



## Variation in first and second language French: the case of ‘parce que’

Mireille Bilger, Henry Tyne

### ► To cite this version:

Mireille Bilger, Henry Tyne. Variation in first and second language French: the case of ‘parce que’. K. Beeching, N. Armstrong, F. Gadet. Sociolinguistic Variation in Contemporary French, John Benjamins, pp.233-259, 2009. hal-00416549

**HAL Id: hal-00416549**

**<https://hal.science/hal-00416549>**

Submitted on 21 Sep 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**Running title: *Parce que* in first and second language French**

**Variation in first and second language French: The case of *parce que***

Mireille Bilger (University of Perpignan, France) and Henry Tyne (University of Nancy, France)

*This paper looks at the distribution of parce que (subordinating and non-subordinating because) in native and non-native (students at a UK university) speakers of French. In both first and second language data, the use of parce que sequences is found to vary quantitatively from one type of speech situation to another. In the non-native speaker group, this variation is found irrespective of year of study: first year learners are often just as variable (even more so in some instances) as fourth year learners in terms of the distribution of parce que sequences..*

**Introduction**

Over the past few decades, there has been a massive growth in the use of corpus enquiry for the study of language and variation. Large corpora, which include different types of texts or oral data, can be trawled for different features. The study of syntax, in particular, has come to the fore in corpus studies. Numerous works have shown how grammatical features are distributed unevenly according to the type of text or situation.

In this chapter, which focuses on the distribution of sequences introduced by *parce que* ('because'), we look at variation in native (French L1) and non-native (French L2) speech. We identify areas where there is little difference between native and non-native speakers, irrespective of the particular year of study. We also identify areas of divergence between the groups of speakers.

## Looking at variation

### *Distributional frequency of features and text-types*

Different approaches have been developed to describe or classify types of situation or text. These often typically relate to either the general situation in which the language occurs or the ‘discourse type’ or activity (description, narration, evaluation, etc.). However, they remain quite general descriptors (e.g. formal *vs.* informal typically encountered in approaches to L2 variation), and do not allow us to get a detailed bearing on the many language forms we find and their different distributions.

The notions of *genre* (Malrieu & Rastier 2001; Kerbrat-Orecchioni & Traverso 2004, *Langage et Société* 1999), *register* (Biber 1988) or *style* (Gadet & Tyne 2004), though quite polysemous, can also be a starting point. Whilst it would be wrong to consider these as neatly interchangeable synonyms (in particular, because of the particular history of each of them<sup>1</sup>), what they have in common is a conceptualization of variation in relation to the constraints imposed by the type of communication rather than simply seeing variation in language as the reflection of given extralinguistic differences (e.g. attention to speech in response to the formal setting).

Numerous authors have attempted to classify text types and many have proposed to look at specific elements in relation to these (for French: Bronckart *et al.* 1985; Adam 1992; Malrieu & Rastier 2001). However, these works are typically built upon pre-existing classifications and they generally only set out to study the differences between these pre-defined types (though this might involve modifying the original classification).

---

<sup>1</sup> If we take *genre*, for instance, Branca-Rosoff (1999) has shown how this term, typically associated with the study of literary texts, where it is a category of literary composition, can be extended to cover speech.

Other approaches purport to study variation in a more inductive way. This is the case for the work of Biber (e.g. 1988) on English, in which the classification criteria are established according to the features found in the texts themselves. Thus, it is possible to observe in a far more detailed manner the ‘external’ and ‘internal’ properties of texts, thereby gaining a more subtle appreciation of variation in terms of the different distributions of features. For example, at a very simple level, we can expect not to find much reference to the first and second persons in academic writing (see Chafe 1982) since it does not involve face to face contact with the interlocutors, nor does it require overt expression of the speaker’s personal involvement. This feature, considered alongside others in a bundle, would allow us to identify a particular type of written or spoken text. We believe this type of approach could be useful for understanding learners’ speech and for looking at practice in language teaching.

### *Variation in L2*

Variation in L2 is often associated with the notion of sociolinguistic competence. However, it is by no means the preserve of sociolinguistically motivated studies and has been studied from a developmental perspective (particularly unsystematic variation), whereby the learner’s interlanguage is seen to be an unstable variety that develops over time. Ellis (1985), for example, has argued that variation is an essential condition for interlanguage development as the learner first internalises new linguistic forms, progressively organises form-function relationships, then eliminates redundant forms. Others (e.g. Tarone 1982) have looked at variation in relation to the degree of attention or casualness in speech. Whereas these studies have typically concentrated on variable use of language where native speakers show invariant use, the more recent paradigm focusing on sociolinguistic competence (see Mougeon *et al.* 2002) marks a break with this and

variant forms are studied where those forms are also used by native speakers of the target language.

Researchers working on L2 sociolinguistic competence have typically applied variationist methodology (Dewaele 2007). Thus, sociolinguistic competence is to do with learners' ability to produce target language-like variability (identified in the target community). The variables that have typically been studied are in phonology, lexis and selected grammatical features such as *ne* omission<sup>2</sup>. However, as Towell (2004) has observed, relatively little work has been carried out in areas beyond feature sets in sociolinguistically motivated studies of L2 acquisition and the wider study of syntax is generally associated with acquisition (i.e. not sociolinguistic) studies: for example, both the UG approach (Hawkins 2004) and the functional / pragmatic approach (Klein & Perdue 1992) concentrate on syntax.

