

Topic shifts in contrast: Ways to change the subject in French and German

Benjamin Fagard, Peter Blumenthal

▶ **To cite this version:**

Benjamin Fagard, Peter Blumenthal. Topic shifts in contrast: Ways to change the subject in French and German. Journal of Pragmatics, In press, 10.1016/j.pragma.2019.07.016 . halshs-02376118

HAL Id: halshs-02376118
https://shs.hal.science/halshs-02376118

Submitted on 22 Nov 2019

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.

Benjamin Fagard & Peter Blumenthal. 2019. Topic shifts in contrast: Ways to change the subject in French and German, *Journal of Pragmatics*. https://doi.org/10.1016/j.pragma.2019.07.016

Author version

0. Introduction

The starting point of this article is a well-known fact: each sentence, each segment of any globally informative text brings along new elements, simultaneously picking up on semantic elements which are already present, having appeared previously. Only by combining new and old can speakers/writers create texts in which topics move on without resulting in a redundant (old-only) or incoherent (new-only) text. Such topic shifts may be left implicit. Alternatively, they may be signaled explicitly, for instance with the help of (more or less) specialized discourse markers (henceforth DMs), such as "topic orientation markers" (Fraser 2009) or "digressive discourse markers" (Traugott 2019), which we will call here "topic shifters".

Our question in this paper is how topic shifts are managed in French and German, or rather to what extent they are signaled by topic shifters. More specifically, the aim of the study is to (a) identify typical topic shifters in the two languages, as well as the syntactic patterns they follow, (b) identify other possible ways of signaling topic shifts, (c) find commonalities and differences between the two languages, given known typological differences between the two – and, if possible, to illustrate the tendency for French (and Romance languages in general) to signal discourse relations more explicitly than German (and Germanic languages). In order to do this, we used English as a pivot language, focusing on three DMs – a typical topic shifter, *by the way* (Traugott, ibid.), and two less typical topic shifters, *furthermore* and *besides* – and analyzed their translation equivalents[1] in French and German, focusing on the ones most frequently found in our corpus: German *zudem* 'besides, moreover', *ferner* 'furthermore', *außerdem* 'besides, moreover', *darüber hinaus* 'moreover', *übrigens* 'by the way, incidentally', *im Übrigen* 'incidentally, by the way' and French *d'ailleurs* 'moreover, besides', *en passant* 'in passing, by the way', *à (ce) propos* 'by

[1] We use the phrase *translation equivalent* to indicate that a given sequence is found to have (approximately) the same function of a target sequence in a translation corpus, indepently from the direction of the translation and the source language. For instance, if *by the way* in English is translated as *im Übrigen* in German and *par ailleurs* in French, then *im Übrigen* and *par ailleurs* are translation equivalents. Thus, it does not mean that one was translated into the other. This proved to be a useful concept for us, because the Europarl7 database contains texts translated in various directions, and the direction of the translation is not always known or easy to recover (i.e. what the source language is, and whether a pivot language was used, e.g. if a speech in Albanian is translated first into English, then from English into French).

the way', *en outre* 'moreover, besides', *par ailleurs* 'furthermore, besides', *de plus* 'moreover, furthermore', *du reste* 'moreover, besides'.[2]

The paper is organized as follows: in Section 1, we provide the state of the art, concerning both topic shifts and expected differences between French and German. In Section 2, we describe the methodology used in the main corpus study. In Section 3, we present the results of this corpus study. In Section 4, we present the results of a series of additional corpus studies. We conclude in Section 5.

1. State of the art: Signaling topic shifts

It is well-known that DMs are highly multifunctional (see for instance Métrich & Faucher 2009), but there is of course some specialization, as noted in the existing classifications such as Fraser's (2009) typology. While DMs (or 'pragmatic markers'[3]) have been the focus of many studies, few have addressed the issue of *digression* or *topic shift*, apart from studies on topic orientation markers (Fraser 2009), on additive markers (Forker 2016, König 2017), digressive markers (Pons Bordería and Estelles Arguedas 2009, Traugott 2019, forthcoming, Charolles, submitted) and topic shifting in general (Van Kuppevelt 1995). We are interested here in topic shifters, such as *(oh) by the way* in example (1).

(1) De Hagen tried to smile. He was not going to get into a discussion of how his parents had developed this particular programming device … It was so typical of his parents to develop a new technology to compensate for their failings as human beings, and then to make a fortune out of it. **Oh, and by the way**, Richie boy, sorry about the side-effects. (Young, 2004, *Ultraviolet night* [COHA]; example (2) in Traugott, submitted)[4]

Topic shifters may be defined as follows: if we take a topic as being "what a proposition is about and what is at the center of current interest and about which then further information is provided" (Forker 2016:74, after Lambrecht 1994, Krifka 2007), a TOPIC SHIFTER[5] is a DM which explicitly signals a change of topic, i.e. it highlights the fact that the "center of current interest" has shifted. In example (1), the topic shift is obvious, from De Hagen's parents to the side-effects of the new technology.

The few studies on topic shifting have shown that topic shifts can be signaled by a variety of means (Cortés Rodríguez and Hidalgo Downing 2015), including vocatives, aboutness markers (Prévost 2011), additives (König 1991, Forker 2016), zero marking (see Suzuki 1995,

[2] These glosses are tentative, given the polyfunctional nature of DMs. Due to space limitations, we do not present a systematic analysis of each one of these DMs.

[3] E.g. *pragmatic markers* (Fraser 1996, Brinton 1996), *discourse particles* (Schourup 1983, Hansen 1998, Aijmer 2002, Fischer 2006) or *discourse markers* (Schiffrin 1987, Schourup 1999).

[4] Traugott further specifies the context, indicating that "a doctor has commented on the uniqueness of Richie De Hagen's programming implant. This implant was the invention of De Hagen's parents."

[5] Semantic and functional categories, such as TOPIC SHIFTER, ADDITIVE or DIGRESSIVE, are, from this point on, written in small caps.

Zufferey 2016 on factors at play in explicit vs zero marking). Among these strategies, a typical way of signaling topic shifts is the use of additives (Author 1, submitted): in that case, *additives* are used pragmatically to add a new topic, typically when in clause-initial position. Another typical way of signaling topic shifts is the use of digressives, i.e. topic shifters which introduce a temporary topic shift, albeit the *temporary* status of such topic shifts is debated. Indeed, the line between digression and topic shift is hard to draw: digressions are expected to be closed quickly (Charolles, submitted), but topic shifts can be temporary or definitive (Van Kuppevelt 1995: 821, Pons Bordería and Estelles Arguedas 2009). For instance, topic orientation markers (e.g. *incidentally*, *before I forget*, *to change the topic*) can signal both topic shifting and digression (Fraser 2009); Traugott (2019) shows the same for what she calls *digressive discourse markers* (see also Mittwoch et al. 2002: 779). Finally, syntax seems to play a role, at least in English, with clause-initial uses triggering topic shift readings (Fraser 2009, Traugott 2019). This importance of the sentence periphery is not limited to English, however, since it has been established for French (see for instance Lamiroy and Charolles 2004, Author 2) and for German (Eisenberg 1998: 380, 652). It is also true for other languages; for instance, "initial positions are the preferred place for discourse marking in Spanish", as well (Pons Bordería 2018:373). But the initial position is not the only salient one for DMs. As adverbials with extraclausal uses, i.e. affecting discourse organization and management (Dik 1997), DMs may also have parenthetical uses. There are thus at least two salient syntactic patterns for DMs, sentence-initial (more integrated) and parenthetical (less integrated). These can be seen as points on a cline from conversation management functions (DMs then functioning more or less as independent utterances) to connecting functions (DMs then being more integrated into their host sentence), as claimed by Fischer (2006).

Beyond these considerations on topic shifting, one side issue in this paper is the extent of the differences between French and German. On the one hand, there is a very strong cultural proximity between European languages, not without consequences on linguistic structures, as shown in Haspelmath (2001: 1506-1507; see also Horst 2013: 182) – to the point that it is possible to speak of 'Standard Average European' (Whorf 1956 [1939], Van der Auwera 2011: 291): we can thus expect a similar behavior in the two languages. On the other hand, it has been shown that there is a tendency for French speakers to make discourse relations explicit, more frequently than in some other languages (Lamiroy and Vanderbauwhede 2016, 2019; Lewis 2006; Iordankaja and Mel'čuk 1999). There are thus also reasons to expect important differences in the signaling of topic shift.

2. Methodology – a contrastive corpus study of topic shifters

We adopt a contrastive approach, believing it can help us tease out different uses of the main DMs which play a key role in the signaling of topic shifts. As highlighted in Lamiroy and Vanderbauwhede (2019), "contrastive studies are valuable not only from a typological or comparative point of view, but [...] also contribute substantially to a better understanding of the individual languages, as they function as a discovery procedure of certain facts which

may go unnoticed when languages are observed in isolation (Aijmer & Altenberg 1996, Lewis 2006, Johansson 2007, Vanderbauwhede 2012, Lamiroy 2014)".

