The representation of modal meaning of French sentence adverbs in a qualitative and quantitative approach*

Corinne Rossari (Neuchâtel)

Abstract

Our aim is to give a representation of the modal contribution of sentence adverbs in comparison to other forms conveying modal meaning, such as tenses and modal verbs. Our analysis will focus on modal sentence adverbs conveying epistemic meaning. These will be compared with the modal verbs *pouvoir* (can) and *devoir* (must) as well as with some uses of the future tense (called epistemic uses), with the purpose to present a model allowing to apprehend modal meanings transmitted by lexical and grammatical forms in order to differentiate their functioning. We will then substantiate our qualitative analysis by quantitative studies on the collocates that the modal sentence adverbs co-occur with in contemporary corpora constituted of 21ˢᵗ century newspapers, as well as in two other corpora representing two different genres and time periods: Universalis Encyclopedia and the digital edition of the Encyclopedia of Diderot and d'Alembert.

1   Introduction

Our aim is to give a representation of the modal contribution of sentence adverbs in comparison to other forms conveying modal meaning, such as tenses and modal verbs. Our analysis will focus on modal sentence adverbs conveying epistemic meaning. This category includes the following items1: *assurément* (certainly), *certainement* (certainly), *certes* (admittedly), *indubitablement* (undoubtedly), *peut-être* (maybe), *probablement* (probably), *sans doute* (undoubtedly), *sûrement* (surely), *vraisembablement* (probably). These will be compared with the modal verbs *pouvoir* (can) and *devoir* (must) as well as with some uses of the future tense (called epistemic uses). Any form in this paradigm (adverb, verb, or tense) is likely to convey an epistemic evaluation by the speaker.

1.   « *L'heure de l'antidisciplinaire est peut-être venue* » *(Internet)*
     'Maybe the time has come for antidisciplinarity'

---

1 Since the values transmitted by a form are only partially shared, the translations are purely indicative. For the examples, we add a "literal" translation when no equivalent form is available in English.

In this use, the sentence adverb *peut-être* can be replaced by any modal form conveying an epistemic evaluation:

2. L'heure de l'antidisciplinaire doit être/peut être/sera venue
   'The time must/may/will have come for antidisciplinarity'

Other modal sentence adverbs could also be used:

3. L'heure de l'antidisciplinaire est probablement/certes/certainement/sans doute/… venue
   'The time has probably/certainly/undoubtedly come for antidisciplinarity'

Although all of the above-mentioned forms can convey a similar modal meaning, which can be considered as being epistemic in the sense that the speaker gives an evaluation about the plausibility of the state of affairs, they do not share the same usage. Thus, the purpose of this article is to present a model to apprehend modal meanings transmitted by lexical and grammatical forms in order to differentiate their functioning. First, we will distinguish modal sentence adverbs from other modal forms, such as the future tense and the modal verb (*pouvoir*), cf. sect 1.1 to 1.3. Then, we will go deeper into the analysis of modal sentence adverbs looking at the properties that distinguish the use of some of them, cf. sect 1.4 and 1.5. In the second part, we will present quantitative analyses of the collocates that the modal sentence adverbs co-occur with in contemporary corpora constituted of 21st century newspapers (sect. 2). We will interpret these results (sect. 2.2.) by comparing them with two other corpora representing two different genres and time periods: Universalis Encyclopedia (sect. 2.2.1.) and the digital edition of the Encyclopedia of Diderot and d'Alembert (sect. 2.2.2). Finally (sect. 3), we will discuss the correlations between qualitative and quantitative analyses of the adverbs in comparison with a similar quantitative analysis of the modal verb *pouvoir* since it shares, as we will see, one particular use with three among the modal sentence adverbs.

2 The triadic model of representation of modal meanings

We assume that three parameters are involved in the interpretation of modal meaning.

P1 The core meaning.

P2 The activation mode of the core meaning: standard or diverted.

P3 The rhetorical nuance of the form related to its diverted use.

2.1 P1: The core meaning

Each form has one core meaning. We assume that a modal form conveys the same core meaning in all uses. For instance, the adverb *peut-être* has two different uses, but in both uses it conveys the same core meaning, which corresponds to an indication of "possibility":

4. Il est peut-être malade, il était très pâle hier
   'He may be sick, (lit. Perhaps he is sick) he was very pale yesterday'

5. Je suis peut-être une femme, mais je sais changer une roue de voiture
   'I may be a woman (lit. Perhaps I am a woman), but I can change a tire'

Only in (4) does *peut-être* convey an epistemic meaning. In (5), the gender of the speaker cannot be doubted, but *peut-être* can nevertheless be used. The two states of affairs in (5) show some incompatibility (to be a woman is seen as incompatible with the capacity of changing a tire) which is resolved by *mais*. In such a use the adverb does not by itself endorse a concessive value which is conventionally triggered by *mais*, but we will see that its functioning is compatible with such a value. According to our analysis, these differences of use exemplified by (4) and (5) depend on the mode in which the core meaning is activated (cf. parameter 2), but not on a difference regarding the core meaning in itself.[2]

Such an assumption is also valid for a verbal modal form such as *pouvoir* or for a temporal modal form such as the future.

We assume that the verb *pouvoir* conveys an indication of "conceivability" in the following different uses:

6. Il peut avoir la grippe, il y a une épidémie actuellement
    'He may have the flu, there's an epidemic going on at the moment'

7. Il peut avoir la grippe, mais je l'ai vu nager comme un poisson à la plage
    'He may have the flu, but I saw him happily swimming at the seaside'

In (6) *pouvoir* can be replaced by *probablement* since it conveys a clear epistemic meaning, whereas this is not the case in (7), where it can commute with the *peut-être* used in (5).

We also assume that the future tense conveys the same indication of "posteriority" in any of the following uses:

8. I Demain, il aura vingt ans
    'Tomorrow he'll be twenty years old'

9. Il est trop tard pour lui téléphoner, à cette heure elle dormira
    'It is too late to call her, she will be sleeping at this hour'

10. Federer sera une fois de plus le meilleur, il vient de remporter la victoire à New York
    'Once again Federer has proven to be the best (lit. will be the best), he's just won the match in New York'

In (9), the future conveys an epistemic stance of the speaker, but not in (10), since the speaker knows the result of the match, nor in (8), where it conveys a genuine temporal meaning.

As previously said, the above differences concerning the interpretation of the form do not depend on the core meaning, but on the mode in which it is activated. There are two modes of activation of the core meaning: a "standard" mode when the form applies to the propositional content of the utterance, or a "diverted" mode when it applies to the occurrence of the utterance, as explained in the next section.

---

[2] For further discussion regarding the complex interaction between the conventional implicature expressed by *mais* and the semantics/pragmatics of concessive structures constituted by the entire sequence *X, mais Y*, see Rossari (2014) as well as Potts (2012), who takes into account multidimensional approaches to the study of meaning.

