à Alain Juppé
- Non seulement l'administration américaine n'a pas l'intention de renoncer devant les oppositions que sa stratégie de toute puissance provoque, mais **il semble** même **qu'**elle les ait intégrées.
- **Il semble** en effet **qu'**Henri Darré, directeur de l'usine depuis septembre 2001, ait mis un point d'honneur à se signaler par une attitude le plus souvent qualifiée de "provocatrice": fouilles, interdiction de se parler, de s'asseoir
- Mondial 2002 Pour les retransmissions des matchs de la Coupe du monde, " **il semble**, a priori, **que** nous ayons trouvé un accord avec toutes les radios nationales
- Des discussions ont en effet eu lieu et **il semble que** nous soyons suivi sur tous les points, détaille Thierry Jarlet,

Are the uses of the indicative then a case of performance errors, or an illustration of variation where usage is changing? A much larger corpus would be required to provide data for a serious study of this phenomenon. The question which arises in the context of this study, however, is how the teacher can deal with such occurrences. For less advanced learners it may be tempting to omit the examples which do not conform to the rules in order to avoid confusion. For more advanced learners, however, they provide a richer and more genuine learning experience than the sanitised presentation of rules which course books and grammars understandably present as unproblematic. Conrad (2004: 80) makes the case for including variation in language teaching, pointing out that "teaching practices that are informed by variation are the exception rather than the rule". As Sinclair (2004: 274) so aptly sums up the situation, "Variation is a nuisance".

### *3.3 The imperfect and pluperfect subjunctive*

Teaching the imperfect and the pluperfect subjunctive to advanced learners presents an interesting challenge for the teacher of French. As these forms are rarely used in French, it is not surprising that they do not feature prominently in course books, grammars and language learning web sites. The web site french.about.com gives the following advice:

The imperfect subjunctive is a literary verb form used in formal writing such as literature, journalism, and history. Like all literary verb forms, you really only need to be able to recognize it, not use it. ([http:// french.about.com](http://french.about.com), 14.05.2013)

Similarly L'Huillier (1999: 159) does not include these forms in her grammar, giving the following advice: "The imperfect is rarely used and conjugations should be checked in a dictionary or book of conjugations". She also notes that the imperfect and pluperfect subjunctive "are almost never used in conversation or everyday writing, particularly as their forms tend to sound very strange and even sometimes ridiculous, particularly in the 1<sup>st</sup> and 2<sup>nd</sup> persons" (p.160). She cites two examples of the use of the tenses for comic effect (pp.147 and 161), a literary text and a text from France-Inter. As we shall see, a third example of this is found in the corpus data. Grevisse, however, challenges the accepted ideas of the comic potential of the tenses in that they sound peculiar, and of the fact that they are difficult to learn:

Sauf cas particuliers, [he has just quoted a text where the author explicitly refuses to use "embrassasse"], l'idée selon laquelle le subjonctif imparfait où plus-que-parfait serait contraire à l'euphonie est subjective: c'est seulement la rareté de ces formes qui les rend surprenantes, et *fascinassions* n'est pas moins harmonieux que *fascination*. - C'est aussi leur rareté qui les rend difficiles, et non leur difficulté qui les rend rares. (1993: 1271)

The imperfect and pluperfect subjunctive are thus neglected by grammarians, considered a subject of ridicule by literary authors and the media, and rarely used. The implication is that they therefore present no challenge to the teacher and the advanced learner of French; they can simply be ignored.