For the main corpus study reported on in this paper, we used English as a *tertium comparationis*, in order to limit the effects of the translation bias. This is an important point, given that DMs are notoriously difficult to translate (Aijmer 2008: 95). We started by choosing English DMs which can be used to signal topic shifts (2.1), then selected all occurrences of these DMs in our dataset (2.2), identified French and German equivalents (2.3), and finally analyzed the use types in all three languages (2.4).

2.1. Choice of DMs

Our first step was to select English DMs which are (more or less) typically used to signal topic shifts. We selected one DM for which the topic-shifting function seems central, *by the way* (Traugott 2019, to appear), as well as two other DMs which can also be involved in topic-shifting (Author 3): *furthermore*, which is also used as an additive marker (Forker 2016, König 2017), and *besides*, which is also used as an argumentative marker (Charolles, submitted). We thus expected the contrasts between these DMs to yield interesting clues.
Accordingly, we looked for occurrences of these DMs and their translation equivalents in French and German in Europarl7, a parallel corpus of debates in the European Parliament.

2.2. Corpus

The methodology adopted for study is the following: with the help of the SearchEngine interface, we accessed the Europarl corpus (Koehn 2005), a parallel corpus with data in all official European languages, gathered from the European Parliament Proceedings. We used version 7 (2015), which includes texts in 21 European languages. In this version of Europarl, texts are not always directly translated from the source language (Cartoni et al. 2013: 36), but sometimes through a pivot language, generally English – which explains our choice of English as a pivot language in this study.

We extracted tokens of the selected English DMs, with the surrounding context, and the corresponding context in French and German. This means that we extract English DMs, which can have been originally said or written in English, or only translated from another language; the database then provides the context[6] of this DM in English, as well as the corresponding contexts in French and German – which can likewise be originals, or translations (from any other language of the European Parliament). We then eliminated duplicated utterances and incomplete excerpts, as well as those in which the sequences were found to have other uses (e.g. *by the way* can be found in sequences in which it is not used as a DM, such as *I was amazed by the way they did it*). The resulting dataset contains 8,217 occurrences of the three DMs and their equivalents in the two other languages (i.e. a total of 24,651 tokens).

[6] The context is generally limited to one sentence, which means that, in order to be sure of our interpretation, we had to retrieve the rest either from the corpus or directly from the European Parliament website. In any case, the context remains limited, since the speaking time is generally limited to one or two minutes.

DM	Europarl7
furthermore	7,353
besides	350
by the way	514
Total	**8,217**

Table 1: Main dataset (*furthermore*, *by the way* and *besides* in Europarl7)

2.3. Coding and annotation

We analyzed the dataset, identifying the translation equivalents of all three English DMs in the French and German excerpts. This lead to the exclusion of some occurrences, as noted above, whenever context was insufficient or whenever the sequences had non-DM uses. However, the number of tokens excluded from the initial dataset was quite limited. For instance, in the Europarl7 database, we found a total of 8,440 occurrences of the three English DMs; our dataset contains 8,217 occurrences from that total (a difference of 2.6%).

In most cases, the translation equivalents of the three English DMs were adverbials such as French *par ailleurs* 'by the way, besides' or German *im Übrigen* 'incidentally, by the way'. Sometimes, however, we found no equivalent. Thus, in example (2), the German DM *im Übrigen* is translated into English (with the DM *by the way*), but the French translation contains nothing marking the sentence as having a topic shift: *Ce point de vue est soutenu par la Finlande* simply translates as 'This also enjoys the support of Finland'.

(2)

a. We therefore need to stem the flood of rubbish and demand a longer life for products. That is why we have introduced an amendment providing for a 5-year guarantee, so as to finally achieve this breakthrough to sustainability. This also enjoys the support of Finland, **by the way**, which makes it clear that even in the Council of Ministers there is support for sustainability in production. We want to see that both the average life and the rate of consumption of products is disclosed with the manufacture and use of the product.

b. (French) Nous devons endiguer la marée des déchets et promouvoir la longévité des produits. C'est pour cette raison que nous avons déposé une proposition d'amendement prévoyant une période de garantie de cinq ans, de manière à réaliser cette percée pour que l'on en arrive à imposer la longévité des produits. Ce point de vue est soutenu par la Finlande. Il apparaît ainsi clairement que même au sein du Conseil de ministres, l'option d'une plus grande longévité des produits disposera d'un appui.

c. (German; source language) Wir müssen also die Müllfluten eindämmen und die Langlebigkeit der Produkte fördern. Deshalb haben wir einen Änderungsantrag eingebracht, der eine Garantiezeit von fünf Jahren vorsieht, um endlich diesen Durchbruch zur Nachhaltigkeit zu erreichen. Das wird **im übrigen** auch von Finnland unterstützt. Damit wird auch deutlich, daß es sogar im Ministerrat eine Unterstützung hin zu mehr Nachhaltigkeit in der Produktion gibt.

In other cases, there is a translation equivalent, but it takes the form of a more complex sequence, for instance a semi-lexicalized construction as in (3) (source language not indicated): *à cela s'ajoute le fait que* (lit. 'to this REFL adds the fact that') in French, *dazu kommt, daß* (lit. 'there.to comes, that') in German.

(3)

a. **Furthermore**, with the decline in swine fever, products from the Netherlands are fully back on the market again.

b. **À cela s'ajoute le fait que** la production néerlandaise est totalement revenue sur le marché après le recul de la peste porcine.

c. **Dazu kommt, daß** auch die niederländische Erzeugung nach dem Abklingen der Schweinepest wieder voll auf den Markt gekommen ist.

Sometimes, a whole clause plays the part of the English DM, as in (4). In this example, a simple English DM (*furthermore*) is used to translate a whole sentence in the Spanish original – *añado algo más* (lit. 'add.1SG something more'). The French translation sticks to the original, with *j'ajoute encore quelque chose* (lit. 'I add still some thing'), while the German translation is somewhere in between, with a non-verbal sentence: *und noch etwas* (lit. 'and still something').

(4)

a. (Spanish; source language) **Añado algo más.** La Asamblea Parlamentaria de la OTAN, que no se caracteriza por su filiación socialista, en su Comité Permanente del día 6 de abril ha planteado una resolución que creo acorde con lo que podemos plantear nosotros.

b. **Furthermore**, in its Standing Committee of 6 April, the Parliamentary Assembly of NATO, which is not known for its Socialist tendencies, tabled a resolution which I believe is in accordance with what we ourselves would propose.

c. (French) **J'ajoute encore une chose.** L'assemblée parlementaire de l'OTAN, qui n'est pas caractérisée par sa filiation socialiste, a proposé au sein de son comité permanent du 6 avril une résolution qui à mon avis correspond à ce que nous pouvons établir.

d. (German) **Und noch etwas.** Die parlamentarische Versammlung der NATO, die nicht für ihre sozialistische Einstellung bekannt ist, hat am 6. April in ihrem Ständigen Ausschuss einen Entschließungsantrag gestellt, der meines Erachtens insoweit mit dem, was wir fordern können, übereinstimmt.

Examples such as (2-4) were not excluded from the dataset, since the English adverbial or sequence is used as a DM in all three. After identifying the translation equivalents, we coded the type of syntactic configuration, distinguishing at the most four different configurations: INITIAL, PARENTHETICAL, FINAL and IN PARENTHETICAL CLAUSES, coding as OTHER all other cases. We coded both the English DMs and their translation equivalents in French and German. The rationale behind this is that, as noted in Section 1, initial uses are usually associated to discourse management. We thus expected topic-shifting uses to be linked to initial uses. We coded as such all cases in which the DM directly followed a strong pause (colon, semi-colon, etc., as in (5a-c)).[7]

(5)

[7] This means that we did not distinguish between extraclausal (i) and initial (ii) uses :

(i) Ich weiß nicht, wer Sie sind. **Im Übrigens**, es interessiert mich auch nicht besonders.
"I do not know who you are. In fact, I am not really interested" (our translation)

(ii) Wir müssen gemeinsam über eine Lösung nachdenken. **Im Übrigens** bin ich deshalb nicht gekommen (...).
"We should look together for a solution. Actually, this is not the reason I came". (our translation)

Both examples are taken from Métrich & Faucher (2009: 879), who specify that the initial, non extraclausal use (i.e. in the *Vorfeld*, without inversion of the standard subject-verb order) of *im Übrigens* is by far the most frequent.

a. **By the way**, no judicial verdict has been reached in any of these cases yet. Mari language schools are being closed down, and education in the Mari language is only allowed in the primary levels of elementary education. (source language: Hungarian)

b. (French) **Cela dit en passant**, aucun verdict judiciaire n'a été rendu jusqu'à présent dans aucun des cas.

c. (German) **Übrigens** ist bislang in keinem dieser Fälle ein Urteil gesprochen worden.

We also coded as initial cases in which a DM was preceded by a coordinating conjunction (English *and*, French *et*, German *und*) (6b).