2.2 P2: The diverted mode of activation of the core meaning

The assumption according to which a core meaning can be activated in a diverted mode is based on the theoretical background of Ducrot (1984), as indicated in Ricci (2017: 19-20), according to which any utterance can be used by the speaker as an object she/he can comment on. Following this idea, our model assumes that each utterance implies a projection of its occurrence. It is on such a projection, i. e. the OCCURRENCE OF THE UTTERANCE (OCC [U]), that the speaker can make a comment, which can be expressed by means of some modal forms. The core meaning of the modal form applies thus to the projected occurrence of the utterance, that is OCC [U]. For instance, the core meaning that can be associated to *peut-être* is POSSIBILITY (POSS). When *peut-être* is used as conveying an epistemic stance 4, its core meaning applies to the propositional content of the utterance: POSS (P). But when it is used as in (5) the core meaning applies to the projected occurrence of the utterance: POSS (OCC [U]). Thus it is not the propositional content (speaker being a woman) that is qualified by *peut-être*, which would make no sense, but it is the projection of the occurrence of the utterance, which the speaker comments on as being possible. She/he utters a content P but at the same time, he/she produces a rhetorical figure (which could be assessed as a "preterition") to communicate that the occurrence of such an utterance is not actual but only possible. Such a rhetorical "trick" allows him/her to distance herself/himself from the endorsement of the propositional content of the utterance that just occurred. The label "preterition" is used in French to designate formulae that allow the speaker to attract the attention of the hearer to certain contents, by pretending not to say them. The formulae used are: *I do not need to say P, I am not competent to say P, I did not say P…* This explains why such a use of *peut-être* is compatible with a polyphonic interpretation, as assumed in Nølke (1993). According to our model, the polyphony is not intrinsic to *peut-être*: it is a consequence of the diverted use of *peut-être*, whose core meaning applies to the OCC [U] instead of applying to the content of the utterance.

Such an analysis is also relevant for modal and non-modal uses of temporal forms. When the future is used as in (8), its core meaning, which is posteriority (POST), is applied to the verbal predicate. But when it is used as in (9) or (10), it is impossible to apply it to the verbal predicate since it refers to a simultaneous or a past event. However, the core meaning does not change. The future is used in a diverted mode. Instead of applying to the content, it applies to the projection of the occurrence of the utterance: POST (OCC [U]). The future indicates that the occurrence of U is postponed. This rhetorical "trick" allows the speaker to distance himself/herself from the endorsement of the content of their utterance, which can be motivated by different causes: in the case of example (9), the motivation for the distancing is the intention of the speaker to show his/her caution. This motivation gives rise to the nuance of conjecture that the future conveys in this example.

2.3 P3: Rhetorical nuances of the diverted mode of activation of the core meaning

Every diverted use of a specific form implies a motivation which determines the specific value of the form. Such a value can be considered as **semi-conventional**. It is **conventional** in the sense that it is attached to a specific form, but it is **semi-** because not all uses of the form convey such a nuance (unlike the first parameter, which determines the core meaning). The

resulting value depends on the contextualization of the form. This parameter is necessary not only to differentiate or bring together forms pertaining to the same language, but also to capture the differences and similarities between equivalent forms in different languages. For instance, we have seen that in (9) and (10), the future conveys very different nuances. In both cases the future is used in a diverted mode, since it cannot be interpreted as applying to the predicate (the time reference being past for both events), but the motivation involving this "trick" is different. In (9), the indication POST (OCC [U]) – which implies that the endorsement of the content P is postponed – is motivated by the speaker showing his/her caution. This confers a conjectural value to the future. In 10, the same indication is motivated by the speaker showing advanced knowledge of events. Although she/he knows at the moment of the enunciation that Federer has won, she/he indicates that the endorsement of its content is postponed, making it understood as a prediction in hindsight. This confers a value of retrospective prediction to the future. Such a value is not shared by the Italian or German future as shown in Ricci (2017) as well as in Hütsch (2016). In both languages, the future in (10) is understood as conveying a conjectural value. In contrast, in Italian, the future can convey a value which is not transmitted by the future in French, but by the modal sentence adverb *peut-être* as illustrated in (5) or by the modal verb *pouvoir* as shown in (7):

11. Sarò anche una donna, ma posso cambiare una gomma
    'I may be a woman (lit. Perhaps I will be a woman), but I can change a tire'

In the latter example, the nuance conveyed by the future in Italian is equivalent to the one conveyed by *pouvoir* or *peut-être* in French. This is due to the fact that according to the third parameter (P3) the motivation of the speaker is the same: he/she shows that P is not relevant. Such a motivation explains the diverted use of the form, allowing him/her not to endorse the content.

To summarize, the core meaning of each of these three forms suspends the endorsement of the content: the future by indicating that it is postponed, the modal sentence adverb *peut-être* by indicating that it is potential, and the modal verb *pouvoir* by indicating that it is conceivable (hence not done). The resulting value conveyed by each of these three forms is the same: P is in the background of the current discourse.

Future: POST (OCC [U])

Peut-être: POSS (OCC [U])

Pouvoir: CONC (OCC [U])

When *pouvoir* conveys an epistemic value as in (6), its value is the result of a standard use of the form, whose core meaning applies to the propositional content: the form indicates that the predicate "have the flu" is conceivable.[3]

---

[3] We prefer to associate to *pouvoir* the core meaning 'conceivable' instead of 'possibility' or 'abstract possibility' which is proposed by most authors: for instance in their monography Barbet (2014) and Mari (2015) use 'possibility'; de Saussure (2012: 139), Vetters (2004: 662), and Le Querler (2001: 21) use 'abstract possibility'. Without going into more detail about the reasons for this choice (since it is not the topic of this article), we believe that the latter choice is more appropriate to reflect the values of capacity/permission that are specific to *pouvoir* in examples such as: *Marie peut être enceinte* (understood as *Marie has no sterility problems*).

To sum up, an epistemic meaning can be the result of a standard use (as in the case of *peut-être* or *pouvoir* indicating that an event is possible/conceivable) or of a diverted use (as in the case of the future indicating that the content of the utterance is not endorsed, since OCC [U] is postponed). This difference in the mode in which an epistemic interpretation occurs allows to predict that the utterance does not have the same properties. A conjecture made with the future cannot be cancelled (as noted in Ricci 2017: 46), whereas it can be cancelled when it is made with *peut-être* or *pouvoir*:

12. Marie peut être enceinte, mais je n'en suis pas sûr
    'Marie may be pregnant, but I'm not sure'

13. Marie est peut-être enceinte, mais je n'en suis pas sûr
    'Marie may be pregnant (lit. Maybe Marie is pregnant), but I'm not sure'

14. Marie sera enceinte, ??mais je n'en suis pas sûr
    'Marie may be pregnant (lit. Marie will be pregnant), but I'm not sure'

The reason is that, with the future, the occurrence of utterance is presented as postponed. So, its endorsement is understood as something that will happen. Such an indication is not compatible with a subsequent content that cancels the former. With *pouvoir* or *peut-être*, the event is presented as a possibility or as something conceivable, which can be weakened in a subsequent utterance.

2.4   Synthesis

Three parameters are attached to any modal form; the application of these three parameters produces an outcome giving the semantic value that the form eventually conveys. When, according to P2, the form has a standard use, the value depends on its contextualization. There are no particular rhetorical motivations, the use of the form not being diverted. But when the form has a diverted use, this use is motivated by a rhetorical purpose, called "motivation". This motivation is attached to a specific form. It gives rise to a nuance that is considered as semi-conventional.