However, reading the comments of grammarians and researchers about how rare these tenses are, one finds comments to the effect that, when they are used in contemporary French, certain patterns can be observed. L'Huillier (1999: 160-161), for example, notes that the third person is "slightly more 'acceptable'" than the first and second person. Girodet (1986: 865) points out that they mainly occur with the verbs *avoir* and *être*. Hanse (1991) provides a broader range of verbs with which they are used, namely *être*, *avoir*, *pouvoir*, *devoir* and *vouloir*. It is thus tempting to revise the advice of french.about.com as follows for advanced learners: you need to be able to recognise it and to be aware of the limited number of contexts where it is still currently used. It is here that corpus data can be particularly useful, providing the teacher and learner with recent examples of usage. Given the rarity of the uses of the tense, two one-million word corpora have been consulted here, CRFJ and CLARF. These corpora are consulted as a source of examples, and not as a means of comparing language use in the two genres. As we shall see, the data confirm the observations noted above and, more importantly, they provide the teacher and learner with genuine examples of recent usage which are not found in other language teaching resources.

In the two corpora of journalistic French and research articles, two million words in all, there are 116 occurrences of the imperfect and the pluperfect subjunctive. It is important to note that in untagged corpora such as these, finding these occurrences involves searching for words with all the possible endings, and then deleting those which are not relevant (words such as

*passions, connaît, russes* etc.). The use of the first, second and third person is distributed as in Table 2.

	CRFJ	CLARF
First person	0	4
Second person	0	0
Third person	34	78
Total	<b>34</b>	<b>82</b>

Table 2: Occurrences of the imperfect and pluperfect subjunctive in the first, second and third person

The four occurrences of the first person do not contradict L'Huillier's comment, as they consist of two historical quotations and two occurrences in a single article which make fun of the tenses.

The occurrences in the third person, of more practical interest to the teacher and learner, are distributed as in Table 3.

	CRFJ	CLARF
Singular <i>être</i>	16	29
Singular <i>avoir</i>	9	12
Singular <i>devoir</i>	0	1
Singular <i>pouvoir</i>	0	5
Singular others	6	18
Plural <i>être</i>	3	6
Plural <i>avoir</i>	0	4
Plural others	0	3
Total	<b>34</b>	<b>78</b>

Table 3: Occurrences in the third person

The 79 occurrences of the verbs *être* and *avoir* are of particular interest here, as they provide examples of language use which are documented by researchers and grammarians as still existing, but where advanced learners will not find significant numbers of examples to guide them in course books. Examples of these and other verbs are provided in the Appendix. As the concordance lines in the Appendix illustrate, in the occurrences of *avoir*, the use of the pluperfect subjunctive as a conditional is well represented (see Grevisse 1993: 1267):

Le plus-que-parfait du subjonctif s'emploie dans la langue soignée avec la valeur du conditionnel passé [...] cet emploi est particulièrement fréquent à la troisième personne du singulier.

This is also the case with the verb *être*, where inversion of subject and verb is particularly evident (12 occurrences out of 16).

While these examples of use can be helpful to learners, they nonetheless pose once again the challenge of containing performance errors and variation. Firstly the form of the past historic and the imperfect subjunctive are sometimes confused:

des services de mise en relation en ligne disparaissaient corps et biens. **Ce fût le cas** notamment du Deuxième Monde de Canal Plus, gigantesque chat room

Secondly, in both corpora there are a number of examples of the use of the subjunctive with *après que*, possibly influenced by the fact that the subjunctive is used with *avant que*:

#### CRFJ

- **après que** Thierry Henry, bousculé par le défenseur Ivan Cordoba, **se mît** dans l'idée de se venger.
- **Après que** Michel Gomez **ait lancé** un souhait
- **Après que** le Président de la République **ait confirmé** le cap de la baisse d'impôt
- Il a été abattu par des gardes du corps **après que** son complice, Bouraoui Al-Ouaer, un autre Tunisien, **eût fait** exploser
- à terminer cette rencontre virile, qui aurait pu dégénérer en première mi-temps **après que** Thierry Henry, bousculé par le défenseur Ivan Cordoba, **se mît** dans
- La liaison ferroviaire avec Nîmes a ainsi été interrompue **après qu'un** TGV **ait été** frappé par la foudre, avant que la SNCF ne suspende tout

