

The influence of text formality on the syntactic position of French contrastive adverbs

A corpus analysis of *en revanche* and *par contre*

Jorina Brysbaert & Karen Lahousse

Research Foundation – Flanders (FWO) & KU Leuven

Abstract This paper is about the influence of text formality on the syntactic position of contrastive adverbs in French. We analyze the distribution of the two contrastive adverbs *en revanche* and *par contre*, both conveying a meaning similar to English ‘on the other hand’, in three written corpora representing different levels of formality. We show that (i) these adverbs occur more often in clause-initial position as the level of formality decreases, and that (ii) the formal character of *en revanche* and the informal character of *par contre* explain distributional differences between the two adverbs. We conclude that text formality has a significant influence on the position of these two contrastive adverbs, next to factors such as text type and lexical idiosyncrasy, which had been put forward in the linguistic literature. Our findings also corroborate the idea that French is largely characterized by formality-induced variation, even with respect to word order patterns, which have not received much attention in previous research on formality.

1. Introduction

In most languages, contrastive adverbs, such as *en revanche* and *par contre* (both similar to English ‘on the other hand’), can occupy different positions in the clause (1):

- (1) [EN REVANCHE / PAR CONTRE] Cette année-ci [EN REVANCHE / PAR CONTRE] ma sœur [EN REVANCHE / PAR CONTRE] ne veut [EN REVANCHE / PAR CONTRE] pas partir en voyage [EN REVANCHE / PAR CONTRE].
'[ON THE OTHER HAND] This year [ON THE OTHER HAND] my sister [ON THE OTHER HAND] does not want [ON THE OTHER HAND] to travel [ON THE OTHER HAND].'
(lit.)

Several factors have been argued to affect the position of these adverbs: (i) discourse and information structure, e.g. highlighting of a topical shift, new information, etc. (Altenberg 1998, 2006, Dupont 2015, 2019, and Lenker 2014); (ii) general language-specific syntactic properties, e.g. verb-raising versus non-verb-raising (Dupont 2015, 2019); (iii) lexical idiosyncrasy, i.e. placement preferences of individual adverbs (Altenberg 2006, and Dupont 2015, 2019); (iv) mode, i.e. written versus spoken language (Biber et al. 1999); and (v) text type, e.g. newspaper editorials versus research articles (Dupont 2019). As far as we know, the distribution of contrastive adverbs has not yet been studied from the point of view of text formality. Yet, this seems to be a potentially interesting factor, since various syntactic phenomena have been proven to be influenced by the level of formality (e.g. Hoffmann 2005 on preposition pied piping versus stranding in English, and Levin & Garrett 1990 on left- versus right-branching sentences in English). Moreover, the role of text formality seems to be particularly pertinent for French, because it has been claimed that this language is subject to important formality-related variation, which might lead to a situation of diglossia (Rowlett 2013, and Zribi-Hertz 2011). Some well-known examples of this kind of variation are the use of *on* 'one' versus *nous* 'we', the deletion versus retention of the negative particle *ne* 'not', and the form of interrogative sentences (see Etienne & Sax 2009, and Zribi-Hertz 2011 for an overview).

The goal of this paper is to determine the influence of text formality on the distribution of contrastive adverbs in French. To rule out the potential impact of the type of contrast these adverbs express (simple contrast vs. concession vs. correction), we focus on two adverbs that are semantically closely related, but different in terms of formality: *en revanche* has been said to be typical of formal French, whereas *par contre* has been claimed to be characteristic of informal French (Hamma & Haillet 2002).

In what follows we first give an overview of our corpora and extracted data (2.). We then introduce our classification of adverbial positions (3.), followed by a presentation of our results (4.). Next, we discuss the findings of our distributional analysis of *en revanche* and *par contre* (5.), and argue that the influence of

formality is at least as important as the effect of text type (5.1.) and lexical idiosyncratic placement preferences (5.2.).