Below we give a synthesis of the indications related to any of the three parameters for the three forms examined.

Peut-être

P1 POSS

P2 standard use→ outcome: possible event

P2 diverted use → POSS (OCC [U]) implies content of U is not endorsed

P3 motivation: showing P non relevant →outcome: P is in the background of the current discourse

Pouvoir

P1 CONC

P2 standard use→ outcome: realizable event

P2   diverted use →CONC (OCC [U]) implies content of U is not endorsed

P3   motivation: showing P non relevant →outcome: P is in the background of the current discourse

Future

P1   POST

P2   standard use → outcome: forthcoming event with respect to the moment of the utterance
P2 diverted use→POST (OCC [U]) implies content of U is not endorsed

P3   motivation: showing caution →outcome: conjectural value

P3   motivation: showing advance knowledge of events → outcome: retrospective prediction

P3   motivation (only for Italian): showing P non relevant →outcome: P is in the background of the current discourse

According to this model, the modal sentence adverbs have the following properties:

Their core meaning applies to the content of the utterance. Some of them have a diverted use that allows them to occur with events which are not compatible with an epistemic reading. In that case, their core meaning applies to the occurrence of the utterance giving rise to one particular rhetorical value (indicating that P is in the background of the current discourse). The speaker usually utters another utterance which is interpreted as the foreground of his/her discourse, generally introduced by *mais* (but). Therefore the entire sequence corresponds to a concessive structure.

15. Je suis peut-être une femme, mais…
    'I'm a woman, but…'

16. Tu peux être ma mère, mais j'ai dix-huit ans et je fais ce que je veux!
    'You may be my mother, but I'm eighteen and I do what I want'

17. Sarò una donna, ma…
    'I may be a woman, but…'

2.5 The diverted use of modal sentence adverbs

Among the modal sentence adverbs, only some can have a diverted use compatible with their occurrences in non-epistemic contexts. The selection of such adverbs is based on their occurrence with predicates expressing inalienable properties (such as nationality, gender…).

18. A propos des attentats de Paris. Je suis **peut-être** suisse, mais je soutiens totalement la France (Internet)
    'About the attacks in Paris. I may be Swiss, but I fully support France'

19. Je suis **certes** une femme, mais je m'impose. (Internet)
    'I may be a woman, but I impose myself'

20. Je suis **sans doute** une femme mûre, mais je m'entretiens et je suis fière de mon corps.
    'I may be a mature woman, but I look after myself and I'm proud of my body.' (Internet)
    > Rossari and Salsmann (2017)

In addition to *peut-être*, *certes* and *sans doute* are found in such contexts, with some peculiarities for *sans doute*: in the context given by (20), the axiological evaluation given by *mûre* introduces ambiguity allowing for a possible access to an epistemic interpretation (i. e. the speaker can judge herself as *mûre*), even if the latter interpretation is not the most accessible (since the evaluation is made by the speaker on herself). The same example in the third person gives rise to an epistemic interpretation (compatible with any adverb of the paradigm):

21. C'est
    assurément/certainement/indubitablement/probablement/sûrement/vraisemblablement une femme mûre, mais elle s'entretient

Regarding the other modal sentence adverbs listed in the introduction, their non-acceptability in genuine non-epistemic contexts such as (18) and (19) is clear-cut:

22. ??Je suis
    assurément/certainement/indubitablement/probablement/sûrement/vraisemblablement suisse, mais je soutiens totalement la France

23. ??Je suis
    assurément/certainement/indubitablement/probablement/sûrement/vraisemblablement une femme, mais je m'impose

The three adverbs compatible with non-epistemic contexts (*peut-être*, *certes*, and *sans doute*) do not present particular traits that differentiate them from the others.

Firstly, they cannot be particularized in relation to the degree of certainty they express, since they convey different grades of certitude: *sans doute* and *certes* represent strong certainty and *peut-être* weak certainty:

24. Tu m'as convaincue: tu as sans doute/certes[4] raison
    'I'm convinced: you are certainly right'

25. ??Tu m'as convaincue: tu as peut-être raison
    'I'm convinced: you may be right'

Secondly, they cannot be particularized in relation to their dialogical uses, which can be a condition to be compatible with concessive contexts, as assessed in Rossari 2016, since all the modal sentence adverbs can be used to answer a yes/no question, as observed in many studies: Borillo (1976: 87), Molinier (1990: 32–33), Guimier (1996: 110), Rossari et al. (2004: 15).

26. Est-ce que Paul est prêt?
    'Is Paul ready?'
    Assurément./ Certainement./ Indubitablement./ Probablement./ Sûrement./ Vraisemblablement./ Peut-être./ Sans doute./ Certes.

---

[4] Even if *certes* conveys a concessive value in contemporary French, it is still possible to understand it as conveying an epistemic evaluation.

Therefore, isolating these adverbs on the basis of a common property that may not be shared by the other adverbs of the paradigm proves to be particularly tricky. Our quantitative analysis will give us some clues to particularize them.

3   The indications given by the quantitative approach

The purpose of the quantitative analysis is to see if there are some indications that allow us to understand why only some of these modal sentence adverbs have this non-epistemic use. We take into account the three adverbs that have the capacity to occur in non-epistemic contexts (*peut-être*, *sans doute*, and *certes*) and one of those that do not have this capacity (*probablement*). We have selected it for its high frequency (it is the most frequent adverb in the press corpus consulted,[5] after *peut-être*, *sans doute*, and *certes*, with the frequencies following this order in both corpora (cf. Rossari and Salsmann 2017).

We have focused on how these adverbs are used rhetorically. We would like to know if the utterance following the one where the adverb occurs is used to support (with connectives such as *car, parce que*) or to add reservations to the former (with *mais*) and, more precisely, when the adverb is used in a sequence with *mais*, to look at its position: before *mais*, it underlines the background role of the utterance and, after *mais*, its foreground role. Therefore, we have identified the connective collocates taking the adverb as pivot (search term) based on the log-likelihood value (LL) in a span of 10 and 20 tokens in the context to the right of the pivot. We present the results extracted from the press corpora *Le Monde* 2008 (20 410 766 tokens) available on the platform BTLC (http://persan.rom.uni-koeln.de/btlsc) (cf. Diwersy (2013)). Using the query for co-occurrence analysis, we have obtained the lists of specific collocates for each pivot. The specificity of these associations between items is calculated by a statistic measure of significance. This measure compares the observed and estimated frequencies of co-occurrence. Applying a threshold value of 10.83, all co-occurrences are considered to be specific if their log-likelihood values are above it, indicating less than a 0.01-percent chance that the co-occurrences are random.