#### CLARF

- contaminées auraient été écoulés en Europe continentale pendant plus d'un an **après que** les autorités britanniques les **aient proscrites** de leur territoire.
- Le Parlement européen récompense ici l'action d'un leader charismatique **après que** le Comité organisateur du Prix Nobel de la Paix **ait décerné** en 1996 ce
- l'objectif des médias, qui consistait à informer le plus vite possible le public **après que survienne** un événement, fut remplacé par cette exigence inimaginable
- autres sous-cultures musicales existantes à l'époque : découverte soudaine, **après que** chaque nouvelle " mode " **ait déjà su** mettre en place une forme d'esthétique
- affectés de phénomènes de survivance dus à la permanence de toute ville, même **après qu'aient disparu** les facteurs qui furent à l'origine de sa création.

When confronted by these concordance lines the teacher has to decide how to present the results of the corpus search to the learners. For learners at intermediate level it may be tempting to omit the uses of the subjunctive with *après que* as performance errors, but this could be ignoring the reality of language change (see, for example Leeman-Bouix 1994; Ball 2000).

*Après que* is a particularly common example of how variation can exist in the language of educated writers such as novelists and journalists. Desrosiers (2000: 14) cites De Beauvoir and Mauriac using the subjunctive, finding examples of both the indicative and the subjunctive in the latter. Summarising the explanations put forward by linguists, he cites the commonly held view that the subjunctive is used by analogy with its use with *avant que*, but also refers to Leeman-Bouix's observation that prepositions which are followed by the infinitive, such as *avant* and *après*, are used with the subjunctive when used

as conjunctions. In the context of this study it is interesting to note that he also makes the point that, if the indicative were used with *après que*, it would be in the past anterior, a tense which he describes as:

un temps que l'usage courant tend à délaisser, le trouvant vieilli, très soutenu, sinon littéraire. Sa décadence est accentuée par le fait qu'il est en principe employé dans la subordonnée en corrélation avec un passé simple dans la principale, lequel est tombé en désuétude dans la langue parlée. En insistant en faveur de l'indicatif, les grammairiens donnent ainsi l'impression de demander aux locuteurs de s'adonner à une sorte d'acharnement linguistique pour faire survivre un temps de verbe qui, aux yeux d'un bon nombre, a quelque chose de guindé aujourd'hui.

The teacher or learner consulting these easily available corpora may not have the time available to follow up the cases of variation encountered, but will at least be able to pass on to the learners the fact that they exist. The amount of data available in the corpora consulted here is small, but it nonetheless provides examples, many of them relatively recent, of occurrences of the imperfect and pluperfect subjunctive, thus introducing advanced learners to an aspect of native speaker language use which they will encounter in newspapers and literary texts. By using these in class, perhaps in support of a single example occurring in a text, the teacher can give the learners attested examples of actual language use. While there is insufficient data to allow generalisations about which verbs need to be taught, the teacher can at least be confident that the concordance lines presented to the learners represent recent occurrences of these rare forms.

#### 4. Conclusion

This study confirms the hypothesis that the data on the subjunctive are easily available and small corpora provide a rich source of occurrences of naturally occurring discourse to complement the necessarily limited number of examples provided in course books and grammars. It is arguable that this choice of the aspect of the language to study, namely the subjunctive, provides a view of corpus consultation as overly complicated, and does not present a realistic view of the way that many aspects of the language can be presented to learners in concordances which clearly reveal patterns of use and where variation in native speaker language use is not an issue. The aim is to show, however, that the aspects of language use where there is variation pose challenges for the teacher and for the advanced learner which can be seen in a positive light. Firstly they clearly show that language use by native speakers, even educated speakers and well known writers like Alexandre Dumas, does not always conform to grammatical rules, and that a descriptive rather than a prescriptive approach can reveal aspects of use which cannot be learned from traditional resources such as course books, grammars and dictionaries. As McCarthy (1998: 78) notes, teachers and learners become slaves to the rules, forgetting that "discourse drives grammar, not the reverse". Secondly, by clearly illustrating the presence of variation, corpus data have

the potential to increase language awareness for both teachers and learners. Noticing the use of the subjunctive with *après que* in the corpus data, for example, has made me more aware of its existence in the spoken language of French native speakers. Finally, this increase in language awareness can encourage teachers to take greater control of the content of what they teach, in particular of the ways in which they present grammatical rules to advanced learners in areas where variation is not uncommon.