2. Corpora and data

We analyzed data from three written French corpora: *Le Monde* (LM), *Est Républicain* (ER) and *Yahoo Contrastive Corpus of Questions and Answers* (YCC).¹ The *LM corpus* contains all articles that appeared in the newspaper *Le Monde* in 1998 and counts about 25,7 million words. Examples of *en revanche* and *par contre* were provided by Piet Mertens (KU Leuven).² The *Est Républicain corpus* is also a journalistic corpus, consisting of all articles that were published in the newspapers *L'Est Républicain* and *Vosges Matin* in the periods 1999-2003 and 2006-2011. We used the second version of this corpus, which has been compiled at ATILF, under the direction of Gaiffe et al. (2018). Given its big size, we only consulted the most recent years 2010 and 2011, containing about 74 million words in total. Occurrences of the two adverbs were extracted with the concordancer AntConc (Anthony 2018). The *YCC corpus* consists of language data from the online discussion platform *Yahoo! Answers* (<https://answers.yahoo.com/>), where questions can be posted and answered. It has been compiled by De Smet (2009) and contains all messages from the years 2006 to 2009, resulting in a total of about 6,1 million words. Again, examples of the adverbs were extracted using AntConc.

Importantly, the three corpora are representative of three different levels of formality, ranging from very formal (LM) to rather informal (YCC). The LM corpus is definitely the most formal one of the three. *Le Monde* is a well-established national newspaper, known for its rigorous editorial board. Moreover, this newspaper has been actively involved in the publication of several prescriptive language guides, showing that great importance is attached to the use of 'normative' French. On the continuum between 'language of immediacy' (informal) and 'language of distance' (formal) proposed by Koch & Oesterreicher (1985, 2007), LM should be situated towards the end of the 'language of distance' pole. Journalists of LM are expected to report on public events and topics, in a very thoughtful writing style, showing low to zero emotional involvement.

¹ It would of course also be interesting to test the influence of formality in spoken French. This is however challenging, due to the lack of big spoken French corpora of different formality levels.

² We thank Piet Mertens for the extraction of these examples.

With respect to its level of formality, ER can be situated in between LM and YCC. In contrast to *Le Monde*, both *L'Est Républicain* and *Vosges Matin* are regional newspapers, which probably have a less thorough editing process than *Le Monde*. This is confirmed by the relatively large number of (small) language errors that can be found in the ER corpus, whereas those are quasi absent in the LM corpus. On the continuum of Koch & Oesterreicher (1985, 2007), the ER corpus can be placed somewhat further away from the 'language of distance' pole – and hence closer to the 'language of immediacy' pole – than LM. The articles in ER often describe rather local events (e.g. birth of a child, wedding, sports activities, etc.) and journalists generally adopt a much more personal and involved writing style.

YCC is clearly an informal written corpus. Since anyone can publish messages on the *Yahoo* discussion forum, a lot of texts come from inexperienced writers. These texts are mainly written in a rather informal (chat language) style and do not go through an editing process. Hence, they are often full of spelling and grammatical errors. From the three corpora in our study, YCC is closest to the 'language of immediacy' pole discussed by Koch & Oesterreicher (1985, 2007). Writers on this online platform often engage with each other in a very emotional and direct way, and their posts deal with more private topics than those in the two newspapers.

Table 1 presents an overview of the absolute and relative frequencies of *en revanche* and *par contre* per corpus.

Table 1. Absolute and relative frequencies of *en revanche* and *par contre* per corpus

	LM (+ formal) (25.7 million words)	ER (± formal) (74 million words)	YCC (- formal) (6.1 million words)
<i>en revanche</i>	64/million (1636)	45/million (3353)	47/million (289)
<i>par contre</i>	2/million (46)	14/million (1025)	248/million (1515)

The adverb *par contre* is typical of informal French (YCC), whereas *en revanche* is most often used in very formal French (LM) (Brysbaert & Lahousse 2020). Note that *par contre* is extremely infrequent in the LM corpus. The results of the distributional analysis of this adverb in LM must therefore be interpreted with cautiousness.