Taking into account these two spans allows us to extract connectives which have a high probability to be linked to the utterance where the adverb occurs. We present below the lexicograms of each adverb for both spans extracted by the platform BTLC from the corpus *Le Monde* 2008.
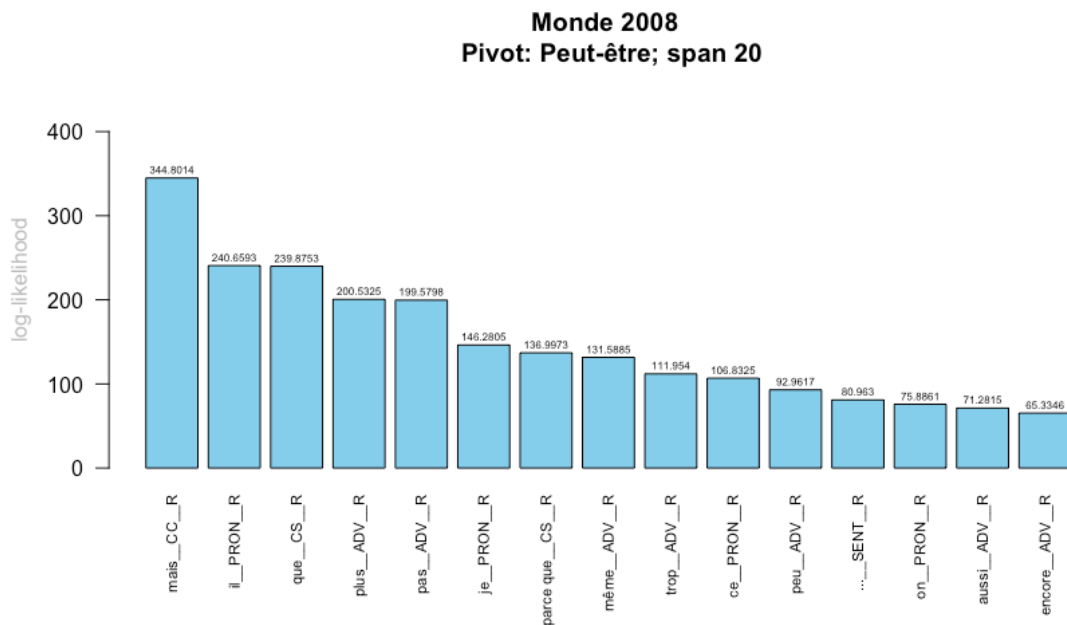
---

[5] We only present the results extracted from Le Monde (2008), but in addition to this corpus, we have gathered the results extracted from Le Figaro and Sud-Ouest, to have a variety of press discourses. They were presented at the conference referred under Rossari and Salsmann (2017). The results being very similar for the three newspapers, we present here only the lexicograms of Le Monde 2008.
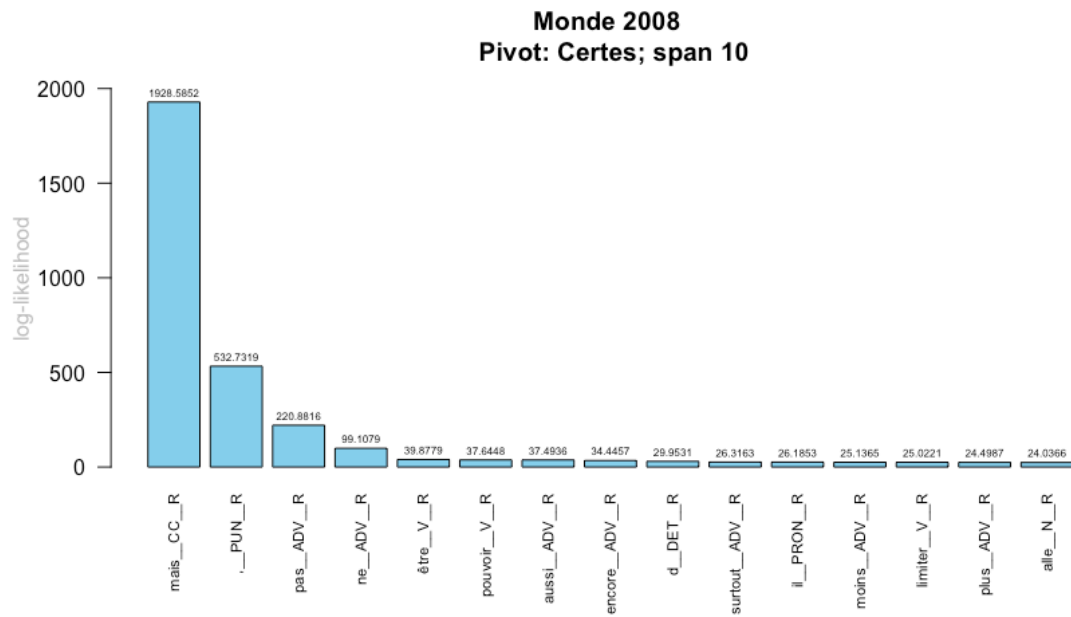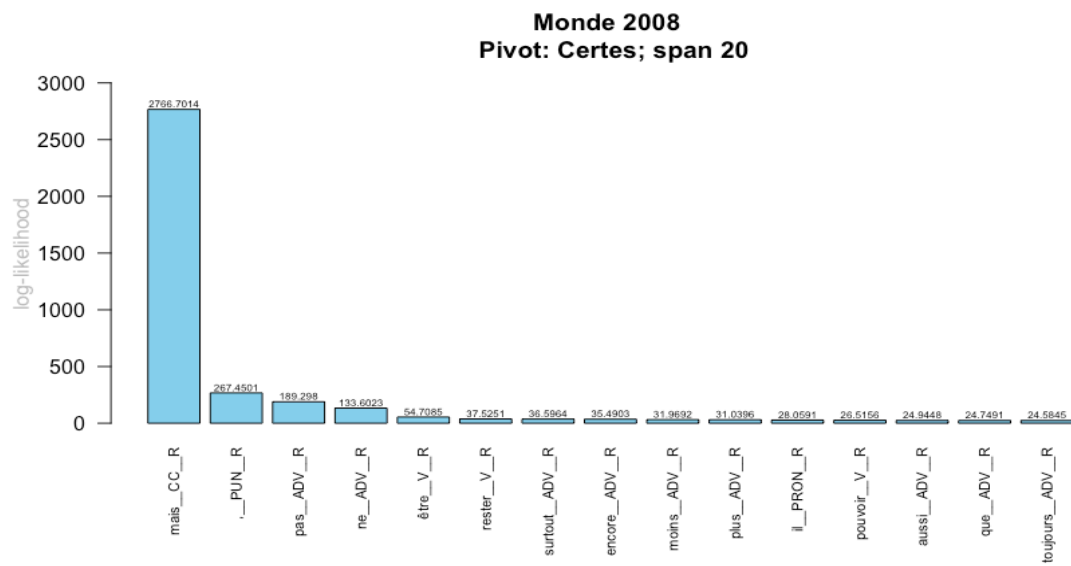
3.1 Press corpus