Certain L2 studies have looked at syntax in relation to complexity and units of production (e.g. Crookes 1990; Skehan 1996) from a psycholinguistic perspective (planning and processing) in relation to task types. Other approaches have looked at typological constraints on event representations in discourse planning, whereby learners use differing degrees of granularity (i.e. cutting the event up into a series of micro stages or processes) in the L2 according notably to the typical ways and means of retelling events in their L1 (ex. see Noyau *et al.* 2005). Elsewhere, in an attempt to identify distinct stages in acquisition, a range of features (not all syntactic) have been looked at by Bartning & Schlyter (2004). For example, the initial stage in the acquisition of L2 French is characterised by nominal utterances and limited grammatical expression whereas, at the other end of the spectrum, the upper advanced stage is characterised by stabilised inflectional forms, the use of different connectors and enhanced informational packaging. Whilst these authors situate variation (they refer to "*une richesse d'expressions qui permettent un choix dans*

---

<sup>2</sup> See Mougeon *et al.* (2002) for an overview of the different features that have been studied in French L2.

*le répertoire*”, 2004: 296, ‘a rich variety of expressions that allow the speaker to choose from a repertoire’) at the advanced level, they do not specifically take variation as a means of identifying specific stages.

#### Parce que *sequences*

Syntactic units may vary considerably in length from one speech situation to another. Among the various ‘devices’ for unit lengthening, the presence of subordinating conjunctions is often signalled in the literature. Givón (1979), for example, suggested that informal, unplanned discourse tends to favour the pragmatic mode of communication as opposed to the syntactic mode, with a tendency for “*loose coordination*” over “*tight subordination*”. In the study of L2, multi-clausal utterances are typically taken as an indicator of complexity and L2 competence (e.g. Skehan 1996; Foster *et al.* 2000). We look here at the use of sequences introduced by *parce que* using the Pronominal Approach (henceforth PA – see Blanche-Benveniste *et al.* 1987; Blanche-Benveniste *et al.* 1990; Van den Eynde 1998).

According to PA, the verb is the basic starting point for analysis as it selects the number of valency-bound elements (complements) and determines the relations between them. An important identification procedure in PA involves reducing all valency-bound elements, no matter how long they are, to minimal forms (proportionality testing) as in the following example<sup>3</sup>:

---

<sup>3</sup> This example and subsequent ones are taken from our data. Transcription conventions are from DELIC (see *Recherches sur le français parlé* 18): silent pauses are indicated by ‘+’ and overlapping speech is indicated using chevrons. Corpus codes are given in brackets at the end of each example. English translations are included as an aide to comprehension only.

- (1) a. j'utilise **une une enquête qui est qui est fait chaque année ou tous les deux ans par l'entreprise et qui s'appelle Vous et votre entreprise**  
(KarenFC)  
'I use **a a survey that is that is carried out each year or every two years by a company and that is called You and your company**'
- b. je l'utilise / j'utilise **ça**  
'I use **it** / I use **that**'

Beyond the valency kernel, 'optional', non-valency elements are also identified. Traditionally, 'optional' sequences introduced by *parce que* are treated as subordinating conjunctions (e.g. Hawkins & Towell 2001). However, this term covers quite a range of forms and grammatical relations. PA enables us to distinguish between those sequences that are "governed"<sup>4</sup> and those that are "non-governed". We can identify governed sequences in that they share certain syntactic properties with valency elements:

Proportionality with a pronominal paradigm (personal pronoun, interrogative pronoun, etc.)

- (2) a. c'était facile de trouver beaucoup d'informations **parce que j'étais à Paris** (LisaFP)  
'it was easy to find lots of information **because I was in Paris**'
- b. **pourquoi** c'était facile de trouver beaucoup d'informations  
'**why** was it easy to find lots of information'

---

<sup>4</sup> This is adapted from the original terminology by Bilger and Campione (2002). See also Tyne (forthcoming).

## Modality contrasts

- (3) c'était facile de trouver beaucoup d'informations **pas parce que** j'étais à Paris  
**mais parce que** j'avais un accès Internet à la maison  
'it was easy to find lots of information **not because** I was in Paris **but because** I  
had an Internet connection at home'

## Extraction using *c'est...que / qui*

- (4) **c'est** parce que j'étais à Paris **que** c'était facile de trouver beaucoup  
d'informations  
'**it's** because I was in Paris **that** it was easy to find lots of information'

Certain non-governed sequences can be analysed from a discourse or macro-syntax point of view. This is particularly the case for those sequences introduced by *parce que* which, as Debaisieux has shown (2002, 2004), can have a distinct 'non-subordinating' nature, where they are detached (physically and / or semantically) from the main predicative construction. In the following extract, *parce que* introduces a remark or a commentary on the term *décharges sauvages*. None of the above properties that apply to governed sequences can be identified in this case.

- (5) en premier lieu il s'agit de sensibiliser + les populations + déjà pour éviter les  
décharges sauvages + **parce que des décharges sauvages nous en rencontrons**



**malheureusement toujours dans nos campagnes + dans nos ravins n'est-ce pas** (Val1PUB)

‘first of all it’s a matter of informing + the public + already so as to avoid fly-tipping + **because fly-tipping we still see it unfortunately in our countryside + in our valleys don’t we**’

This study focuses on the syntactic description (i.e. governed *vs.* non-governed *parce que* sequences), and no attempt is made to look at finer semantic issues, pragmatic functions and cross-speaker construction<sup>5</sup>: we concentrate, then, on the ‘simple’ opposition between governed and non-governed.

## Data

The L1 data we use here come from the *Corpus de Reference du Français Parlé* (CRFP – see *Recherches sur le français parlé* 18); they consist of three extracts of a similar size, selected randomly, for each of three different situations: public speech<sup>6</sup> (*parole publique*), professional speech (*parole professionnelle*), private speech (*parole privée*). The L1 dataset comprises approximately 100 000 words for seven different speakers.