3. Classification of adverbial positions

In the linguistic literature, there are mainly two approaches to the distribution of French contrastive adverbs. Some researchers make a distinction between clause-

initial (2), clause-medial (3) and clause-final (4) contrastive adverbs (e.g. Csúry 2001, and Hamma & Haillet 2002):

- (2) *Le vin blanc est bon ; **par contre**, le vin rouge pique un peu.* (Hamma & Haillet 2002:107)
 ‘The white wine is good; **on the other hand**, the red wine stings a little.’
- (3) *Il n’a pas eu son permis ; il a, **par contre**, rencontré l’amour de sa vie.* (Hamma & Haillet 2002:107)
 ‘He did not obtain his license; he did, **however**, find the love of his life.’
- (4) *J’ai acheté le bouquin ; il n’y a rien sur l’exclamation, **par contre**.* (Hamma & Haillet 2002:107)
 ‘I bought the book; there is nothing on exclamation, **though**.’

The main disadvantage of these analyses is the fact that the sentence-medial position does not distinguish between adverbs occurring immediately after the subject, after the inflected verb (as in (3)), after a complement of the verb, etc. (but see e.g. Altenberg 2006, who identifies several sentence-medial sub-positions in English and Swedish).

Other researchers, such as Dupont (2015, 2019), use a classification based on thematic-structural criteria. She starts from a theme-rheme analysis of the clause and distinguishes five adverbial positions, illustrated in example (5): thematic 1 (T1) (i.e. adverbs occurring at the very beginning of the clause), thematic 2 (T2) (i.e. adverbs occurring after the interpersonal theme³), rhematic 1 (R1) (i.e. adverbs occurring after the topical theme⁴), rhematic 2 (R2) (i.e. adverbs occurring within the verb phrase) and rhematic 3 (R3) (i.e. adverbs occurring after all obligatory complements in the verb phrase):

- (5) [**However**_(T1)] interestingly [**however**_(T2)] my sister [**however**_(R1)] liked [**however**_(R2)] to feed the birds [**however**_(R3)]. (our example)

This thematic-structural classification works well in formal texts, where clauses are generally well-structured and have a clear start (capital letter) and ending

³ An interpersonal theme provides “some information about the stance of the writer or speaker towards the message” (Dupont 2015:92).

⁴ According to Dupont (2015:91), who refers to Hasselgård (2004:65-66), the topical theme is “the first clause element which has a function in transitivity” and indicates an entity “of the outside world such as participants, and circumstances”.

(punctuation mark). However, in more informal corpora, such as the YCC (see section 2.), this is often not the case, which makes the classification difficult to apply.

We use a classification of adverbial positions based on syntactic criteria. We focus on clauses presenting an SVO word order and determine the position of the adverb with respect to the subject (S) (full underline in the following examples) and the inflected verb (V) (dotted underline).⁵ Three possible positions are distinguished, i.e. before S ((6)-(7)), between S and V (8), and after V ((9)-(10)):

- (6) *En revanche, la viande bovine laisse apparaître une grande diversité, avec des morceaux de choix et de qualité.* (ER)
'On the other hand, beef shows great diversity, with delicate and qualitative pieces.' (lit.)
- (7) *À la campagne en revanche, l'enfant peut être exposé à de nombreux allergènes comme les moisissures, les pollens, ou encore les poils de certains animaux.* (ER)
'In the countryside on the other hand, children can be exposed to many allergens such as molds, pollens, or the hair of certain animals.' (lit.)
- (8) *La production de zinc, en revanche, avait été temporairement arrêtée en raison de la faiblesse actuelle des cours.* (LM)
'The zinc production, on the other hand, had been temporarily halted due to the current weakness in prices.' (lit.)
- (9) *Cette interview contient, en revanche, de précieuses informations sur l'évolution de la situation du cinéma en Algérie depuis l'indépendance.* (LM)
'This interview contains, on the other hand, valuable information on the evolution of the cinema situation in Algeria since independence.' (lit.)
- (10) *Sur cette même distance du 200m, le Villarois Nacim Fahchouch pouvait sourire, en revanche.* (ER)
'On this same distance of the 200m, the Villarois Nacim Fahchouch could smile, on the other hand.' (lit.)

⁵ About 9% of the extracted data had to be excluded, because there was no subject and/or inflected verb, or because the clause did not present SVO word order.

Note that we do not set apart adverbs occurring immediately after the inflected verb (9) from those appearing at the very end of the clause (10). This distinction seems to be less important, since French contrastive adverbs are almost never used in clause-final position (1-3% in the corpora of Dupont (2019)).