As mentioned above, studies of the use of corpora by learners, particularly through giving learners direct access to the corpus data, have elicited negative as well as positive responses from learners. This study has tried to mitigate such reactions by suggesting that teachers could use the data in search of additional teaching materials. Finally, while, from an applied linguistic or pedagogical point of view, the focus of this study may appear to fly in the face of more current communicative or task-based approaches, it is important to remember that it is not intended to be seen as an approach to language learning, but rather as a way in which grammatical points can be included in a class where a communicative approach is used. The corpus-based examples would thus have the advantage of providing the learners with examples of actual use from both journalistic and literary sources. The fact that the resources are freely available online and easy to consult with minimal training, particularly in the case of Lextutor, should make them attractive to teachers and learners, although empirical studies would be necessary to confirm this and to investigate the reactions of both teachers and learners to examples of variation.

## REFERENCES

- Aston, G. (1997). Small and large corpora in language learning. In: B. Lewandowska-Tomaszczyk & J.P. Melia (eds.), *Practical applications in language corpora* (pp. 51-62). Lodz: Lodz University Press.
- Aston, G. (2001). Text categories and corpus users: a response to David Lee. *Language Learning & Technology*, 3, 1, 73-76.
- Ball, R. (2000). *Colloquial French grammar: a practical guide*. Oxford: Blackwell.
- Bernardini, S. (2002). Exploring new directions for discovery learning. In: B. Kettemann & G. Marko (eds.), *Teaching and learning by doing corpus analysis* (pp. 165-182). Amsterdam/New York: Rodopi.
- Biggs, P. & Dalwood, M. (1976). *Les orléanais ont la parole*. London: Longman.
- Boulton, A. (2010). Learning outcomes from corpus consultation. In: M. Moreno Jaén, F. Serrano Valverde & M. Calzada Pérez (eds.), *Exploring new paths in language pedagogy: lexis and corpus-based language teaching* (pp. 129-144). London: Equinox.
- Boulton, A. (2012). Corpus consultation for ESP: A review of empirical research. In: A. Boulton, S. Carter-Thomas & E. Rowley-Jolivet (eds.), *Corpus-informed research and learning in ESP: issues and applications* (pp. 261-291). Amsterdam: John Benjamins.