The L2 data was collected at the University of Surrey in 2003. It includes recordings of twenty students in their first and fourth years of study, taken across three different situations: formal presentation<sup>7</sup>, formal conversation/defence, informal conversation. The L2 data amounts to approximately 65 000 words. In both cases, the samples include roughly equal numbers of

---

<sup>5</sup> As scholars working with PA have shown (e.g. Blanche-Benveniste *et al.* 1990), the number of speakers involved in the construction is not crucial to the syntactic analysis (cf. the notion of “*locuteur collectif*”).

<sup>6</sup> These appear in the graphs (see below) in short: PUB, PRI, PRO.

<sup>7</sup> These situation types appear in the graphs (see below) in short: FP (formal presentation), FC (formal conversation), IC (informal conversation).

male and female speakers. Raw data and average frequencies are presented in tables 1 to 4 in the appendix. Concordances were studied using the programme *Contextes*<sup>8</sup>.

Given that it is difficult to reproduce speech situations from one speaker to the next when working with naturally occurring speech (cf. the problems of working with pre-existing categories), there is no exact comparability of these two studies from the descriptor point of view. Also, as is often the case in studies of L2 data, the learners (from whom it is not always easy to obtain lengthy recordings of flowing speech) form a subset of educated, university-based students aged around twenty. This is not the case for the L1 data, which, whilst also including adults' speech only, is made up of a varied selection of native speakers of French. However, given what we know about the importance of communicative constraints (see Jahandarie 1999; Koch & Österreicher 2001), the tendency for similar communicative constraints is taken as a basis for any attempt to make comparisons here. Our aim here, then, is to bring together two different datasets containing selected samples of oral data which are sufficiently distinct from a functional point of view so as to enable comparisons and contrasts from one type to another. We would add that it is not our intention here to formulate statements about L1 and L2 French in general.

## **Results**

### ***Inter-speaker-group study***

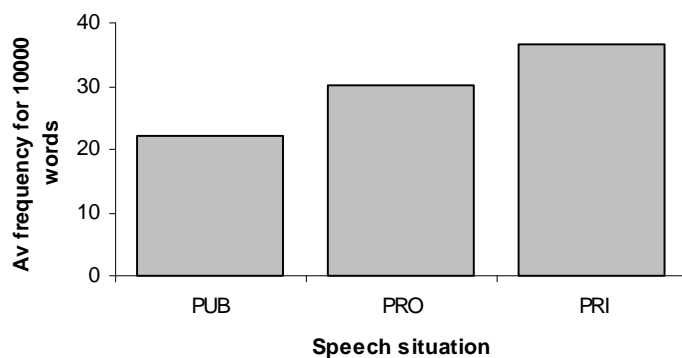
---

<sup>8</sup> Created by J. Véronis (<http://sites.univ-provence.fr/veronis/logiciels/Contextes/index-fr.html>, page accessed 21/10/08).

### *Governed parce que sequences*

The frequency of use<sup>9</sup> of governed *parce que* sequences varies across speech situations in the L1 data (see fig. 1). We note that it is in private speech that the greatest number of occurrences is found. There is a significant<sup>10</sup> difference between public and private speech (see table 4).

Fig 1. L1 governed *parce que*



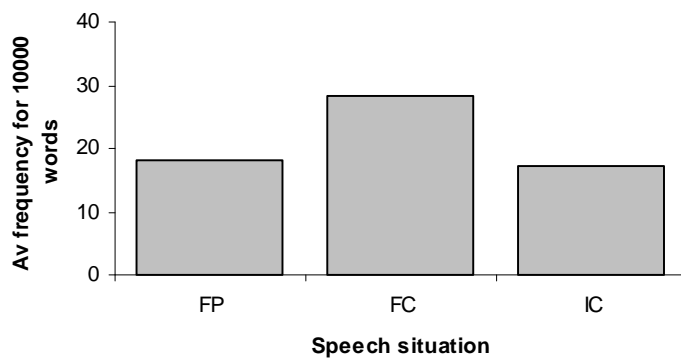
In the L2 data, the frequency of use of governed *parce que* sequences also varies across the three speech situations (see fig. 2). There is a significant difference (see table 5) between formal presentation and formal conversation, and between formal conversation and informal conversation. There is no significant difference between formal presentation and informal conversation (see table 5).

Figure 2. L2 governed *parce que*

---

<sup>9</sup> Following Biber (e.g. 2000), we look at frequency of use per given number of words produced.

<sup>10</sup> Chi<sup>2</sup> values (tables 5 and 6) are given in the appendix.



*Non-governed* parce que *sequences*

Debaisieux found in her L1 data that non-governed use of *parce que* sequences represented the overwhelming majority of cases:

[...] l'observation des données authentiques montre que ces exemples constituent l'écrasante majorité des occurrences relevées dans les corpus oraux (78% des cas), comme nous l'avons montré dans une étude antérieure portant sur plus de 3000 exemples, et qu'ils présentent en outre des configurations syntaxiques beaucoup plus complexes que les cas d'école généralement cités.

(Debaisieux 2004: 53)

'[...] a look at attested data shows us that these examples account for a huge majority of cases found in the spoken corpora (78% of all cases), as we showed in an earlier study of 3000 examples, and that they display, among other things, syntactic features that are far more complex than the textbook cases generally cited.'