4. Results

The results of our positional analysis of the two contrastive adverbs are summarized in Table 2 (*en revanche*) and Table 3 (*par contre*), which provide an overview of the frequency of each position in the three corpora (i.e. three levels of formality). Raw numbers are given between brackets.

Table 2. Distribution of *en revanche* per corpus

	LM (+ formal)	ER (± formal)	YCC (- formal)
Before S	21% (314)	63% (1952)	87% (230)
Between S & V	16% (248)	6% (171)	4% (10)
After V	63% (969)	31% (957)	9% (24)
TOTAL	100% (1531)	100% (3080)	100% (264)

Table 3. Distribution of *par contre* per corpus

	LM (+ formal)	ER (± formal)	YCC (- formal)
Before S	14% (6)	83% (791)	92% (1182)
Between S & V	19% (8)	4% (37)	2% (23)
After V	67% (28)	13% (126)	6% (78)
TOTAL	100% (42)	100% (954)	100% (1283)

It appears that, for both adverbs, the distribution varies depending on the corpus. In LM (+ formal), they show up most often after V (63% - 67%), and almost as frequently between S & V (16% - 19%) and before S (21% - 14%). By contrast, in ER (± formal), their most frequent position is before S (63% - 83%), and they occur much less often after V (31% - 13%). In YCC (- formal), the two contrastive adverbs are almost exclusively used before S (87% - 92%). These differences are highly significant for both adverbs: *en revanche*, $X^2(4, N = 4875) = 909.16, p < .001$, with a moderate effect size, Cramer's $V = 0.31$; and *par contre*, $X^2(4, N = 2279) = 237.69, p < .001$, with a small effect size, Cramer's $V = 0.23$. In other words, as the level of formality decreases, the contrastive adverbs occur more often before S and less often between S & V and after V. The position before S

thus seems to be more typical of rather informal French, whereas the positions between S & V and after V are more characteristic of formal French.

As becomes clear from a comparison of the distribution of *en revanche* (Table 2) and *par contre* (Table 3), our analysis also reveals a considerable difference between the two adverbs. In the ER corpus, *en revanche* (63%) occurs much less often before S than *par contre* (83%), and conversely, *par contre* (13%) is used much less often after V than *en revanche* (31%). This difference is also statistically significant, $X^2(1, N = 3826) = 126.09, p < .001$, with a small effect size, Cramer's $V = 0.18$.

5. Discussion

Our data thus reveal that the level of formality has a significant effect on the position of the contrastive adverbs. In this section, we discuss the results in more detail, showing that the influence of text formality on adverb placement is at least as important as that of two other factors mentioned in the linguistic literature (see section 1.): text type (5.1.) and lexical idiosyncrasy (5.2.).

5.1. Influence of text formality vs. text type

Dupont (2019) shows that French contrastive adverbs occupy distinct preferred positions in academic prose versus newspaper editorials: in newspaper editorials, adverbs occur most frequently within the verb phrase (48%), and less often clause-initially (32%), whereas in research articles, they appear most often in clause-initial position (53%), and less often within the verb phrase (37%)⁶. She argues that this discrepancy is due to the text type (defined in terms of text purpose), and should be explained as a different use of the information structure signaling properties of the adverbs. When an adverb is placed after the inflected verb, it can highlight the boundary between given and new information, and mark off the rheme containing the new information. According to Dupont (2019:319-324), writers of editorials might take more advantage of this 'discursive potential' of

⁶ Dupont's (2019:379-380) corpus of editorials contains texts from four 'quality' newspapers (*Libération*, *Le Figaro*, *Le Monde* and *Le Nouvel Observateur*), whereas her corpus of academic prose consists of research articles from several disciplines within the Humanities. Also recall that she uses a thematic-structural classification of positions, which is different from our syntactic classification (see section 3.).

adverb placement than writers of research articles, in order to direct and influence the readers' perception and interpretation of their texts.