(6)

a. As these subsidies pile up, do you not think that taxpayers will eventually have had their fill of this recklessness? Fifthly, the list of eligible associations contained in the report reveals a motley assortment of organisations, many of which have nothing to do with European citizenship, such as the European Council on Refugees and Exiles. In addition, since it is all so simple, the European Parliament could not resist adding to this list through its amendments. I also wonder, **by the way**, about the Association of the Councils of State and Supreme Administrative Jurisdictions of the European Union. How can it be that an association of leading judges who are supposed to be part of an entirely independent judiciary can receive money to promote active European citizenship? My sixth and final point (...)

b. (French; source language) **Et au passage,** je m'interroge en plus sur l'Association des conseils d'État et des juridictions administratives suprêmes de l'Union européenne.

c. (German) Fragen stelle ich mir auch zur Vereinigung der Staatsräte und der Obersten Verwaltungsgerichte der Europäischen Union.

Parenthetical uses, which are less syntactically integrated, are also expected for topic shifters, given that they operate on a different level with respect to sentence syntax, and can easily be moved around in the sentence (cf. Section 1). We coded as PARENTHETICAL uses cases in which a DM was preceded and followed by commas or hyphens (7), and as IN PARENTHETICAL CLAUSE cases in which a DM was used in a parenthetical clause, including relative clauses (8).

(7)

a. This is one of my favourite topics**, by the way,** which is why I am going to consider all legislative and non-legislative measures to redress the persistent inequalities in decision making. (source language)

b. (French) C'est un de mes sujets favoris, **soit dit en passant**, c'est pourquoi je vais considérer toutes les mesures législatives et non législatives en vue de corriger les inégalités persistantes dans la prise de décision.

c. (German) Dies ist, **nebenbei gesagt,** eines meiner Lieblingsthemen, warum ich auch sämtliche Legislativmaßnahmen und nicht legislativen Maßnahmen prüfen werde, um die immer noch bestehenden Ungleichheiten in Entscheidungspositionen zu beseitigen.

(8)

a. I therefore welcome the fact that Parliament takes such a great interest in this file – and has from the very beginning, **by the way** – which is evidenced by the resolution which Parliament has put on the table. (source language)

b. (French) Je salue par conséquent le fait que le Parlement se soit tant intéressé à ce dossier – et ce dès le tout début, soit dit en passant – comme le montre la résolution déposée par le Parlement.

c. (German) Ich begrüße daher die Tatsache, dass das Parlament so großes Interesse für dieses Dossier zeigt – und das ganz von Anfang an getan hat – was durch die Entschließung bewiesen wird, die das Parlament vorgelegt hat.

We initially coded as FINAL all sentence-final uses, as in (9), but these seemed too rare enough to be significant, except in English. Consequently, in French and German, we included such uses in the last category (OTHER).

(9) Madam President, yes, I also voted against the report on a smoke-free Europe, although by doing so, I am subjecting myself to what *is tantamount to a* witch-hunt, even in the House **by the way.**

3. Results – Common topic-shifting strategies in French and German

3.1. Main strategies in German

The most frequent German translation equivalents of *furthermore*, *by the way* and *besides* in the dataset are, quite expectedly, not the same, though there is some overlap. Thus, the most frequent equivalents of *furthermore* appear to be *darüber hinaus*, *außerdem*, *zudem* and *ferner* (making up roughly 60% of all equivalents, see Table 2).

translation equivalent	absolute frequency	%[8]
darüber hinaus	1,529	21%
außerdem	1,373	19%
zudem	744	10%
ferner	654	9%
auch	506	7%
des Weiteren	452	6%
other	*2,095*	*28%*
total	7,353	100%

Table 2: Translation equivalents of *furthermore* in our dataset (only items above 5%)

The most frequent equivalents of *besides* are *außerdem*, *im Übrigen*, *übrigens* and *darüber hinaus* (making up roughly 60% of all equivalents, see Table 3).

translation equivalent	absolute frequency	%
außerdem	100	29%
im Übrigen	74	21%
übrigens	32	9%
darüber hinaus	27	8%
zudem	27	8%
other	90	*26%*
Total	350	100%

Table 3: Translation equivalents of *besides* in our dataset (only items above 5%)

The most frequent equivalents of *by the way* are *übrigens* and *im Übrigen*, totaling more than 75% of all occurrences (Table 4).

[8] I.e. percentage of all translation equivalents; the same goes for the following tables.

translation equivalent	absolute frequency	%
übrigens	273	53%
im Übrigen	118	23%
other	123	24%
Total	514	100%

Table 4: Translation equivalents of *by the way* in our dataset (only items above 5%)

Thus, for all three English DMs, there seems to be no perfect or natural German equivalent. This could be linked to various factors. For one thing, it is well-known that DMs are hard to translate; there is often no one-to-one correspondence between DMs in different languages, as is well illustrated for instance in Grieve (1996) for a long list of French DMs and their English equivalents. Pragmatics are in any case a difficult realm for L2 speakers, and has given rise to a dedicated field, cross-cultural pragmatic failure (see e.g. Trosborg 1995), which goes beyond the use of DMs alone (Pohl 2004). Besides, as already noted, our dataset does not contain translations from English into German: it is much more complex, since the source texts may be in any official language of the European Union. However, the fact that the three English DMs we selected have various translation equivalents in German is rather to our advantage, since our aim in this paper is to identify both typical topic shifters and other possible ways of signaling topic shifts: we did not set out to identify one equivalent of each English DM, but to use the English DMs as a way to identify a paradigm of DMs with uses as topic shifters.

Accordingly, in the following, we focus on the six DMs which are most frequently found as translation equivalents of the three English DMs: *außerdem*, *darüber hinaus*, *ferner*, *zudem*, *im Übrigen* and *übrigens*. As can be seen in Table 5, almost all appear as translation equivalents of *furthermore*, *besides* and *by the way*, but they display various degrees of specialization. Our hypothesis is that they could be seen as forming a paradigm of German topic shifters.

translation equivalent	**furthermore**	**besides**	**by the way**	**Total**
außerdem	1,373	100	4	1,477
darüber hinaus	1,532	27	2	1,561
ferner	655	6	1	662
zudem	745	27	0	772
im Übrigen	338	74	118	530
übrigens	119	32	273	424
ø	401	11	15	427
other	2,190	73	101	2,364
Total	7,353	350	514	8,217

Table 5: Frequent translation equivalents of *furthermore*, *besides* and *by the way* in our dataset

There is a clear contrast between the translation equivalents of *by the way* and those of *furthermore* (Figure 1): for *by the way*, we found an overwhelming majority of *im Übrigen* and *übrigens* (76%), which represent only 7% of all translation equivalents of *furthermore*. Conversely, for *furthermore*, the most frequent equivalents are *außerdem* and *darüber*

hinaus (40%), but their use as translation equivalents for *by the way*, while not unattested, seems exceptional (1%). *Besides* is an interesting case, in-between the two other markers: both pairs of DMs are frequently used as translation equivalents with similar frequency – *außerdem* and *darüber hinaus* (36%), as well as *im Übrigen* and *übrigens* (30%).

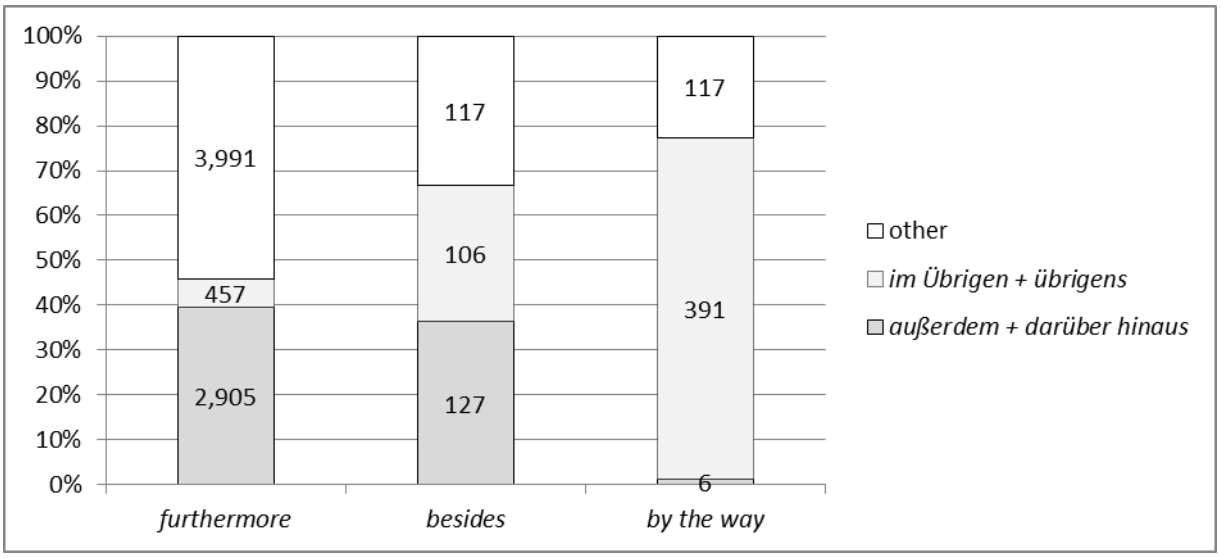


Figure 1: Frequent translation equivalents of *furthermore*, *besides* and *by the way* in our dataset