In our data, the use of governed and non-governed *parce que* sequences is more even, i.e. the overall quantity of governed vs. non-governed sequences is similar (see first two lines of tables 1 and 2). Whereas there is no significant difference in frequency of use for non-governed *parce que* sequences across the different speech situations in the L1 group (see table 4), the formal conversation stands out (significant – see table 5) in the L2 data, being the situation in which the most sequences were used (cf. use of governed *parce que* sequences above). See figures 3 and 4 below.

Fig 3. L1 non-governed *parce que*

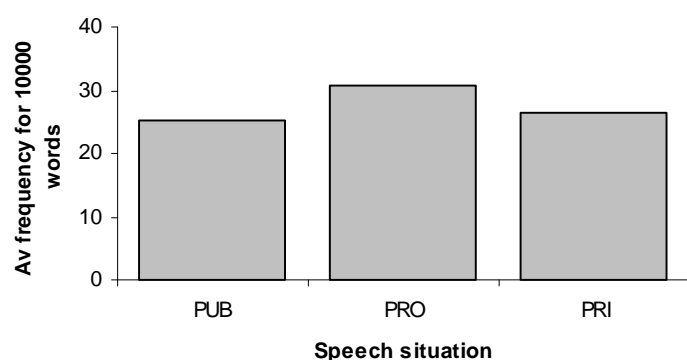
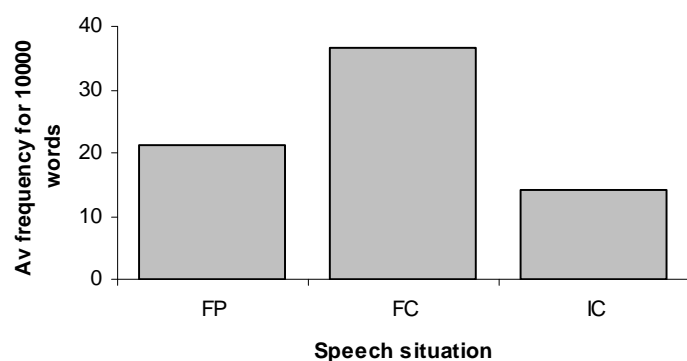


Fig 4. L2 non-governed *parce que*



### Grouping

*Parce que* can introduce a sequence which is itself made up of *quand / si / dès que* / etc. ('when / if / as soon as / etc.') + predicative construction (see examples 6-8). The *parce que* sequence relates to the predicative construction to the left, whether governed or not, whereas the sequence it is introducing (*quand, si*, etc.) is governed by the following predicative construction.

- (6) enfin ça c'est une autre histoire **parce que quand je vois des des vieux de de quatre-vingt-cinq ans qui qui nous donnent des leçons de de jeunesse et qui expliquent aux jeunes comment il faut agir + ça paraît surprenant** (Aix1PRO)  
'well now that's another story **because when I see old people of of eighty who who tell us how to be young and tell young people how to behave + it's a bit surprising**'
- (7) et maintenant dans l'imprimerie ce qu'on demande à un imprimeur c'est non pas + d'être un artiste c'est d'être un gestionnaire + pourquoi **parce que s'il a suffisamment d'argent + il peut acheter un matériel suffisamment + sophistiqué** (Aix1PRO)  
'and now in printing what they ask of a printer it's not + to be an artist it's to be a manager + why **because if he's got enough money + he can buy sufficiently + sophisticated material**'
- (8) c'était le modèle parfait quoi + et puis et en même temps le contact parisien + **parce que dès qu'on avait fini ses classes ici qu'on avait son prix + on avait des contacts avec Paris** (Aix1PUB)

‘he was the perfect role model you know + and then and at the same time the Parisian contact + **because as soon as you’d finished your tuition here and you had your prize [from the conservatoire] + you had contacts with Paris**’

Speakers can vary between the ‘successive’, separated manner and the ‘inclusive’, grouped manner of production:

- (9) il y a quand même des inconvénients quand même **parce que quand** on ressort et puis qu’il fait quarante degrés dehors ben + on (n’)a pas l’habitude en été **quand** je suis en vacances j’ai du mal à supporter la chaleur **parce que** j’ai pas l’habitude + toute la journée je suis au frais donc (Bes2PRI)

‘there are nonetheless negative aspects all the same **because when** you go back out and it’s forty degrees outside well + you’re not used to it in summer **when** I’m on holiday I find it difficult to put up with the heat **because** I’m not used to it + so all the year round I’m in a cool environment’

As Debaisieux (2004) has observed, *parce que* sequences in the case of grouping are generally non-governed. And, whilst we only found a handful of instances of grouping in our data (24 occurrences in all – see tables 3 and 4), most were indeed non-governed (20 out of 24). We would point out three things: firstly, the only occurrences of *governed* use of *parce que* + *comme* / *dès que* / etc., occurred in the L1 data; secondly, no examples of *parce que* + *dès que* and *parce que* + *lorsque* (‘when’) were found in the L2 data; finally, whilst the low number of occurrences means that we cannot check for statistical significance, we would tentatively point

out that there does not appear to be any particular preference in either group in terms of the situation in which grouping occurs (see tables 3 and 4).

#### *Intra-speaker-group study: L2*

##### *Governed *parce que* sequences*

The results show that there is an increase in use of governed *parce que* sequences in the formal conversation in both groups (see figs. 5 and 6). Whilst this difference was significant for the first year group (see table 6), there was no significant variation in the fourth year group.