Span 10

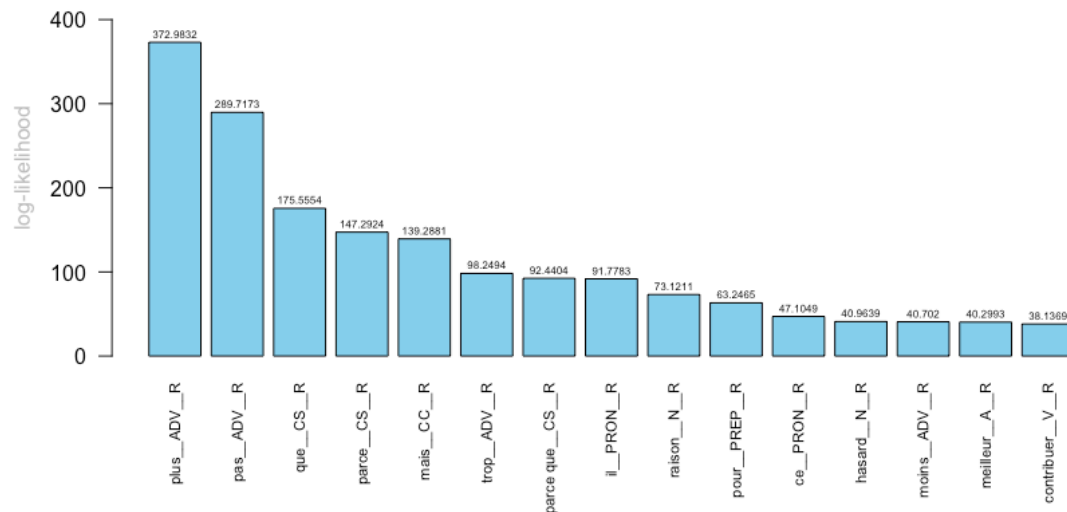**Monde 2008**
Pivot: Peut-être; span 10

Span 20

Monde 2008
Pivot: Peut-être; span 20

Span 10

Monde 2008
Pivot: Certes; span 10
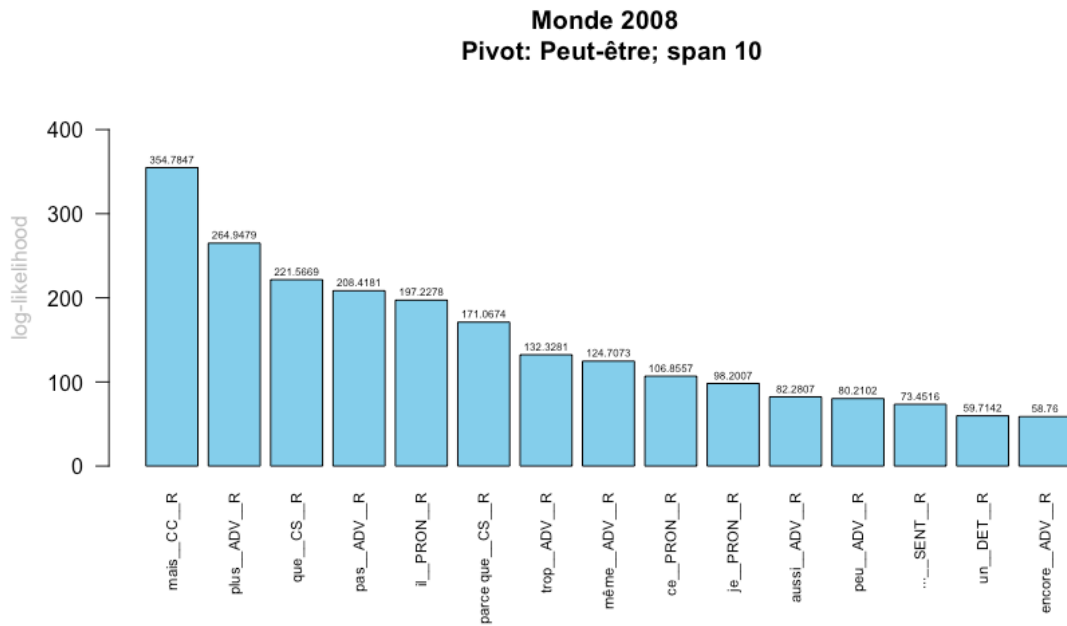
Span 20

Monde 2008
Pivot: Certes; span 20

Span 10

Monde 2008
Pivot: Sans doute; span 10

Span 20

Monde 2008
Pivot: Sans doute; span 20

Span 10
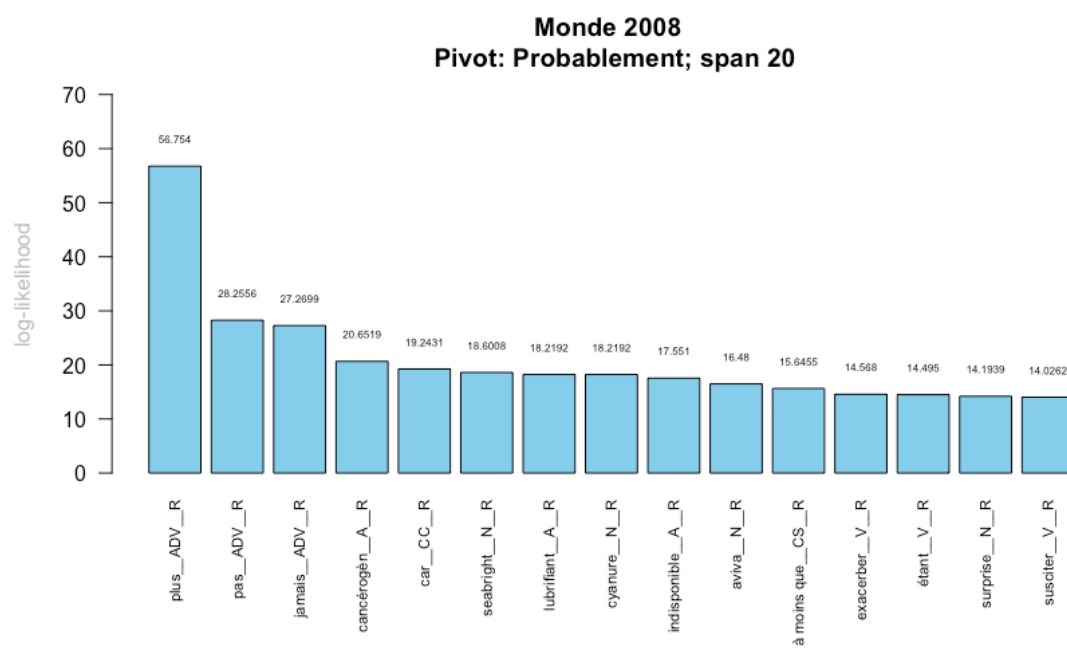
Monde 2008
Pivot: Probablement; span 10