Based on our results, we propose however that this distinct distribution can also be partially linked to differences in formality. Our analysis shows that, within the text type of journalistic French (i.e. corpora LM and ER), contrastive adverbs have a distinct distribution depending on the level of formality. It is more than likely that such differences with respect to formality also play a part in the corpus of Dupont (2019), next to the influence of text type (i.e. text purpose) (see Hundt & Mair 1999, Hyland & Jiang 2017, and Westin 2002 on the level of formality of newspaper language and/or academic prose in English). Hence, we argue that, in written French, there might be a combined effect of text formality and text type on the position of contrastive adverbs.

5.2. *Influence of text formality vs. lexical idiosyncratic placement preferences*

In the linguistic literature, it has been shown that contrastive adverbs display partly idiosyncratic placement patterns, which means that they can have their own preferred positions (e.g. Altenberg 2006 on English and Swedish). For French, Dupont (2019) observes for instance that *en revanche* occurs more often in clause-initial position than *pourtant* (similar to English 'yet'), which in turn is more often used clause-initially than *au contraire* (similar to English 'on the contrary'). However, she does not take into account the fact that, from a semantic point of view, the adverbs in her study are slightly distinct: *en revanche* typically indicates a simple contrast, whereas *pourtant* signals a concession and *au contraire* may convey a correction or a simple contrast (Lewis 2006). According to us, this semantic difference could also intervene in the distribution of the adverbs. This is much less the case with respect to *en revanche* and *par contre*, which both express a simple contrastive relation and are more or less synonymous.

At first sight, our results seem to confirm that contrastive adverb placement is characterized by a certain degree of lexical idiosyncrasy, since the distribution of *en revanche* and *par contre* in the ER corpus is significantly different. However, this discrepancy between *en revanche* and *par contre* can also be explained in terms of formality. As we showed in previous research (Brysbart & Lahousse 2020), *en revanche* can be considered to be a 'formal adverb', in the sense that it is especially frequent in formal French, whereas *par contre* can be seen as an 'informal adverb', because it occurs much more often in informal French. Hence,

in ER, which is the ‘in-between’ corpus with respect to formality, the ‘formal’ *en revanche* shows up more often in the ‘formal’ after V position, and the ‘informal’ *par contre* shows a stronger preference for the ‘informal’ before S position. This implies that, with respect to the distribution of semantically closely related adverbs such as *en revanche* and *par contre*, formality might be of greater importance than lexical idiosyncrasy.

6. Conclusion

In this paper, we analyzed the syntactic position of the French contrastive adverbs *en revanche* and *par contre* in three written corpora, representing different levels of formality: LM (+ formal), ER (\pm formal) and YCC (- formal). Our analysis revealed that the distribution of the two adverbs is influenced by the level of formality. More precisely, the more informal the corpus, the more often the adverbs occur before S. In addition, we observed that in ER (\pm formal), *en revanche* is used much less often before S than *par contre*. We argue that this difference can also be explained in terms of formality, since *par contre* is especially frequent in informal French and might therefore display a stronger preference for the more ‘informal’ before S position than *en revanche*, which is characteristic of formal French.

On the basis of these findings, we argue that text formality can affect the position of contrastive adverbs independently of other factors that have been identified in previous research. More in general, our results support the hypothesis that French is largely characterized by formality-related variation, which had not yet been shown with respect to word order patterns, such as the distribution of contrastive adverbs. In future work, it would of course be interesting to (i) include more contrastive adverbs, and to (ii) build a statistical multifactorial model to study the interaction of the level of formality with other factors that could influence the adverbial position (e.g. mode, syntactic complexity), in order to gain a clear insight into the importance of the different factors at play.

References

- Altenberg, B. (1998). ‘Connectors and sentence openings in English and Swedish’. In S. Johansson and S. Oksefjell, eds., *Corpora and cross-linguistic research: Theory, method, and case studies*, Amsterdam: Rodopi, 115-143.