This suggests that, even though (almost) all DMs listed in Table 5 are sometimes used as equivalents for all three English DMs, some are probably better functional equivalents than others: *im Übrigen*, for instance, might well be a better translation equivalent of *by the way* than of *furthermore*. Besides, the syntactic analysis shows that the uses of a given DM in German as translation equivalent of two different DMs in English is not necessarily an indication of similarity in function between these DMs. For instance, as shown in Table 6, *übrigens* and *im Übrigen* follow different patterns when used as equivalents of *by the way* – in which case they are often found in a parenthetical clause (150 occurrences, or 38%) – and as equivalents of *besides* (4 occurrences, or 4%) or *furthermore* (36 occurrences, or 8%).

	zudem, *ferner*, *außerdem*, *darüber hinaus*				*übrigens*, *im Übrigen*			Total[9]
	in par. clause	initial	other	par.	in par. clause	initial	other	
furthermore	64	3,460	779	2	36	252	169	7,353
besides	0	123	37	0	4	63	39	350
by the way	1	3	3	0	150	103	138	514
Total	65	3,586	819	2	190	418	346	8,217

Table 6: *zudem*, *ferner*, *außerdem*, *darüber hinaus*, *übrigens*, *im Übrigen* as translation equivalents of *furthermore*, *besides*, *by the way* and their syntactic patterns in Europarl7

This suggests that, when we find a given DM as translation equivalent of two different DMs, for instance *im Übrigen* for *by the way* and *furthermore*, we may be dealing with different uses of this DM, linked to different syntactic configurations. The same goes for *außerdem*, *darüber hinaus* and *ferner*, which tend to be in initial position as equivalents of *furthermore*

[9] This column includes all translation equivalents.

and *besides*, but less so when used as equivalents of *by the way*, as shown in Table 7. This seems to indicate that one typical use of these DMs, i.e. the use of *außerdem*, *darüber hinaus* and *ferner* as additive markers (a typical function of *furthermore*), could be linked to a sentence-initial position.

		furthermore	*besides*	*by the way*	Total
außerdem	in par. clause	12	0	1	13
	initial	1,147	81	2	1,23
	other	214	19	1	234
	Total	1,373	100	4	1,477
darüber hinaus	in par. clause	15	0	0	15
	initial	1,314	19	1	1,334
	other	203	8	1	212
	Total	1,532	27	2	1,561
ferner	in par. clause	1	0	0	1
	initial	539	5	0	544
	other	113	1	1	115
	parenthetical	2	0	0	2
	Total	655	6	1	662

Table 7: *außerdem*, *darüber hinaus* and *ferner* as translation equivalents of *furthermore*, *besides*, *by the way* and their syntactic patterns in Europarl7

Similarly, *furthermore* is much more frequently associated with *außerdem*, *darüber hinaus*, *ferner* or *zudem* than with *übrigens* or *im Übrigen*, but the tendency is more pronounced for initial uses of *furthermore* (3,760 occurrences vs 357, or 60% vs 6%), least so for its parenthetical uses (359 occurrences vs 67, or 48% vs 9%), as shown in Table 8. This could be seen as illustrating the importance of the sentence-initial position for argumentative, or additive, uses of DMs.

	additives (zd,fr,au,dh)	**digressives (üb,iÜ)**	**zero**	**other**	**Total**
furthermore					
Initial	3,760	357	296	1,822	6,235
other	186	33	34	114	367
parenthetical, or in parenthetical clause	359	67	71	254	751
Total	**4,305**	**457**	**401**	**2,190**	**7,353**

Table 8: *zudem*, *ferner*, *außerdem*, *darüber hinaus*, *übrigens*, *im Übrigen* as translation equivalents of *furthermore* and their syntactic patterns in Europarl7

3.2. Main strategies in French

The most frequent French translation equivalents of *furthermore*, *by the way* and *besides* in Europarl7 are, quite expectedly, not the same, though there is some overlap. Thus, the most frequent equivalents of *furthermore* appear to be *en outre*, *par ailleurs* and *de plus* (making up roughly 71% of all equivalents, see Table 9).

translation equivalent	absolute frequency	**%**
en outre	2,634	36%
par ailleurs	1,527	21%
de plus	1,019	14%
Total	**7,353**	100%

Table 9: French translation equivalents of *furthermore* in our dataset (only items above 5%)

The most frequent equivalents of *besides* are *en outre*, *par ailleurs* and *d'ailleurs* (making up roughly 65% of all equivalents, see Table 10).

translation equivalent	absolute frequency	%
en outre	107	31%
par ailleurs	70	20%
d'ailleurs	48	14%
du reste	35	10%
de plus	28	8%
Total	350	1

Table 10: French translation equivalents of *besides* in our dataset (only items above 5%)

The most frequent equivalents of *by the way* are *d'ailleurs*, *en passant* and *à (ce) propos*, totaling more than 75% of all occurrences (Table 11).

translation equivalent	absolute frequency	%
d'ailleurs	139	27%
en passant	93	18%
à (ce) propos	67	13%
par ailleurs	46	9%
Total	514	1

Table 11: French translation equivalents of *by the way* in our dataset (only items above 5%)

Thus, for all three English DMs, there seems to be no perfect or natural French equivalent, either, presumably for the same reasons as in German (see above, section 3.1.). In the following, we focus on the DMs which are most frequently found as translation equivalents of the three English DMs: *en outre*, *par ailleurs*, *d'ailleurs*, *du reste*, *de plus*, *à (ce) propos*, and *en passant*. As can be seen in Table 12, and just as we found in German for the most frequent items, almost all appear as translation equivalents of *furthermore*, *besides* and *by the way*, but they display various degrees of specialization.

translation equivalent	**furthermore**	**besides**	**by the way**	**Total**
d'ailleurs	283	48	139	470
du reste	37	35	11	83
de plus	1,019	28	1	1,048
en outre	2,634	107	12	2,753
par ailleurs	1,527	70	46	1,643
en passant	0	1	93	94
à (ce) propos	3	2	67	72
ø	**381**	**7**	**43**	**431**
other	**1,469**	**52**	**102**	**1,623**

Total	7,353	350	514	8,217

Table 12: Frequent translation equivalents of *furthermore*, *besides* and *by the way* in French, in our dataset

As we saw in German, there is a clear contrast between the translation equivalents of *by the way* and those of *furthermore*: for *by the way*, we found a majority of *d'ailleurs*, *en passant* and *à (ce) propos* (67%), which represent only 4% of all translation equivalents of *furthermore*. Conversely, for *furthermore*, the most frequent equivalents are *en outre*, *par ailleurs* and *de plus* (70%), while their use as translation equivalents for *by the way* is quite limited (11%). As we saw in the case of German translation equivalents, *besides* is an interesting case, though not in the same way: in the first set of DMs (*d'ailleurs*, *en passant* and *à (ce) propos*), only *d'ailleurs* is regularly used as translation equivalent of *besides* (14%), while the second set (*en outre*, *par ailleurs* and *de plus*) frequently appears (59%); furthermore, another DM is found as a regular translation equivalent, though it appeared frequently neither with *furthermore* nor with *by the way*: *du reste* (10%). The contrast between the three DMs and their French translation equivalents is perhaps more visible in Figure 2.

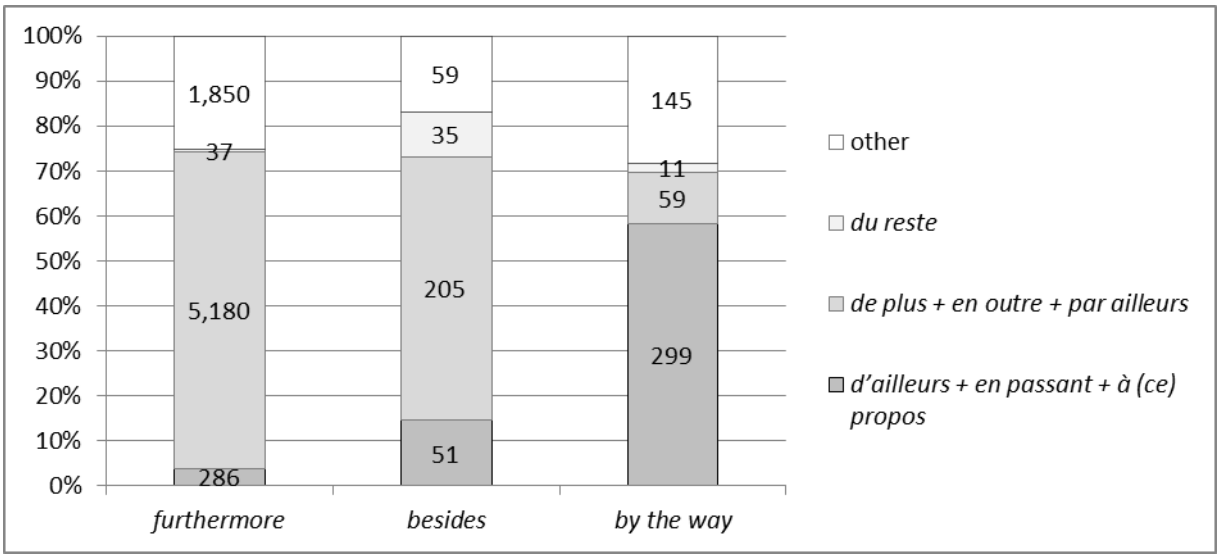


Figure 2: Frequent translation equivalents of *furthermore*, *besides* and *by the way* in French, in our dataset