Span 20
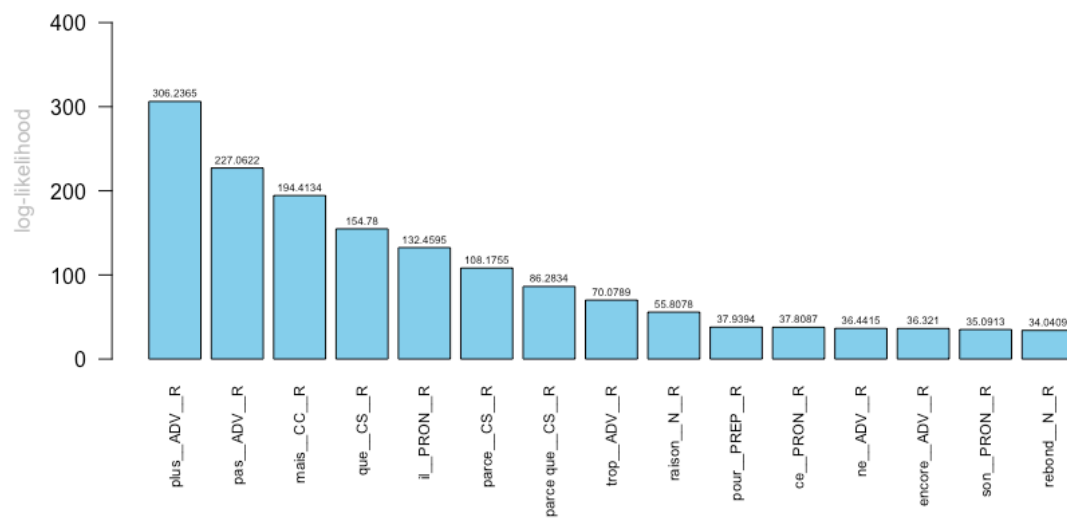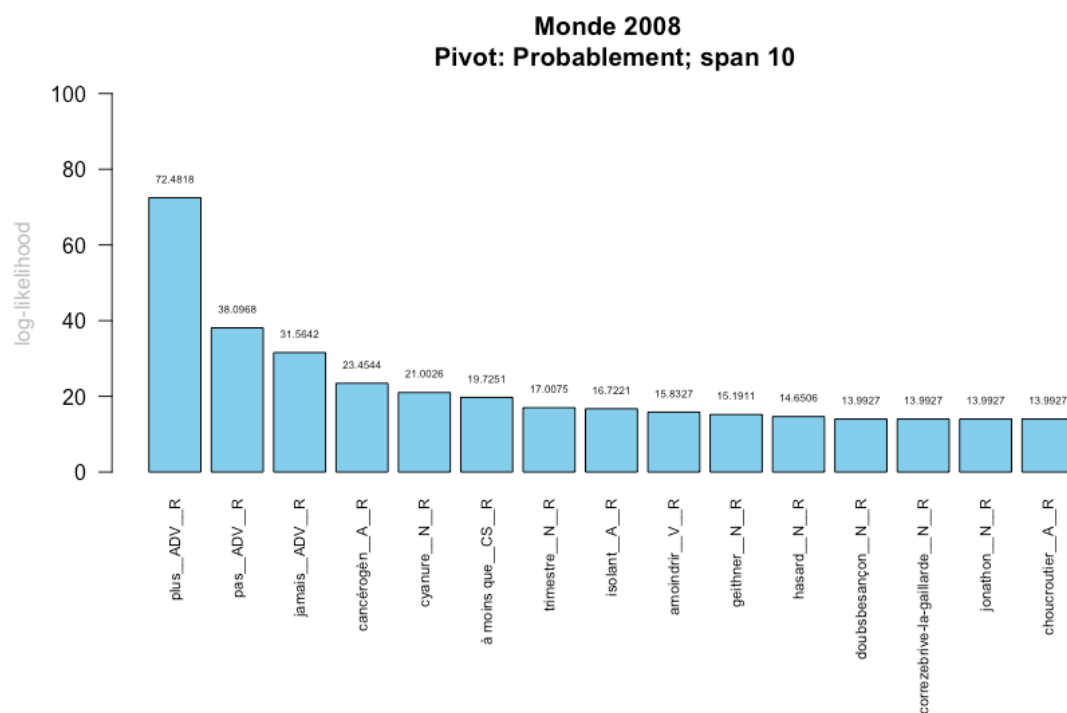
Monde 2008
Pivot: Probablement; span 20

3.2   Interpretation of the quantitative data

The results for these first data are clear. The connective *mais* is the first right collocate as a connective for *peut-être*, *certes*, and *sans doute* in both spans. In contrast, *mais* does not appear among the results for *probablement*. Its first right collocate as a connective is *car* in

span 20. The similar behavior of these three adverbs in comparison with *probablement* confirms the qualitative analysis, which shows that these three adverbs share the property of being used in non-epistemic contexts usually inserted in concessive sequences (sequences where the utterance with the adverb is followed by another utterance introduced by *mais*). Since *mais* indicates that the utterance introduced should be interpreted as foreground (cf. Rossari 2016), the common use of these adverbs before *mais* can be considered as facilitating the understanding of the previous utterance as background, which corresponds to the motivation of their diverted use.

To deepen the interpretation of these data, we can focus on the strength of the attraction between the adverb and *mais* depending on whether the conjunction is positioned to the left or right of the adverb. We make the assumption that the greater the force of attraction with *mais* on the right, the more the adverb is able to occur in non-epistemic contexts, facilitating access to a background interpretation of the sequence. We take into account a span of 10 items for both left and right position of the adverb, which better ensures the semantic link between the adverb and the connective. We present the results for *Le Monde* and refer to Rossari and Salsmann (2017) for the additional press corpus (*Le Figaro* and *Sud-Ouest* show similar results).
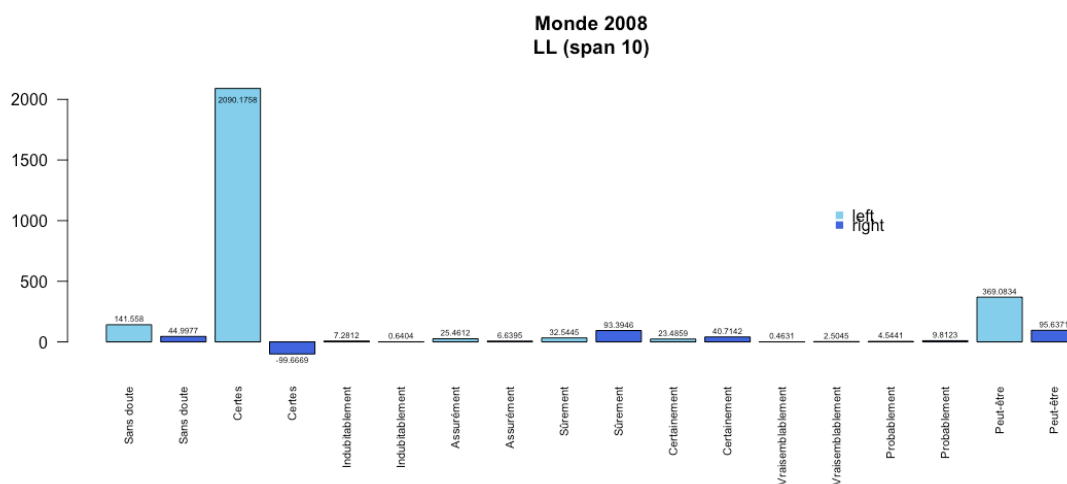
The difference for *peut-être*, *sans doute*, and *certes* between right or left collocates of *mais* is clear cut. In addition, we can note that, in the case of *certes*, there is a dis-attraction between the latter and left *mais*, which can be interpreted as the effect of the conventionalization of the concessive value with *certes* (cf. Rossari et al. 2016). Even if for *sans doute* and *peut-être* the left *mais* can still be considered as specific (the value is over 10.83), the difference between the two positions of *mais* is significant (141 vs 44 for *sans doute* and 369 vs. 96 for *peut-être*). For the other adverbs, the score is either in favor of left *mais* (*sûrement*, *certainement*) or the difference between left and right *mais* is non-significant.