- Altenberg, B. (2006). 'The function of adverbial connectors in second initial position in English and Swedish'. In K. Aijmer and A.-M. Simon-Vandenberg, eds., *Pragmatic markers in contrast*, Oxford: Elsevier, 11-37.
- Anthony, L. (2018). *AntConc (Version 3.5.7)* [computer software]. Tokyo: Waseda University. Available from <https://www.laurenceanthony.net/software>.
- Biber, D., Johansson, S., Leech, G., Conrad, S. and Finnegan, E. (1999). *Longman grammar of spoken and written English*. London: Longman.
- Brysbart, J. and Lahousse, K. (2020). 'Les marqueurs de contraste au contraire, par contre et en revanche en français parlé et écrit, formel et informel'. In F. Neveu, B. Harmegnies, L. Hriba, S. Prévost and A. Steuckardt, eds., SHS Web of Conferences 78 (7e Congrès Mondial de Linguistique Française), Montpellier: EDP Sciences, paper n°01010. <https://doi.org/10.1051/shsconf/20207801010>.
- Csúry, I. (2001). *Le champ lexical de mais: Étude lexico-grammaticale des termes d'opposition du français contemporain dans un cadre textologique*. Debrecen: Kossuth Egyetemi Kiadó.
- De Smet, H. (2009). *Yahoo contrastive corpus of questions and answers* [corpus]. Leuven: Department of Linguistics, KU Leuven.
- Dupont, M. (2015). 'Word order in English and French: The position of English and French adverbial connectors of contrast'. *English Text Construction* 8(1), 88-124.
- Dupont, M. (2019). *Conjunctive markers of contrast in English and French: From syntax to lexis and discourse*. Ph.D. dissertation, Université catholique de Louvain, Louvain-la-Neuve.
- Etienne, C. and Sax, K. (2009). 'Stylistic variation in French: Bridging the gap between research and textbooks'. *The Modern Language Journal* 93(4), 584-606.
- Gaiffe, B., Nehbi, K. and Tonnelier, M. (2018). *Corpus journalistique issu de l'Est Républicain* [corpus] (2 ed.). Nancy: Analyse et traitement informatique de la langue française [ATILF]. Available from ORTOLANG, https://hdl.handle.net/11403/est_republicain/v2.
- Hamma, B. and Haillet, P.P. (2002). 'Par contre: Un type particulier de dynamique discursive'. *Linx* 46, 103-113.
- Hoffmann, T. (2005). 'Variable vs. categorical effects: Preposition pied piping and stranding in British English relative clauses'. *Journal of English Linguistics* 33(3), 257-297.
- Hundt, M. and Mair, C. (1999). 'Agile and uptight genres: The corpus-based approach to language change in progress'. *International Journal of Corpus Linguistics* 4(2), 221-242.
- Hyland, K. and Jiang, F. (2017). 'Is academic writing becoming more informal?'. *English for Specific Purposes* 45, 40-51.
- Koch, P. and Oesterreicher, W. (1985). 'Sprache der Nähe – Sprache der Distanz: Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte'. *Romanistisches Jahrbuch* 36, 15-43.
- Koch, P. and Oesterreicher, W. (2007). 'Schriftlichkeit und kommunikative Distanz'. *Zeitschrift für Germanistische Linguistik* 35(3), 346-375.
- Lenker, U. (2014). 'Knitting and splitting information: Medial placement of linking adverbials in the history of English'. In S.E. Pfenninger, O. Timofeeva, A.-C. Gardner,

- A. Honkapohja, M. Hundt and D. Schreier, eds., *Contact, variation, and change in the history of English*, Amsterdam: John Benjamins, 11-38.
- Levin, H. and Garrett, P. (1990). 'Sentence structure and formality'. *Language in Society* 19(4), 511-520.
- Lewis, D.M. (2006). 'Contrastive analysis of adversative relational markers, using comparable corpora'. In K. Aijmer and A.-M. Simon-Vandenberg, eds., *Pragmatic markers in contrast*, Oxford: Elsevier, 139-153.
- Rowlett, P. (2013). 'Do French speakers really have two grammars?'. *Journal of French Language Studies* 23(1), 37-57.
- SA Le Monde. (1999). *Le Monde sur CD-ROM: CEDROM-SNI* [corpus].
- Westin, I. (2002). *Language change in English newspaper editorials*. Amsterdam/New York: Rodopi.
- Zribi-Hertz, A. (2011). 'Pour un modèle diglossique de description du français: Quelques implications théoriques, didactiques et méthodologiques'. *Journal of French Language Studies* 21(2), 231-256.