The syntactic analysis shows, however, that the uses of a given DM in French as translation equivalent of two different DMs in English is not necessarily an indication of similarity in function between these DMs. For instance, as shown in Table 13, at least for these seven DMs, translation equivalents of *by the way* are much more frequently found in a parenthetical clause (122 occurrences, or 33%) than those of *furthermore* (63 occurrences, or 1%) or *besides* (9 occurrences, or 3%), and much less in initial position (143 occurrences, or 23%, vs respectively 4,384 occurrences (80%) and 227 occurrences (78%)).

	syntax	*d'ailleurs*	*du reste*	*de plus*	*en outre*	*par ailleurs*	*en passant*	*à (ce) propos*	subtotal	%	**Total**[10]
furthermore	par. clause	47	0	0	7	9	0	0	63	1%	**90**
	initial	113	22	944	2,061	1,242	0	2	4,384	80%	**5,166**
	other	108	11	27	400	180	0	1	727	13%	**1,668**
	par.	15	4	48	166	96	0	0	329	6%	**429**
	Total	283	37	1,019	2,634	1,527	0	3	5,503	100%	7,353
besides	par. clause	4	1	1	1	2	0	0	9	3%	**9**
	initial	22	24	26	94	59	1	1	227	78%	**263**
	other	22	10	1	10	5	0	1	49	17%	**72**
	par.	0	0	0	2	4	0	0	6	2%	**6**
	Total	48	35	28	107	70	1	2	291	100%	350
by the way	par. clause	54	4	0	0	13	47	4	122	33%	**147**
	initial	20	5	1	7	19	30	61	143	39%	**180**
	other	58	2	0	3	13	7	2	85	23%	**155**
	par.	7	0	0	2	1	9	0	19	5%	**32**
	Total	139	11	1	12	46	93	67	369	100%	514

Table 13: *d'ailleurs*, *en passant*, *à (ce) propos*, *en outre*, *par ailleurs*, *de plus* and *du reste* as translation equivalents of *furthermore*, *besides*, *by the way* and their syntactic patterns in Europarl7

As we noted in the case of German DMs, this suggests that the use of a French DM as translation equivalent of two different English DMs in our dataset may in fact reveal two distinct use-types of this DM. And these use-types, as we saw in the case of German DMs, may be linked to different syntactic configurations. However, this difference in syntactic behavior is very clear for some DMs, less for others. For instance, translation equivalents of *by the way* are quite often found in parenthetical clauses (33%), *d'ailleurs* and *en passant* accounting for most of these occurrences in a parenthetical clause: this is perfectly in line with our assumption that *by the way* is a typical topic shifter, given that parenthetical clauses are precisely dedicated to side comments, information given as afterthoughts, etc. Moreover, the case of *par ailleurs* when used as a translation equivalent of *by the way* is interesting: though the syntactic pattern of *par ailleurs* in our dataset is strongly biased toward sentence-initial uses (overall 1,320 occurrences, i.e. 80% of the total) and away from uses in a parenthetical clause (overall 28 occurrences, i.e. 1%), when used as a translation equivalent of *by the way*, *par ailleurs* displays a very different syntactic pattern, much more typical of a digressive marker, with 13 occurrences (28%) in a parenthetical clause and 19 (41%) sentence-initial occurrences. Conversely, translation equivalents of *furthermore* are very often sentence-initial (80%), and *d'ailleurs* can be seen to provide a mirror image of what we just showed for *par ailleurs*: when used as a translation equivalent of *furthermore*,

[10] Including other translation equivalents.

d'ailleurs is much more frequently in sentence-initial position than when it is a translation equivalent of *by the way* (40% vs 14%)[11].

As in German, *furthermore* is typically associated with a restricted set of DMs – in French, *en outre*, *par ailleurs* and *de plus* – but this tendency is much more pronounced for the initial uses of *furthermore* (in that case, the three French DMs account for 4,614 occurrences, or 74% of all translation equivalents) than for its parenthetical uses (391 occurrences, or 52%), as shown in Table 14. Conversely, *d'ailleurs* is found in only 3% of cases for initial *furthermore*, and 8% for its parenthetical uses. Together with what we noted concerning the uses of *d'ailleurs* and *par ailleurs* in sentence-initial position and in parenthetical clauses, this indicates a strong association between use-types (argumentative vs digressive), syntactic patterns (sentence-initial vs in parenthetical clause) and specific DMs (here, *d'ailleurs* vs *par ailleurs*).

furthermore	initial	parenthetical or in parenthetical clause	other
en outre	2,324	203	107
par ailleurs	1,363	120	44
de plus	927	68	24
d'ailleurs	195	61	27
du reste	29	7	1
à (ce) propos	2	1	0
zero	220	101	60
other	1,175	190	104
total	6,235	751	367

Table 14: French translation equivalents of *furthermore* depending on its syntactic pattern, in Europarl7

A similar phenomenon can be observed concerning the translation equivalents of *by the way*. For instance, *à (ce) propos* is found as translation equivalent mostly when *by the way* is in initial position, *d'ailleurs* when it is final or parenthetical (see Table 15).

by the way	*d'ailleurs*	*du reste*	*de plus*	*en outre*	*par ailleurs*	*en passant*	*à (ce) propos*	*zero*	other	Total
Final	24	4	0	0	1	14	0	9	12	64
Initial	30	2	1	4	12	24	46	5	31	155
other	5	0	0	0	2	6	2	3	7	25
par.	80	5	0	8	31	49	19	26	52	270
Total	139	11	1	12	46	93	67	43	102	514

Table 15: French translation equivalents of *by the way* depending on its syntactic pattern, in Europarl7

[11] The difference, calculated for (uses in parenthetical clauses * initial uses * other)*(*d'ailleurs* as translation equivalent of *by the way* * of *furthermore*), is statistically significant ($chi^2 = 38.8$, $p < 0.01$). The same can be said of *par ailleurs* (same values; $chi^2 = 257.1$, $p < .01$).

Taken together, these observations suggest that there are more or less typical uses of each DM. Typical uses of *furthermore* are sentence-initial and argumentative, typical uses of *by the way* are in a parenthetical clause and digressive. In both French and German, these uses are more closely associated, in turn, with DMs which are argumentative and digressive, respectively. These first results warrant a further investigation into syntactic patterns.

3.3. Syntactic patterns

The syntactic patterns found for the DMs under study are slightly different. *Furthermore* and *besides* are frequently used initially, while *by the way* is more frequently parenthetical (Table 16).

English	Initial	parenthetical	other	in parenthetical clause	Final	**Total**
furthermore	6,235	586	367	165	0	**7,353**
besides	316	24	10	0	0	**350**
by the way	155	263	25	7	64	**514**
Total	6,706	873	402	172	64	**8,217**

Table 16: Syntax of *furthermore*, *besides* and *by the way* (Europarl7)

In line with Lamiroy and Charolles (2004), and Fischer (2006), we consider that the frequent initial uses of *furthermore* and *besides* indicates their tendency to highlight logical relations, and take on an ARGUMENTATIVE use. The frequent parenthetical uses of *by the way*, on the other hand, are consistent with its analysis as a genuine DIGRESSIVE marker, or topic shifter (Traugott 2019).

As noted in section 3.2., French DMs globally appear to follow quite closely the syntactic distributions of their English counterparts, with a high frequency of initial uses, especially for *de plus*, *par ailleurs* and *en outre* (roughly 80% to 90% of their uses, and more or close to 1,000 sentence-initial occurrences). However, the proportion of initial uses is slightly lower than in English, while that of parenthetical uses is slightly higher (Table 17).

French	initial	other	parenthetical	in parenthetical clause	Total
d'ailleurs	155	188	22	105	470
du reste	51	23	4	5	83
en outre	2,162	413	170	8	2,753
par ailleurs	1,320	198	101	24	1,643
de plus	971	28	48	1	1,048
en passant	31	7	9	47	94
à (ce) propos	64	4	0	4	72
other	855	603	113	52	1,623
zero	0	431	0	0	431
Total	5,609	1,895	467	246	8,217

Table 17: Syntax the most frequent French translation equivalents of *furthermore*, *besides* and *by the way* (Europarl7)

German DMs also follow closely the syntactic distributions of their English counterparts, albeit with a lower frequency of initial and parenthetical uses (Table 18). The DMs which are most often found in sentence-initial position are *außerdem*, *darüber hinaus* and *ferner* (all three above 80% of initial uses, with more than 500 sentence-initial occurrences), while only *übrigens* and *im Übrigen* have relatively frequent parenthetical uses.