- Cappeau, P. & Sejjido, M. (2005). *Les corpus oraux en français (inventaire 2005 v.1.0)*. [http://www.culture.gouv.fr/culture/dglf/recherche/corpus\\_parole/Inventaire.pdf](http://www.culture.gouv.fr/culture/dglf/recherche/corpus_parole/Inventaire.pdf) (07.05.2012).
- Chambers, A. (2005). Integrating corpus consultation in language studies. *Language Learning & Technology*, 9, 2, 111-125.
- Chambers, A., Farr, F. & O'Riordan, S. (2011). Language teachers with corpora in mind: from starting steps to walking tall. *Language Learning Journal*, 39, 1, 85-104.
- Chambers, A. & Le Baron, F. (éds.) (2007). *Le corpus Chambers-Le Baron d'articles de recherche en français*. <http://www.ota.ahds.ac.uk/headers/2527.xml> (14.05.2013).
- Chambers, A. & Rostand, S. (éds.) (2005). *Le corpus Chambers-Rostand de français journalistique*. Oxford (Oxford Text Archive). <http://www.ota.ahds.ac.uk/headers/2491.xml> (14.05.2013).
- Cheng, W., Warren, M., & Xun-feng, X. (2003). The language learner as language researcher: putting corpus linguistics on the timetable. *System*, 31, 2, 173-186.
- Cobb, T. (1997). Is there any measurable learning from hands on concordancing? *System*, 25, 3, 301-315.
- Conrad, S. (2004). Corpus linguistics, language variation, and language teaching. In: J. Sinclair (ed.), *How to use corpora in language teaching* (pp. 67-85). Amsterdam: John Benjamins.
- Cook, G. (2001). "The philosopher pulled the lower jaw of the hen". Ludicrous invented sentences in language teaching. *Applied Linguistics*, 22, 3, 366-387.
- Council of Europe (2001). *Common European framework of reference for languages: learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Desrosiers, J. (2000). *Après que et le subjonctif. L'Actualité langagière*, 33, 14.
- Ellis, N. (2002). Frequency effects in language processing. A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.
- Firth, J.R. (1957). *Papers in linguistics. 1934-1951*. London: Oxford University Press [cited in Cook, 2001].
- Gabrielatos, C. (2005). Corpora and language teaching: just a fling or wedding bells? *Teaching English as a Second Language – Electronic Journal*, 8, 4, 1-35. <http://tesl-ej.org/ej32/a1.html> (14.05.2013).
- Girodet, J. (1986). *Pièges et difficultés de la langue française*. Paris: Bordas.
- Grevisse, M. (1936 [ed. Goosse, A. (1993)]). *Le bon usage. Grammaire française*. Paris, Louvain-la-Neuve: Duculot.
- Hanse, J. (1991). *Nouveau dictionnaire des difficultés de la langue française*. Paris: Duculot.
- L'Huillier, M. (1999). *Advanced French grammar*. Cambridge: Cambridge University Press.
- Johns, T. (1988). Whence and whither classroom concordancing? In: T. Bongaerts, P. De Haan, S. Lobbe & H. Wekker (eds.), *Computer applications in language learning* (pp. 9-27). Dordrecht: Foris.
- Leeman-Bouix, D. (1994). *Les fautes de français, existent-elles?* Paris: Seuil.
- McCarthy, M. J. (1998). *Spoken language and applied linguistics*. Cambridge: Cambridge University Press.
- McCarthy, M. J. & Carter, R. (1995). Spoken grammar: what is it and how can we teach it? *ELT Journal*, 49, 3, 207-218.
- Meunier, F. & Gouverneur, C. (2009). New types of corpora for new educational challenges: Collecting, annotating and exploiting a corpus of textbook material. In: K. Aijmer (ed.), *Corpora and language teaching* (pp. 179-201). Amsterdam: John Benjamins.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.

- Schmidt, R. (1995). Consciousness in foreign language learning: A tutorial on the role of attention and awareness in learning. In: R. Schmidt (ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Hawai'i: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Scott, M. (2004). *Wordsmith tools version 4.0*. Oxford: Oxford University Press.
- Sinclair, J. (1988). *New directions in English dictionaries*. Unpublished manuscript [cited in Cook, 2001].
- Sinclair, J. (2004). New evidence, new priorities, new attitudes. In: J. Sinclair (ed.), *How to use corpora in language teaching* (pp. 271-299). Amsterdam: John Benjamins.
- Sweet, H. (1899 [ed. Mackin, R. (1964)]). *The practical study of languages: a guide for teachers and learners*. London/Oxford: Oxford University Press [cited in Cook, 2001].
- Timmis, I. (2010). Teachers telling tales: Exploring materials for teaching spoken language. In: F. Mishan & A. Chambers (eds.), *Perspectives on language learning materials development* (pp. 63-85). Oxford: Peter Lang.
- Wen-Ming, H. & Hsien-Chin L. (2008). A case study of corpus-informed online academic writing for EFL graduate students. *CALICO*, 26, 1, 28-47.
- Widdowson, H. G. (1983). *Learning purpose and language use*. Oxford: Oxford University Press.
- Yoon, H. & Hirvela, A. (2004). ESL student attitudes towards corpus use in L2 writing. *Journal of Second Language Writing*, 13, 257-283.