These quantitative studies show results converging with the qualitative analysis. Adverbs that share the property of being able to be used in non-epistemic contexts also share the property of having the connective *mais* as a right collocate. The strength of the attraction (greater for *certes* and *peut-être* than for *sans doute*) is congruent with the fact that the non-epistemic

context in which *sans doute* can occur should be more ambiguous than that for *peut-être* and *certes*, as can be seen in example (20) or in the following example:

27. Je suis sans doute trop gâté, c'est un fait.
    'I may be too spoiled, it is a fact'

The ambiguity is due to the addition of "it is a fact" that makes the epistemic judgment expressed by *sans doute* disappear.

2.2.1 *Encyclopedia Universalis* corpus

However, such clear results are not observable in any corpus. If we take into account an encyclopedic corpus, the outlines are not entirely convergent with the qualitative analysis. We present the results extracted from *Encyclopedia Universalis* 2005 (49 859 864 tokens) available on BTLC concerning the attraction with right or left *mais*.
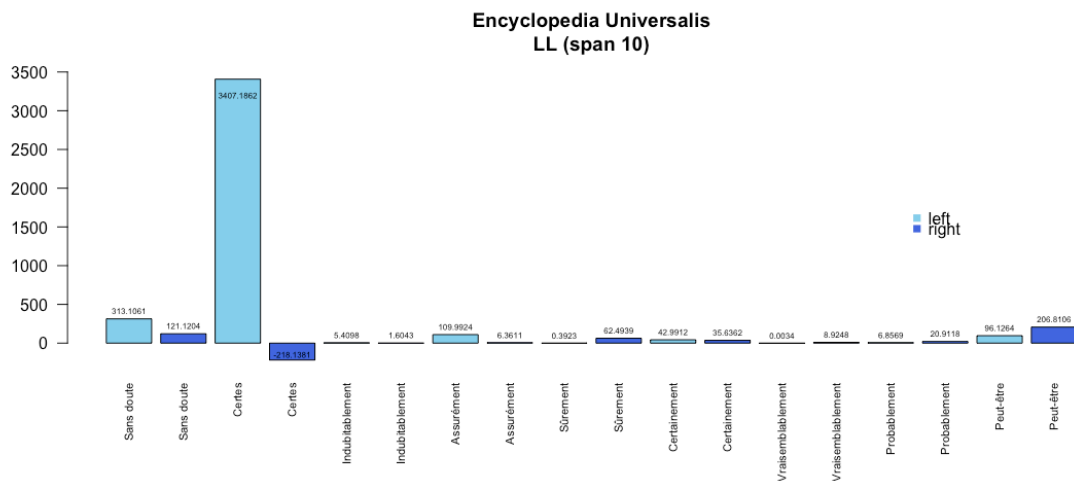

The use of *peut-être* is different. Its attraction with *mais* is inverted in comparison with the corpus *Le Monde*. It is stronger with left *mais* than with right *mais*. With regard to *sans doute* and *certes* the results are similar in both corpora. So the attraction between one adverb and a particular connective seems to depend on the genre. However, it is not totally unpredictable that the attraction of *peut-être* with right *mais* is better represented in press discourse than in encyclopedic discourse. The former represents a more colloquial style than the latter and concessive uses of *peut-être* (which are represented by the sequence with right *mais*) fit better with such a style.

2.2.2 *Encyclopedia of Diderot and d'Alembert* corpus

Taking another diachronic period into account (the 18th century) with the *Encyclopedia of Diderot and d'Alembert* (DDA)[6] shows a different picture of the relation between the adverb and *mais*.

---

[6] The digital edition of Diderot and d'Alembert (23 940 181 tokens) available in BTLC has been established on the basis of the 1751-1772 edition which was realized by ARTFL under the direction of Robert Morrissey. We thank him for giving us access to it.

Encyclopedia D.&D'A.
LL (span 10)
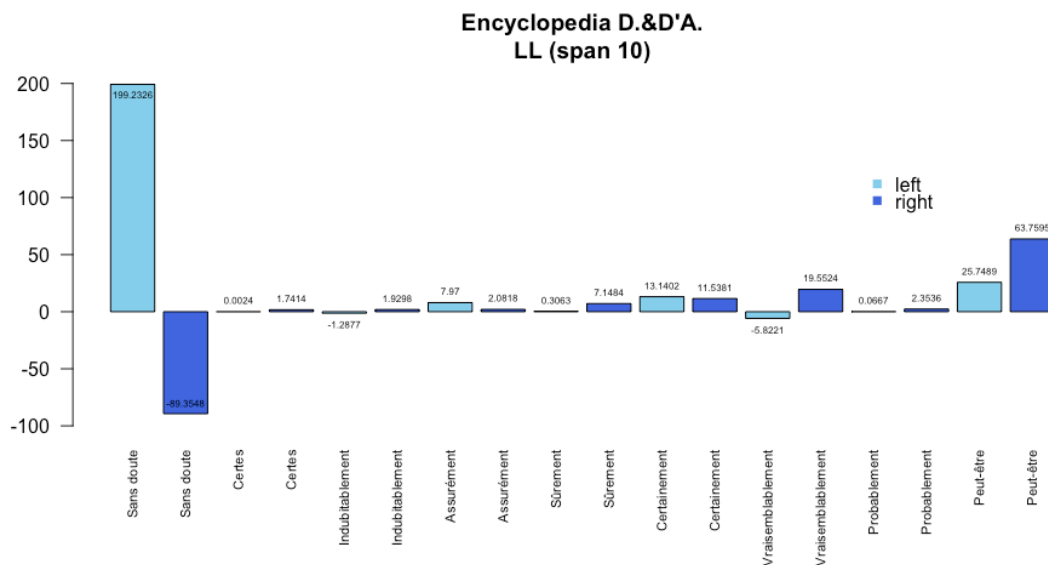
The results are different from those observed in the corpus *Le Monde* 2008. The only clear attraction between a modal adverb and *mais* as right collocate is with *sans doute*. The other adverbs showing strong attraction in *Le Monde* 2008 (*certes* and *peut-être*) do not show any particular attraction with right *mais*: for *peut-être* both right and left *mais* show attraction and for *certes* there is even a dis-attraction for right *mais* whereas we saw the contrary in the press corpus.

Therefore, the only attraction observed in the three corpora is the one between *sans doute* and right *mais*, but the strength of this attraction is weaker in the 21st corpora. There is no indication of attraction between *peut-être* and right *mais* and, for *certes*, the attraction with right *mais* seems to point at a sudden change (i. e. non-gradual) since the log-likelihood scores show dis-attraction in the 18th corpus and the highest score in the 21st corpus.

The attraction of one adverb with *mais* as right collocate is not only a question of discourse genre but also a question of historical period. The fact remains that the results obtained in the press corpora allow us to consider the connective *mais* as shaping the value of the adverb making it compatible with "new" contexts such as the non-epistemic ones.