German	initial	other	in parenthetical clause	parenthetical	Total
außerdem	1,230	234	13	0	1,477
darüber hinaus	1,334	212	15	0	1,561
ferner	544	115	1	2	662
zudem	478	258	36	0	772
im Übrigen	311	152	67	0	530
übrigens	107	194	123	0	424
other	1,583	671	83	27	2,364
zero	0	427	0	0	427
Total	5,587	2,263	338	29	8,217

Table 18: Syntax the most frequent German translation equivalents of *furthermore*, *besides* and *by the way* (Europarl7)

These results confirm the observation of a link between DM, use-type and syntactic distribution. In both French and German, in our dataset, some DMs tend to be used sentence-initially and have argumentative uses, while others tend to be used in a parenthetical clause and have digressive uses.

3.4 German and French translation equivalents

There thus seems to be a close correspondence between the two sets of DMs under study, i.e. between the French and German translation equivalents of *furthermore*, *besides* and *by the way*. We have seen that there are multiple contrasts, for instance the initial uses of *furthermore*, or the parenthetical uses of *by the way*. As noted in sections 3.2. and 3.3, we believe these could be seen as related to typical discourse functions. More specifically, we think initial *furthermore* corresponds to ADDITIVE uses, and parenthetical *by the way* to DIGRESSIVE uses. Though all three DMs are associated to a wide variety of translation equivalents, those linked to these uses are much more limited. Thus, initial uses of *furthermore* are associated to the French DMs *en outre*, *par ailleurs*, and *de plus*, and to the German DMs *außerdem*, *darüber hinaus*, *ferner* and *zudem*. Parenthetical uses of *by the way* are linked to the French DMs *d'ailleurs* and, more loosely, *par ailleurs*, and to the German DMs *übrigens* and *im Übrigen*.

This suggests a stronger correspondence for some of these French and German DMs than for others. This is confirmed by a direct analysis of their uses in our dataset. As shown in Tables 19 and 20, there is a strong overlap on the one hand between *en outre* (also *de plus*, *du reste*) and *außerdem*, *darüber hinaus*, *ferner* and *zudem* – all typically displaying ADDITIVE uses –, on the other between *im Übrigen*, *übrigens* and *par ailleurs*, *d'ailleurs*, *à (ce) propos* and *en passant* – all typically displaying DIGRESSIVE uses.

German-French	*en outre*	*de plus*	*par ailleurs*	*d'ailleurs*	*en passant*	*à propos*	*du reste*
außerdem	45%	17%	18%	2%	0%	0%	0%
darüber hinaus	46%	16%	20%	2%	0%	0%	0%
ferner	42%	16%	22%	0%	0%	0%	0%
zudem	39%	17%	20%	2%	0%	0%	0%
im Übrigen	10%	4%	32%	27%	3%	2%	7%

übrigens	7%	2%	13%	37%	11%	11%	3%

Table 19: German-French translation equivalents in our dataset (shades of green[12] indicate the strength of the correlation)

French-German	*außerdem*	*darüber hinaus*	*ferner*	*im Übrigen*	*übrigens*	*zudem*
de plus	24%	24%	10%	2%	1%	12%
en outre	24%	26%	10%	2%	1%	11%
par ailleurs	16%	19%	9%	10%	3%	9%
d'ailleurs	6%	5%	0%	31%	34%	4%
du reste	8%	4%	1%	47%	17%	2%
en passant	1%	0%	0%	19%	48%	0%
à (ce) propos	0%	1%	0%	19%	65%	0%

Table 20: French-German translation equivalents in our dataset (shades of green indicate the strength of the correlation)

Figure 5 below presents a schema of these correspondences, based on the hypothesis that the French and German DMs under study can be placed along a cline from ADDITIVE to DIGRESSIVE functions. This does not mean that a DM on one side of the cline cannot take on the opposite function: as noted in Sections 1 and 3.3, discourse functions are also linked to syntactic position, so the DIGRESSIVE use of, say, German *außerdem* remains possible, but in a PARENTHETICAL position.

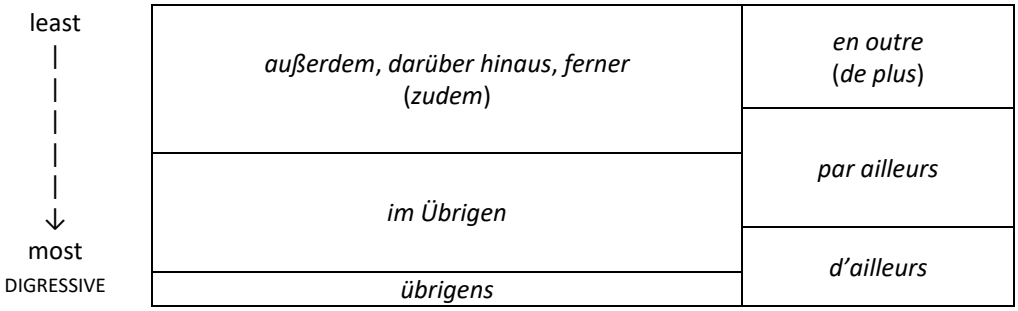


Figure 5: Correspondence between French and German DMs signaling topic shifts

4. Complementary corpus studies

In Section 3, we analyzed the uses of a series of French and German DMs, on the basis of their occurrences as translation equivalents of *furthermore*, *besides* and *by the way* in the parallel corpus Europarl7. Our results suggest that they can be placed along a cline from more ADDITIVE (and connective) functions to more DIGRESSIVE ones. However, the focus on three English DMs, and the nature of the corpus, invites caution: is this a good characterization of the DMs under study? In order to evaluate how well our description fits, we ran a few more tests, with complementary corpus studies. First, we tested the correspondence between the French and German DMs under study, on the same dataset (Section 4.1). We then tried to see if our method for identifying TOPIC SHIFTERS was successful, by looking at the translation equivalents of another typical TOPIC SHIFTER, *incidentally*. Finally, we looked for *genre effects* (Section 4.3).

[12] Provided by an Excel function.

4.1. Correspondence between French and German DMs

We tested the validity of the correspondences between French and German DMs (described in Section 3.4) with a complementary corpus, checking for the presence of the six German DMs and seven French DMs as translation equivalents of one another. The goal of this second experiment is specifically to identify the importance of the use-types we identified among all uses of a given DM: for instance, if we consider *im Übrigen* and *übrigens* in German and *d'ailleurs*, *en passant* and *à (ce) propos* in French a good examples of DIGRESSIVE markers, we may expect them to show a good deal of overlap, i.e. to be frequent translation equivalents of one another. More importantly, we may expect this overlap to account for an important portion of their uses; in other words, our hypothesis is that looking for translation equivalents of *im Übrigen* and *übrigens* should yield not only many instances of *d'ailleurs*, *en passant* and *à (ce) propos*, but also few instances of other DMs. We have the same expectations for the ARGUMENTATIVE DMs identified in section 3.

In order to check this assumption, we looked for occurrences of these DMs in the Europarl7 corpus. Given the high frequency of most of these DMs (some with more than 15,000 occurrences), we checked only a random sample of 1,000 occurrences for each DM (with the exception of *en passant*, which appears only 329 times in the corpus). This yielded a second dataset, described in Table 21.

sequence	**occurrences**
im Übrigen	1,000
außerdem	1,000
darüber hinaus	1,000
ferner	1,000
übrigens	1,000
zudem	1,000
d'ailleurs	1,000
à (ce) propos	1,000
de plus	1,000
du reste	1,000
en outre	1,000
par ailleurs	1,000
en passant	329
Total	**12,329**

Table 21: Complementary dataset (most frequent French and German translation equivalents of *furthermore*, *by the way* and *besides* in the Europarl7 database)

We then checked for the presence of the translation equivalents in the other language. The results, with the frequency of DMs found as translation equivalents, are presented in Table 22 (German translation equivalents of French DMs) and Table 23 (French translation equivalents of German DMs). These tables show that the equivalence found with the help of English DMs is good enough for some DMs, not so much for others. Among the French DMs under study, for instance, there is a good overlap for *en outre*, *par ailleurs* and *en passant*,

not so much for *d'ailleurs* and *du reste*, and the overlap seems actually very poor for *de plus* and *à (ce) propos*, which thus probably have many uses besides those identified in section 3.

#	*außerdem*	*darüber hinaus*	*ferner*	*im Übrigen*	*übrigens*	*zudem*	other	Total
en outre	220	203	71	25	9	86	386	1,000
par ailleurs	126	104	44	88	38	45	555	1,000
en passant	4	1	0	53	116	1	154	329
d'ailleurs	10	8	3	180	158	26	615	1,000
du reste	20	5	10	207	98	16	643	999
de plus	13	9	7	3	1	7	959	999
à (ce) propos	8	2	4	4	4	7	971	1,000
Total	**401**	**332**	**139**	**560**	**424**	**188**	**4,283**	**6,327**

Table 22: Complementary dataset – German translation equivalents of French DMs

Similarly, the overlap is not the same across the German DMs: it is best for *im Übrigen* and *übrigens* (for which the overlap with *d'ailleurs* is very important), *darüber hinaus*, worst for *ferner* (Table 24).