Fig 5. Y1 governed *parce que*

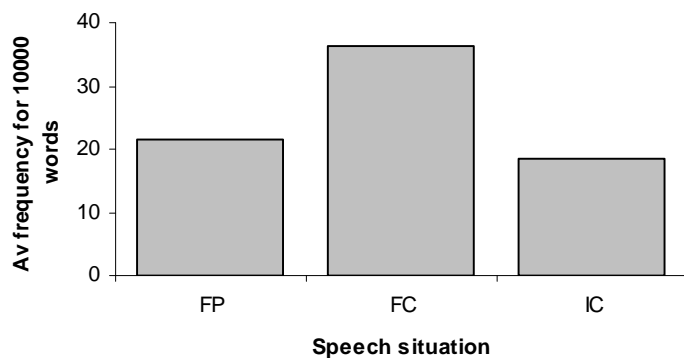
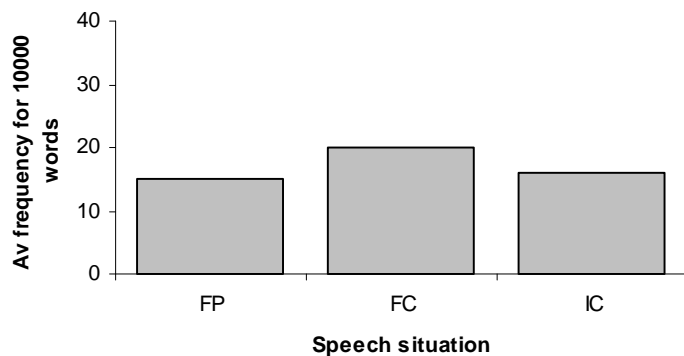


Fig 6. Y4 governed *parce que*





### *Non-governed parce que sequences*

Concerning non-governed *parce que* sequences, the intra-group study shows that whilst the formal conversation stands out as the situation in which the greatest number of non-governed *parce que* sequences was used for both groups (significant – see table 6), the difference is most marked in the first year data (compare figs. 7 & 8). In the fourth year data, however, the overall use of these sequences is slightly higher than that for the first year group. There is no significant difference in use between the formal presentation and the informal conversation in the fourth year group (see table 6).

Fig 7. Y1 non-governed *parce que*

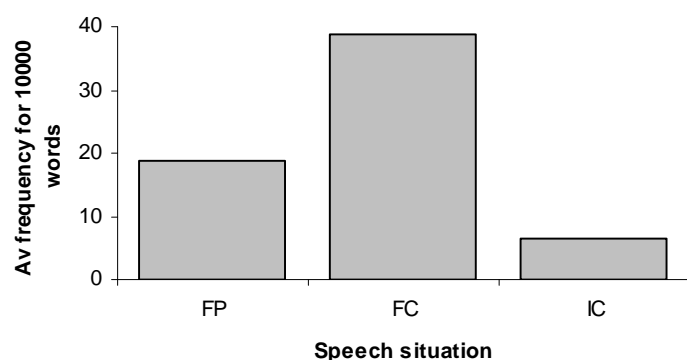
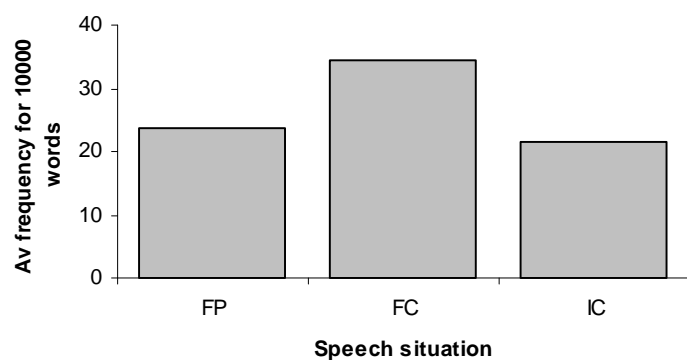


Fig 8. Y4 non-governed *parce que*



### *Grouping*

There were 12 occurrences of grouping in the L2 data; 9 of these were in the fourth year group (see table 4). And, whilst the occurrences in the first year data were of the same type (*parce que* + *quand*), three different types were found in the fourth year data: *parce que* + *comme* / *quand* / *si* ('because + like / when / if'). As for the distribution of this phenomenon across speech situations, we would again stress that, given the low number of occurrences, no sound conclusions can be reached.

### **Discussion**

Whilst the greatest use of governed *parce que* sequences occurred in the private speech situation in the L1 group, it occurred in the formal conversation in the L2 group (although the difference was only significant in the first year group). The highest number of non-governed *parce que* sequences was also found in the formal conversation in the L2 group (first and fourth years). This situation typically requires the speaker to explain, nuance or hedge as (s)he attempts to justify or readjust his/her position. Non-subordinating *parce que* would appear to be particularly suited to this activity (see example 10 below).

- (10) Sam: euh est-ce que est-ce que + enfin le fait d'avoir fait toutes ces recherches est-ce que ça est-ce que ça a changé votre avis ou est-ce que ça va changer vos habitudes euh gastronomiques + vos habitudes alimentaires +

Debbie: ben ça ça a changé euh + comme j'ai dit ça ça a changé euh mes idées  
**parce que au début je croyais qu'il y avait un gros problème + au niveau de la**  
**sécurité alimentaire mais en fait euh +**

Philippe: hm

Debbie: **ben je crois que c'est <plutôt euh>**

Philippe: <ah oui> (DebbieFC)

Sam: 'er has has + I mean the fact that you did all of this research has that has that  
altered your opinion or will that alter your dining habits + your eating habits +

Debbie: well that that has altered er + like I said that that has altered my ideas  
**because at first I thought there was a big problem + with food safety but in in**  
**fact er +**

Philippe: hm

Debbie: **well I think that it's <more likely er>**

Philippe: <ah yes>'