## Appendix

### Verbs in the imperfect and pluperfect subjunctive

#### *Le Corpus Chambers-Rostand de français journalistique*

##### *Avoir and être*

1. par la loi, pas par le dialogue social; par la norme, pas par le contrat. Or il eût été essentiel, et pour le coup véritablement novateur, de tout faire
2. les Tchéchènes avaient eu le temps de mettre leurs menaces à exécution, le résultat eût été pire. Mais quoique les Russes aient eu l'initiative de l'attaque, on éprouve
3. BADEN-BADEN (Allemagne) de notre envoyée spéciale Alphonse Allais eût été content : à Baden-Baden, la ville est bien à la campagne.
4. Shéhérazade est de tous les temps et de tous les pays. Sans doute eût-elle reçu le Prix Nobel de littérature si le prix avait existé en ces temps
5. proportion du SMIC, a-t-il précisé. Une déclaration qui laisse un goût amer. L'eût-il fait au moment où la gauche était aux affaires

1. Lionel Jospin a compris, avec retard, que la négociation, aussi incertaine fût-elle, était le seul chemin envisageable.
2. Nicolas Philibert ne s'en est jamais caché, faisant valoir qu'un cinéaste - fût-il documentariste - n'est pas journaliste.
3. lors de la saison 2001-2002, c'est peut-être trop pour un joueur de football, fût-il du calibre du milieu de terrain international d'Arsenal.
4. il tombe, il attend longtemps la voiture de dépannage, s'en irrite fût-ce avec sa courtoisie naturelle,
5. il n'existe pas d'oasis épargnées par les flux qui agitent la planète - fût-ce une salle de classe auvergnate. Pas plus qu'il n'existe d'êtres épargnés
6. Il est moins lié à l'extrême droite qu'il ne le fût, et beaucoup plus à la situation internationale
7. détermination du gouvernement à juguler l'immigration clandestine, fût-elle de transit vers l'Eden anglais, montre ici ses limites.
8. contre l'Irak. On ne saura jamais quel a été le principal tourment moral qui fût assez intense pour conduire le scientifique britannique David Kelly
9. L'impérialisme libéral, fût-il animé des meilleures intentions, risque d'alimenter un ressentiment
10. " Le Tour, évidemment, ne ressemble en rien à une classique, fût-elle la plus complète de toutes. Mais Guimard n'a pas tort, car Lance Armstrong
11. était réputée pour son sens de l'accueil de l'autre, aussi étrange et étranger fût-il.
12. Reste qu'on est en droit d'attendre d'un chef de gouvernement, fût-ce au cœur de la trêve estivale, plus de détermination et de précision
13. Encore faudrait-il que la République fût juste, généreuse, attentive à tous ses enfants. Elle ne l'est pas.
14. Mais, pas de chance, le dernier de poule est un compétiteur coriace, fût-il italien. Trévise a quand même défait les London Wasps (32-12)
15. Quoi qu'il en fût, cette découverte de l'homosexualité supposée de Molière a visiblement beaucoup
16. Iran soit assez flexible pour changer du jour au lendemain à cause d'un prix, fût-il prestigieux comme le Nobel de la paix.

##### Other verbs

1. Et elle préférerait que la Maison Blanche se préoccupât un peu plus de la crise économique
2. Trois petits quarts d'heure suffirent pour que tout rentrât dans l'ordre.
3. Le simple bon sens aurait voulu que l'on s'attachât depuis vingt ans à réduire progressivement puis à arrêter cette pêche

4. ne vaudrait-il pas mieux que l'ONU accompagnât, pour " l'encadrer ", une intervention militaire considérée comme inévitable
5. a répondu à l'annonce d'une vieille dame du coin, qui souhaitait que quelqu'un mît de l'ordre dans ses archives.
6. après que Thierry Henry, bousculé par le défenseur Ivan Cordoba, se mît dans l'idée de se venger.