#	*en outre*	*par ailleurs*	*du reste*	*de plus*	*en passant*	*à (ce) propos*	*d'ailleurs*	*other*	*Total*
im Übrigen	49	204	49	26	20	21	324	307	1,000
übrigens	31	78	41	13	35	32	437	333	1,000
darüber hinaus	262	98	1	116	1	3	19	500	1,000
außerdem	185	68	1	71	1	4	17	653	1,000
zudem	169	62	3	91	0	3	18	654	1,000
ferner	141	64	0	55	0	8	10	722	1,000
Total	**837**	**574**	**95**	**372**	**57**	**71**	**825**	**3,169**	**6,000**

Table 23: Complementary dataset – French translation equivalents of German DMs

In our view, these results are linked to the multifunctionality of DMs: a DM may have ARGUMENTATIVE and/or DIGRESSIVE uses among many other use-types. This means that the description provided in section 3 covers only a portion of the uses of the DMs under study – apparently a small portion in the case of *à (ce) propos* and *de plus* in French, but a fair portion nonetheless for a few others, among which German *im Übrigen* and *übrigens* and French *en outre*. They also show that the distinction between ARGUMENTATIVES and DIGRESSIVES seems to hold, since in both directions (German translation equivalents of French DMs and vice-versa) the translation equivalents of ARGUMENTATIVES are much more frequently ARGUMENTATIVES than DIGRESSIVES (German translation equivalents: 986 occurrences vs 469; French translation equivalents: 1,387 occurrences vs 84), and vice-versa (German translation equivalents: 515 occurrences vs 74; French translation equivalents: 869 occurrences vs 491)[13].

[13] In both cases, the difference is statistically significant: for German > French, $chi^2 = 1071.4$, $p < .01$; for French > German, $chi^2 = 511.8$, $p < .01$.

4.2. Identifying topic shifters: translation equivalents of *incidentally*

The third experiment was designed to test the validity of our hypothesis concerning French and German topic shifters, i.e. to see whether we had succeeded in correctly identifying sets of French and German DMs used to mark topic shifts, especially concerning DIGRESSIVES: German *im Übrigen* and *übrigens* and French *d'ailleurs*, *en passant* and *à (ce) propos*.

In order to do this, we looked for translation equivalents of *incidentally*, which we chose as being another English DM typically used as a DIGRESSIVE (Traugott to appear). Our hypothesis was that its translation equivalents would be primarily those identified in section 3 as digressive markers, especially when it is used parenthetically. Following the procedure described in Section 2, we retrieved all occurrences of *incidentally* in Europarl7 and the corresponding contexts in French and German.

The results tend to confirm our hypothesis: the main translation equivalents are, in French, *d'ailleurs* and *en passant* (as well as, less often, *par ailleurs*, *du reste* and *en outre*) (table 25); in German, *übrigens* and *im Übrigen* (table 26). Besides, in both languages, the overlap between *incidentally* and its most frequent translation equivalents is rather good: the DMs we identified as rather ADDITIVE than DIGRESSIVE were rarely found as translation equivalents of *incidentally*, and the category "other" is less important than what we saw in Section 4.1 for some DMs.

translation equivalent	parenthetical	initial	other	Total
d'ailleurs	187	81	7	275
en passant	82	46	7	135
par ailleurs	37	37	3	77
du reste	29	13	2	44
en outre	16	30	1	47
de plus	2	6	1	9
à (ce) propos	2	15	1	18
other	191	147	12	350
Total	546	375	34	955

Table 25: French translation equivalents of *incidentally* depending on its syntactic patterning (Europarl7)

translation equivalent	parenthetical	initial	other	Total
übrigens	255	156	18	429
im Übrigen	136	124	6	266
außerdem	5	5	0	10
zudem	4	5	0	9
darüber hinaus	2	1	0	3
ferner	0	2	0	2
other	144	82	10	236
Total	546	375	34	955

Table 26: German translation equivalents of *incidentally* depending on its syntactic patterning (Europarl7)

The importance of parenthetical uses for frequent translation equivalents of *incidentally*, both in French (66% for *d'ailleurs*, 61% for *en passant*) and in German (59% for *übrigens*, 51% for *im Übrigen*) is another indication than these DMs are indeed used as DIGRESSIVES.

4.3. Genre effects

All corpus studies presented so far are based on the Europarl corpus. This means that our results describe the use of DMs as topic shifters not in French and German as a general rule, but only in this specific corpus. This might be a strong bias, given the tendency of specialized discourses to become homogeneous across European languages (e.g. for legal language, see Grzmil-Tylutki 2014: 39). Though we cannot hope to remedy this with a single additional experiment, we believe comparing our main dataset with similar data from a different corpus – i.e. one with texts from different genres – can provide an indication of the extent of this restriction.

In order to see whether there might be an important difference between our datasets and the use of DMs in literary text, we used the Parasol corpus (von Waldenfels et al. 2006), restricting our search to texts available in English, French and German: this limited the corpus to three novels, *The master and Margarita* by Mikhail Bulgakov (translated from Russian into all three languages) and *Harry Potter* (volumes one and two) by J. K. Rowling (translated from English into the two other languages). One limitation of this corpus is its size: we found only 58 occurrences in total, one of *furthermore*, nine of *by the way*, and 48 of *besides*. The results are thus only qualitative[14]. They confirm the possibility of having no translation equivalent for a given DM, e.g. *by the way* in the French translation in example (10): *si tu veux savoir* (lit. 'if you want to know') translates *did you know?*, while *by the way* is left implicit.

(10)

a. "All right — I only came in here because people outside are behaving very childishly, racing up and down the corridors," said Hermione in a sniffy voice. "And you've got dirt on your nose, by the way, did you know?"

b. J'étais venue vous voir parce que les autres ne font que des bêtises, ils courent dans le couloir comme des idiots et toi, tu as une saleté sur le nez, si tu veux savoir.

c. (...) Und **übrigens**, du hast Dreck an der Nase, weißt du das?

Another result seems to be confirmed: the wide variety of strategies available to speakers/writers, the most frequent markers totalling less than 25% of all translation equivalents in French (*en outre*), less than 30% in German (*außerdem*).

But perhaps the most interesting result is the low frequency of two of the DMs we chose, *by the way* (with nine occurrences) and *furthermore*. Testing the frequency of the French and

[14] Note that there is an important contrast, within this small corpus, between *The master and Margarita* (which was written in Russian, as specified, and for which the English DMs are therefore themselves translations) and *Harry Potter* (in which the English DMs are original). As noted by one of the reviewers, there could be some effect on the results. Unfortunately, these results are quantitatively insufficient to reasonably test this hypothesis, which we will have to pick up on in future research.

German DMs (not only as translation equivalents of the three English DMs), we saw that some of them, too, were hardly to be found. This is the case, for instance, of French *par ailleurs*. This could mean that the use of DMs in Europarl is mostly genre-related. We tested this hypothesis for *par ailleurs*, by checking its frequency across genres. As shown in Table 27, *par ailleurs* is indeed very specialized, and displays a high frequency in the press and in essays, much higher than in novels – its frequency in the Press corpus is ten times higher than in the Novels corpus. The same holds for *en outre*, but the difference is much less important. *D'ailleurs*, on the other hand, is more frequent in the Novels corpus.

DM	FrTenTen	Frantext – novels (20th century)	Frantext – essays (20th century)	BTLC – Press (*Le Monde* 2008)[15]
par ailleurs	32.6 (373,017)	8.2 (363)	43.4 (336)	84.1 (1,717)
d'ailleurs	144.7 (1,656,374)	295.2 (13,109)	278.9 (2,159)	122.1 (2,493)
en outre	31.3 (358,168)	18.1 (802)	40.6 (314)	50.1 (1,022)
corpus size (millions)	9,890	44.4	7.7	20.4

Table 27: Frequency of a few French DMs across genres and corpora – frequency per million (number of occurrences)

This suggests that there is indeed a strong genre effect for at least some DMs.

5. Conclusion

In this paper, we attempted to draw a map of French and German DMs used to signal topic shifts. In order to do this, we analyzed the uses of a series of French and German DMs, on the basis of their occurrences as translation equivalents of three English DMs, *furthermore*, *besides* and *by the way*, in a parallel corpus (Europarl7). Our results suggest that they can be placed along a cline from more ADDITIVE (and connective) functions to more DIGRESSIVE ones, and that, furthermore, some DMs are more typically associated with these functions than others. For instance, in French, *d'ailleurs* and to a lesser extent *par ailleurs* seem to be good instances of TOPIC SHIFTERS; this seems to be true for German *übrigens* and *im Übrigens*. At the other end of the cline, French *en outre* and German *außerdem* can be used to signal topic shifts, especially when they follow specific syntactic patterns, but it is not their primary function. However, after complementary tests, all DMs were found to be multifunctional; besides, a qualitative analysis on another corpus seems to indicate that there are strong genre effects.