The results are different from those obtained for the L1 group, where there is no significant difference in the amount of non-governed *parce que* sequences used. Whilst this may be evidence of an L1-L2 difference (i.e. non-native speakers 'overuse' *parce que* whereas native speakers have a greater range of connectors from which to choose), it may rather simply be a difference between L1 and L2 data, since the pattern of use of non-governed *parce que* sequences is found in *both* first and fourth year L2 speakers. This is important since other researchers (e.g. see in Dewaele & Mougeon 2002) have shown that very advanced speakers of L2, i.e. in their final year of study at university, having spent time in the target language community, can

generally be expected to use sociolinguistic forms in ways approaching those found in native speakers of the target language. One of the problems for this type of study is the risk of continually creating sub-styles of styles as different patterns of distribution are observed for marginally different (i.e. in terms of the extralinguistic description) types of speech or situation. In fact, it would be possible, with very fine-grained analysis, to continue to find uniqueness in every new sample. For the study of L2 we appear to be faced, then, with a problem: at what stage do we consider variant use of forms to be situation-specific (i.e. functionally motivated) and, conversely, at what point do we consider the distributions that we do find to be explicable by level or degree of acquisition? We cannot control situations of naturally occurring speech so that all samples are one hundred percent comparable and so there is bound to be intra-group variation. It seems, then, that for this type of study to become a plausible way forward, careful choice of data is paramount both for L2 data and for L1 control group data.

In some instances, no significant difference in percentage use of *parce que* sequences (governed or non-governed) was found between situations which are quite clearly different in terms of their extralinguistic description. However, it is not because we find a similar distribution for a given feature that the situations are the same (Biber 1988, for example, looks at the distribution of some 67 different linguistic features). Rather, it means that the frequency of *parce que* sequences is not a quantifiably variable feature in this case.

As Bartning & Schlyter (2004) have pointed out in their taxonomy, lower levels of learner (stages one and two in their model) do not have a rich palette of expression: general connectors (*et, mais, puis*, 'and, but, then'), for example, are found (but not *donc* and *enfin*, roughly translatable as 'so' and 'well') and 'simple subordination' using *parce que* and *quand*. Concerning our data, the findings for *parce que* + *comme* / *dès que* / etc. (see section 2.2.3 above) would appear to follow this: although the figures were very low, we nevertheless found

that grouping was used both more frequently and more variably in the fourth year group than in the first year group. One possible explanation of this could be the distribution of certain types of combinations in the target language (i.e. typically in oral data) and subsequently their absence from typical presentations of the language as an L2. For example, when we look at numbers of occurrences in three large L1 French corpora<sup>11</sup> for the three types of grouping used by the L2 speakers, comparing oral and written data, this type of grouping (without looking at the particular function of *parce que*) is much rarer in the written data than in the oral data (see table 7 in appendix). It is worth noting that those occurrences in the written data appeared all to relate to oral (or ‘oral-like’) situations: e.g. theatre dialogue (example 11), reported speech in the written media (example 12) and Internet communication (example 13).

- (11) précisément ! ... c’est pour cela qu’il faut rompre avec lui, peu à peu ... **parce que si** vous restez intimement liés, suivez-moi bien, comme il est pauvre, il vous ...  
(Theatre)  
‘precisely! ... that is why you must stop seeing him, little by little ... **because if** you remain intimate, hear me now, as he is poor, he will ...’
- (12) un commerce encore rare en Russie. “Je veille à m’habiller très chaudement **parce que si** je me recroqueville, cela dissuade les clients. Il faut les accueillir à ...”  
(Press, *La Tribune de Genève*)

---

<sup>11</sup> The oral component is made up of CRFP (450 000 words) and CORPAIX (1 000 000 words). The written component comes from the DELIC corpus (4 000 000 words).

‘a commercial activity still rare in Russia: “I make sure I wear warm clothes **because if** I’m doubled up with cold, that puts the customers off. They must be welcomed at ...”’

- (13) je suis contre le port du voile, en tout cas au moins quand il est forcé **parce que si** on est majeur et libre on fait ce qu’on veut, mais ça ne m’empêche pas d’...  
(Internet forum)

‘I’m against women wearing Islamic headscarves, at least when it’s forced upon them **because if** you’re grown-up and can act freely you can do what you want, but that doesn’t stop me from ...’

## Conclusion

Whilst the study of *parce que* sequences alone does not allow us to identify specific speech styles in any great detail, this type of approach would, nonetheless, appear to suggest that general patterns of variation, similar to those of the target language (though not necessarily in line with traditional prescriptive grammars of the language, where *parce que* is typically subordinating) are present in L2 and these are found in the speech of learners of different levels of proficiency. Contrary to the idea that less advanced learners are lesser users of variation, in particular pre-year-abroad (supported by the findings of many existing studies of variable features – e.g. see overview in Dewaele 2007), this simple exploration of English learners of French would suggest that first year students may vary just as much as fourth years in some respects (see also Tyne, 2009). Continued analysis of the distribution of forms (and not just alternation between features in an increasingly rich palette of variables) across speech situations within a general framework

will help give a more nuanced picture of the extent of mastery of variation in L2 at different stages of acquisition.