Further studies on French and German TOPIC SHIFTERS should thus include a wider range of genres. Another interesting perspective is that mentioned by Traugott (2019): she shows the existence of different patterns of DM formation at different periods (a spatial pattern in Early Modern English, and manner adverbs in the 19th century). It would be interesting to see in what ways patterns of DM formation may have contributed to the selection of the DMs we looked at, to the exclusion of others. Among the topic shifters we analyzed, for

[15] Diwersy (2005).

instance, two are quite recent (as DMs): German *im Übrigen* and French *par ailleurs*, both adverbials based on a prepositional phrase. This could be another instance of the convergence of European languages, and certainly deserves further study.

References

[excluding author references]

Aijmer, K., Altenberg, B., 1996. Introduction, in: Aijmer, K., Altenberg, B., Johansson, S. (Eds.), *Languages in contrast. Text-based cross-linguistic studies*. Lunds Universitets Publikationer, Lund, 11-16.

Aijmer, K., 2008. Translating discourse particles: a case of complex translation, in: Anderman, G., Rogers, M. (Eds.), *Incorporating Corpora: The Linguist and the Translator*. Multilingual Matters, Clevendon, 95-116.

Aijmer, K., 2002. *English discourse particles. Evidence from a corpus*. John Benjamins, Amsterdam, Philadelphia.

Brinton, L.J., 1996. *Pragmatic Markers in English: Grammaticalization and discourse functions*. Mouton de Gruyter, Berlin, New York.

Cartoni, B., Zufferey, S., Meyer, T., 2013. Using the Europarl corpus for cross-linguistic research. *Belgian Journal of Linguistics* 27, 23-42.

Charolles, M., (submitted). Discourse Topics and Digressive Markers.

Cortés Rodríguez, L.M., Hidalgo Downing, R., 2015. Indicadores de cambio temático en el discurso parlamentario: análisis de un discurso de Rodríguez Zapatero [2011]. *Revista Signos. Estudios de lingüística* 89, 279-306.

Dik, S.C., 1997. Pragmatic Functions, in: *The Theory of Functional Grammar* Part I: *The structure of the clause.* Mouton-de Gruyter, Berlin, New York.

Diwersy, S., 2015. *Kölner Romanistische Korpusdatenbank*. Romanisches Seminar der Universität zu Köln, Köln.

Eisenberg, P. 1998. *Duden. Grammatik der deutschen Gegenwartssprache*. Dudenverlag, Mannheim.

Fischer, K. (Ed.), 2006. *Approaches to Discourse Particles*. Elsevier, Amsterdam.

Forker, D., 2016. Toward a typology for additive markers. *Lingua* 180, 69-100.

Fraser, B., 1996. Pragmatic markers. *Pragmatics* 6(2), 167-190.

Fraser, B., 2009. Topic Orientation Markers. *Journal of Pragmatics* 41(5), 892-898.

Grzmil-Tylutki, H., 2014. Langues spécialisées ou langues discursives – Question terminologique ou choix épistémologique ? *Roczniki Humanistyczne* 62(8), 27-47.

Hansen, M.-B. M., 1998. *The Function of Discourse Particles*. John Benjamins, Amsterdam, Philadelphia.

Haspelmath, M., 2001. The European linguistic area: Standard Average European. *Language Typology and Language Universals*. De Gruyter, Berlin, 1492–1510.

Horst, J. van der, 2013. *Taal op drift: lange-termijnontwikkelingen in taal en samenleving.* Meulenhoff, Amsterdam.

Iordanskaja, L., Mel'čuk, I., 1999. Textual connectors across languages: French 'en effet' *vs*. Russian 'samom dele', in: Mey, J.-L., Boguslawski, A. (Eds.), *E Pluribus Una*. *Festschrift Anna Wierzbicka*. *RASK* 9-10, 305-347.

Johansson, S., 2007. *Seeing through multilingual corpora: on the use of corpora in contrastive studies*. John Benjamins, Amsterdam, Philadelphia.

Koehn, P., 2005. Europarl: A parallel corpus for statistical machine translation. Conference proceedings: 10th Machine Translation Summit, AAMT, Phuket, Thailand, Vol. 5, 79-86.

König, E., 1991. *The Meaning of Focus Particles: A Comparative Perspective*. Routledge, London.

König, E., 2017. Syntax and semantics of additive focus markers from a cross-linguistic perspective. A tentative assessment of the state of the art, in: De Cesare, A.-M., Andorno, C. (Eds.), *Focus on Additivity. Adverbial modifiers in Romance, Germanic and Slavic languages*. John Benjamins, Amsterdam, Philadelphia, 23-43.

Krifka, M., 2007. Basic Notions of Information Structure. *Interdisciplinary Studies on Information Structure* 6, 13-55.

Lambrecht, K., 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge University Press, New York.

Lamiroy, B., Charolles, M., 2004. Des adverbes aux connecteurs: 'simplement', 'seulement', 'malheureusement', 'heureusement'. *Travaux de linguistique* 49, 57-79.

Lamiroy, B., Vanderbauwhede, G., 2016. Les marqueurs de discours 'en fait', 'de fait', 'en effet' et 'en réalité' et leurs équivalents néerlandais: indices de grammaticalisation, in: Sarda L., Vigier, D., Combettes, B. (Eds.), *Connexion et indexation. Ces liens qui tissent le texte.* ENS Éditions, Lyon, 195-208.

Lamiroy, B., 2014. Contrastive linguistics as a discovery procedure. *IVth International Conference on Using Corpora in Contrastive and Translation Studies*. University of Lancaster. (keynote lecture)

Lamiroy, B., Vanderbauwhede, G., 2019 (forthcoming). On two French discourse markers and their Dutch equivalents: *d'ailleurs* and *par ailleurs*. *Journal of Pragmatics*.

Lewis, D., 2006. Contrastive analysis of adversative relational markers, using comparable corpora, in: Aijmer, K., Simon-Vandenbergen, A.-M. (Eds.), *Pragmatic markers in contrast*. Elsevier, Amsterdam, 139-153.

Métrich, R., Faucher, E. (with J. Albrecht), 2009. *Wörterbuch deutscher Partikeln: unter Berücksichtigung ihrer französischen Äquivalente*. De Gruyter, Berlin, New York.

Mittwoch, A., Huddleston, R., Collins, P., 2002. The clause adjuncts, in: Huddleston, R., Pullum, G. (Eds.), *The Cambridge Grammar of the English Language*, Chapter 8. Cambridge University Press, Cambridge, 663-784.

Pohl, G. 2004. Cross-cultural pragmatic failure and implications for language teaching, *Second Language Learning and Teaching* 4 [http://eprints.usq.edu.au/id/eprint/27625].

Pons Bordería, S., 2018. Paths of grammaticalization: beyond the lp/rp debate, in: Pons Bordería, S., Loureda Lamas, O. (Eds.), *Beyond Grammaticalization and Discourse Markers – New Issues in the Study of Language Change*. Brill, Leiden, Boston, 334-383.

Pons Bordería, S., Estellés Arguedas, M., 2009. Expressing digression linguistically: Do digressive markers exist? *Journal of Pragmatics* 41, 921-936.

Prévost, S. 2011. *À propos*: From verbal complement to 'utterance marker' of discourse shift, *Linguistics* 49 (2), 391-413.

Schiffrin, D., 1987. *Discourse Markers*. Cambridge University Press, Cambridge.

Schourup, L.C., 1983. Common Discourse Particles in English Conversation. *Ohio State University Working Papers in Linguistics 28*.

Schourup, L.C., 1999. Discourse markers, *Lingua* 107, 227-265.

Suzuki, S., 1995. The functions of topic-encoding zero-marked phrases: A study of the interaction among topic-encoding expressions in Japanese. *Journal of Pragmatics* 23, 607-626.

Traugott, E.C. (forthcoming). Expressions of stance-to-text: Discourse management markers as stance markers.

Traugott, E.C., 2019 (forthcoming). The development of "digressive" discourse-topic shift markers in English. *Journal of Pragmatics*.

Trosborg, A., 1995. *Interlanguage Pragmatics: Requests, Complaints, and Apologies*. Mouton de Gruyter, Berlin, New York.

van der Auwera, J., 2011. Standard Average European, in: Kortmann, B., van der Auwera, J. (Eds.), *The Languages and Linguistics of Europe: A Comprehensive Guide*. Mouton de Gruyter, Berlin, New York, 291–306.

Van Kuppevelt, J., 1995. Main structure and side structure in discourse. *Linguistics* 33(4), 809-833.

Vanderbauwhede, G., 2012. Les emplois référentiels du SN démonstratif en français et en néerlandais: pas du pareil au même. *Journal of French Language Studies* 22, 273-294.

von Waldenfels, R., Meyer, R., 2006. *ParaSol, a Corpus of Slavic and Other Languages*. Available at parasol.unibe.ch. Regensburg, Bern.

Whorf, B., 1956. *Language, Thought and Reality. Selected Writings of Benjamin Lee Whorf. Edited by John B. Carroll*. MIT Press, Cambridge.

Zufferey, S., 2016. Discourse connectives across languages: Factors influencing their explicit or implicit translation. *Languages in Contrast* 16, 264-279.