## References

- Adam, J.-M. 1992, *Les textes, types et prototypes: récit, description, argumentation, explication et dialogue*. Paris: Nathan
- Bartning, I. & Schlyter, S. 2004. Itinéraires acquisitionnels et stades de développement en français L2. *Journal of French Language Studies* 14(3): 281-299.
- Biber, D. 1988. *Variation Across Speech and Writing*. Cambridge: CUP.
- Biber, D. 2000. Corpus-Based Analysis of Grammar: Variability in the form and use of English Complement Clauses. In *Corpus: méthodologie et applications linguistiques*, M. Bilger (ed), 224-236. Paris & Perpignan: Honoré Champion & Presses Universitaires de Perpignan.
- Bilger, M. & Campione, E. 2002. Propositions pour un étiquetage en 'séquences fonctionnelles'. *Recherches sur le français parlé* 17: 117-136.
- Blanche-Benveniste, C., Bilger, M., Rouget, C. & Van den Eynde, K. 1990. *Le français parlé: études grammaticales*. Paris : Editions du CNRS.
- Blanche-Benveniste, C., Deulofeu, J., Stéfanini, J. & Van den Eynde, K. 1987. *Pronom et syntaxe: l'approche pronominale et son application au français* [2<sup>nd</sup> ed.]. Paris: SELAF.
- Branca-Rosoff, S. 1999. Des innovations et des fonctionnements de langue rapportés à des genres. *Langage et Société* 87: 115-129.
- Bronckart, J.-P. et al., 1985. *Le fonctionnement des discours. Un modèle psychologique et une méthode d'analyse*. Paris: Delachaux et Niestlé.

- Chafe, W. 1982. Integration and Involvement in Speaking, Writing, and Oral Literature. In *Spoken and Written Language: Exploring Orality and Literacy*, D. Tannen (ed.), 35-53. Norwood, New Jersey: Ablex.
- Crookes, G. 1990. The Utterance, and Other Basic Units for Second Language Discourse Analysis. *Applied Linguistics* 11(2): 183-99.
- Debaisieux, J.-M. 2002. Le fonctionnement de *parce que* en français parlé: étude quantitative sur corpus. In *Romanistische Korpuslinguistik – Korpora und gesprochene Sprache, Romance Corpus Linguistics, Corpora and Spoken Language*, C. Pusch, C. & W. Raible (eds), 349-362. Tübingen, Gunter Narr Verlag.
- Debaisieux, J.-M. 2004. Les conjonctions de subordination: mots grammaticaux ou mots de discours? Le cas de *parce que*. *Revue de sémantique et pragmatique* 15/16: 51-67.
- Dewaele, J.-M. 2007. Diachronic and/or Synchronic Variation? The Acquisition of Sociolinguistic Competence in L2 French. In *French Applied Linguistics*, D. Ayoun (ed), 208-236. Amsterdam: John Benjamins.
- Dewaele, J.-M. & Mougeon, R. (eds.). 2002. L'acquisition de la variation par les apprenants du français langue seconde. *AILE* 17.
- Ellis, R. 1985. Sources of Variability in Interlanguage. *Applied Linguistics* 6(2): 118-131.
- Foster, P. *et al.*, 2000. Measuring Spoken Language: A Unit for All Reasons. *Applied Linguistics* 21(3): 354-375.
- Gadet, F. & Tyne, H. (eds.). 2004. Le style comme perspective sur la dynamique des langues. *Langage et société* 109.
- Givón, T. 1979. From Discourse to Syntax: Grammar as a Processing Theory. In *Syntax and Semantics. Vol. 12: Discourse and Syntax*, T. Givón (ed), 81-112. New York: Academic Press.



- Hawkins, R. 2004. The Contribution of the Theory of Universal Grammar to our Understanding of the Acquisition of French as a Second Language. *Journal of French Language Studies* 14 (3): 233-255.
- Hawkins, R. & Towell, R. 2001. *French Grammar and Usage* [2<sup>nd</sup> ed.]. London: Arnold.
- Jahandarie, K. 1999. *Spoken and Written Discourse: A Multidisciplinary Perspective*. Norwood, New Jersey: Ablex.
- Kerbrat-Orecchioni, C. & Traverso, V. 2004. Types d'interactions et genres de l'oral. *Langages* 153: 41-51.
- Klein, W. & Perdue, C. 1992. *Utterance Structure: Developing Grammars Again*. Amsterdam: John Benjamins.
- Koch, P. & Österreicher, W. 2001. Gesprochene Sprache und geschriebene Sprache / Langage parlé et langage écrit. *Lexicon des Romanistischen Linguistik* 1(2), 584-627. Tübingen: Max Niemayer Verlag.
- Langage et Société* 87, 1999. Types, modes et genres.
- Malrieu, D. & Rastier, F. 2001. Genres et variations morphosyntaxiques. *TAL* 42(2): 547-577.
- Mougeon, R. Nadasdi, T. & Rehner, K. 2002. Etat de la recherche sur l'appropriation de la variation par les apprenants avancés du FL2 ou FLE. *AILE* 17: 7-50.
- Noyau, C., de Lorenzo, C., Kihlstedt, M., Paprocka, U., Sanz, G. & Schneider, R. 2005. Two dimensions of the representation of complex event structures: granularity and condensation. Towards a typology of textual production in L1 and L2. In Hendriks, H. (ed.) *The Structure of Learner Varieties*. Berlin: Mouton de Gruyter, pp. 157–201.
- Recherches sur le français parlé* 18, 2004. Autour du corpus de référence du français parlé.
- Skehan, P. 1996. Syntactic and Pragmatic Modes in Task-Based Foreign Language Learning. ESRC End of Award Report (R000221528).

- Tarone, E. 1982. Systematicity and Attention in Interlanguage. *Language Learning* 32(1): 69-84.
- Towell, R. 2004. Research into the Second Language Acquisition of French: Achievements and Challenges. *Journal of French Language Studies* 14(3): 357-75.
- Tyne, H. (2009). Style in L2: The Icing on the Cake? In *The Advanced Learner Variety: the Case of French*, E. Labeau & F. Myles (eds) 243-268. Bern: Peter Lang.
- Van den Eynde, K. 1998. From Verbal to Nominal Valency: Some Methodological Reflections. In *Studies in Valency IV: Valency and Verb Typology*, K. Van Durme & L. Schøsler (eds), 147-161. Odense: Odense University Press.

Van den Eynde, K. 1998. From verbal to nominal valency: Some methodological reflections. In *Studies in Valency IV: Valency and Verb Typology*, K. Van Durme & L. Schösler (eds), 147–161. Odense: Odense University Press.

## Appendix

Table 1. L1 *parce que* sequences

Sequence type	PUB		PRO		PRI	
	Tot <sup>13</sup>	Av per 10 000	Tot	Av per 10 000	Tot	Av per 10 000
<i>Parce que</i> governed	26	22.22	31	30.21	43	36.71
<i>Parce que</i> non-governed	26	25.27	32	30.68	30	26.47

Table 2. L2 *parce que* sequences

Sequence type		FP		FC		IC	
		Tot	Av per 10 000	Tot	Av per 10 000	Tot	Av per 10 000
all	<i>Parce que</i> governed	29	18.28	35	28.28	58	17.31
	<i>Parce que</i> non-governed	39	21.30	55	36.61	60	14.13
Y1	<i>Parce que</i> governed	14	21.56	12	36.44	20	18.52
	<i>Parce que</i> non-governed	14	18.85	14	38.77	7	6.58
Y4	<i>Parce que</i> governed	15	15.00	23	20.13	38	16.10
	<i>Parce que</i> non-governed	25	23.76	41	34.44	53	21.68

Table 3. L1 grouping

Sequence type	PUB		PRO		PRI	
	Tot	Av per 10 000	Tot	Av per 10 000	Tot	Av per 10 000
<i>Parce que</i> + <i>comme</i>	0	0.00	1	1.05	0	0.00
<i>Parce que</i> + <i>dès que</i>	1	1.00	0	0.00	0	0.00

13. These are raw figures. However, given that the amount of data contributed by each speaker to the dataset is variable, all comparisons were made on the basis that speakers contribute evenly to the data set, i.e. we calculated each individual's production in terms of the number of occurrences per 10 000 words. The new total was then used to produce an average number of occurrences per 10 000 words for each corpus (L1 and L2).

<i>Parce que + lorsque</i>	2	1.65	0	0.00	0	0.00
<i>Parce que + quand</i>	0	0.00	1	1.05	2	1.52
<i>Parce que + si</i>	1	0.65	2	1.84	2	1.52

Table 4. L2 grouping

Sequence type		FP		FC		IC	
		Tot	Av per 10 000	Tot	Av per 10 000	Tot	Av per 10 000
all	<i>Parce que + comme</i>	1	0.65	0	0.00	1	0.22
	<i>Parce que + dès que</i>	0	0.00	0	0.00	0	0.00
	<i>Parce que + lorsque</i>	0	0.00	0	0.00	0	0.00
	<i>Parce que + quand</i>	1	0.52	1	0.75	2	0.70
	<i>Parce que + si</i>	1	0.65	3	1.39	2	0.37
Y1	<i>Parce que + comme</i>	0	0.00	0	0.00	0	0.00
	<i>Parce que + dès que</i>	0	0.00	0	0.00	0	0.00
	<i>Parce que + lorsque</i>	0	0.00	0	0.00	0	0.00
	<i>Parce que + quand</i>	1	1.05	1	1.50	1	0.82
	<i>Parce que + si</i>	0	0.00	0	0.00	0	0.00
Y4	<i>Parce que + comme</i>	1	1.31	0	0.00	1	0.44
	<i>Parce que + dès que</i>	0	0.00	0	0.00	0	0.00
	<i>Parce que + lorsque</i>	0	0.00	0	0.00	0	0.00
	<i>Parce que + quand</i>	0	0.00	0	0.00	1	0.57
	<i>Parce que + si</i>	1	1.31	3	2.79	2	0.74

Table 5. Is there a significant difference in the distribution of *parce que* sequences between speech situations? (L1)

Sequence type	PUB-PRO		PUB-PRI		PRO-PRI	
	$\chi^2$ df = 1	Signif p <.05	$\chi^2$ df = 1	Signif p <.05	$\chi^2$ df = 1	Signif p <.05
<i>Parce que</i> governed	1.93	no	5.96	yes	1.13	no
<i>Parce que</i> non-governed	0.96	no	0.05	no	0.57	no

**Table 6.** Is there a significant difference in the use of *parce que* sequences between speech situations? (L2)

Sequence type		FP-FC		FP-IC		FC-IC	
		$\chi^2$ df = 1	Signif p <.05	$\chi^2$ df = 1	Signif p <.05	$\chi^2$ df = 1	Signif p <.05
all	<i>Parce que</i> governed	5.29	yes	0.06	no	6.43	yes
	<i>Parce que</i> non-governed	9.38	yes	2.67	no	21.33	yes
Y1	<i>Parce que</i> governed	8.03	yes	0.41	no	11.91	yes
	<i>Parce que</i> non-governed	19.47	yes	11.50	yes	55.03	yes
Y4	<i>Parce que</i> governed	2.22	no	0.11	no	1.35	no
	<i>Parce que</i> non-governed	3.86	yes	0.17	no	5.60	yes

**Table 7.** Grouping in oral / written corpora (number of occurrences per million words)

Sequence type	Oral	Written			
		Press	Forum	Sci	Lit
<i>Parce que</i> + <i>comme</i>	24	0	1	1	0
<i>Parce que</i> + <i>quand</i>	71	0	3	0	1
<i>Parce que</i> + <i>si</i>	63	3	2	1	3