4 The degree of attraction of *pouvoir* with *mais*

To challenge the model and assess its predictability concerning the properties of the modal sentence adverbs having non-epistemic uses, we conducted the same quantitative queries for the modal verb *pouvoir* since it shares with these adverbs the possibility to be used in non-epistemic contexts, cf. (7) or (16).

We measure the degree of attraction of *pouvoir* with *mais* as right or left collocate, in a first test without specifying a particular tense, and, in the second test, by conjugating *pouvoir* in the present tense only, since this tense seems much more natural in non-epistemic contexts:

28. ??Il pouvait/a pu avoir la grippe, mais je l'ai vu nager comme un poisson
   'He may have had the flu, but I saw him happily swimming'

This allows a close comparison with the data obtained for the paradigm of epistemic adverbs. Our test concerns *pouvoir* and *devoir*, since the latter, like *probablement* for instance, cannot be used in non-epistemic contexts:

29. ??Il pouvait/a pu avoir la grippe, mais je l'ai vu nager comme un poisson
    'He may have had the flu, but I saw him happily swimming'

30. ??Il a probablement la grippe, mais je l'ai vu nager comme un poisson à la plage
    'He probably has the flu, but I saw him happily swimming'

31. ?? Tu dois être ma mère, mais j'ai dix-huit ans et je fais ce que je veux!
    'You must be my mother, but I'm eighteen and I do what I want'

32. ??Tu es probablement ma mère, mais j'ai dix-huit ans et je fais ce que je veux!
    'You are probably my mother, but I'm eighteen and I do what I want'

One could expect that the results for *pouvoir* and *devoir* in the present tense show similar differences to those that appear for *probablement* vs. *certes, peut-être,* or *sans doute.* We limited our research to the corpus in which the difference between the adverbs is the clearest, i. e. the corpus *Le Monde* 2008.

***Pouvoir/devoir* all tenses**
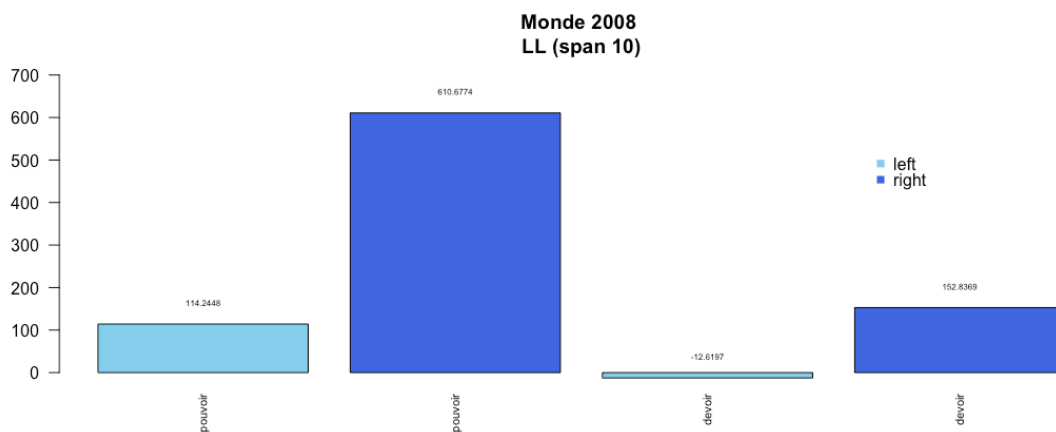
***Pouvoir/devoir* present tense**
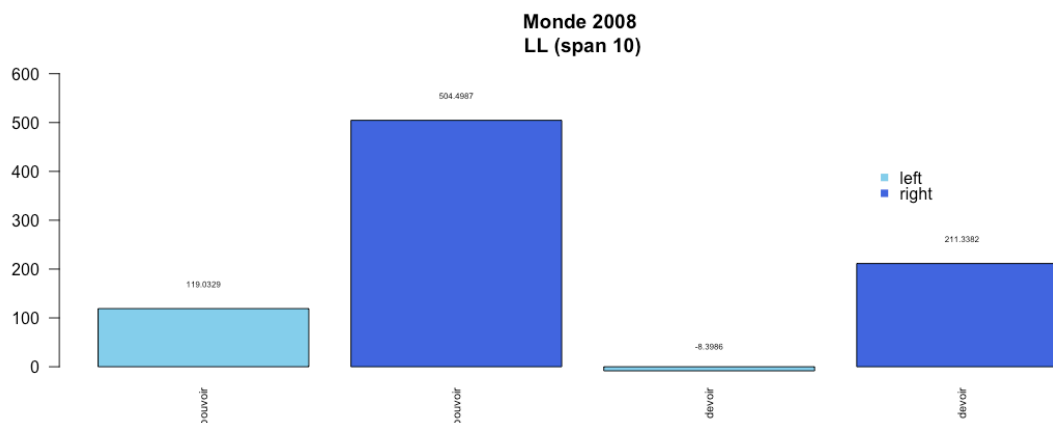
Our expectations are partially met. On the one hand, the behavior of the modal *pouvoir* is very different from that of *peut-être*, *certes*, and *sans doute*, since the latter are much more attracted by the left *mais* than by the right *mais*. *Pouvoir* on the contrary shows much more attraction for the right *mais*. But, on the other hand, there are three aspects that met our expectations.

(i) Even if it is lower than with right *mais*, the attraction with left *mais* is still positive (well above 10.83) for *pouvoir*.

(ii) For *devoir* the association *devoir…mais* is instead negative.

(iii) The present tense shows a slightly higher attraction for *pouvoir* with the right *mais* than when it is used in any tense (LL 119 vs. 114).

The fact remains that the high attraction of *pouvoir* with left *mais* weakens the relevance of the association of "modal item… *mais*" as an indication of the non-epistemic use. Extensive research is required in order to be able to differentiate the values of *pouvoir* in relation to the right or left *mais* while at the same time looking at the proportion of *pouvoir* used in a diverted mode.

## 5. Conclusions

5.1. Synthesis

We have first presented a model that aims to describe how a modal value of a particular form is transmitted. We have seen that such a model allows to differentiate epistemic meaning transmitted by modal sentence adverbs or *pouvoir* and epistemic meaning transmitted by the future tense. The former is conveyed by the standard functioning of the form (parameter 2) whereas the latter is conveyed by a diverted functioning of the form, whose core meaning applies to the projection of the occurrence of the utterance (OCC [U]). Furthermore, the modal sentence adverbs can also have a diverted functioning. In that case, they can occur in non-epistemic contexts (like the future tense in Italian or *pouvoir*). Those contexts are usually

interpreted as concessive, since they can easily be followed by a sequence introduced with *mais*. Only some of the modal sentence adverbs (*certes*, *peut-être*, and *sans doute*) can have this functioning. The purpose of the quantitative section has been to find regularities that particularize the functioning of these three adverbs. We have highlighted their attraction with right *mais* in the press corpus. But we cannot generalize our results for these three adverbs to corpora representing other discourse genres or another diachronic span. We have observed that in the encyclopedic corpus of the 21st *peut-être* is more attracted by left *mais* than right *mais*, and that in the *Encyclopedia of Diderot and D'Alembert* only *sans doute* is attracted by right *mais*. *Certes* is neither attracted by left nor by right *mais* and *peut-être* is more attracted by left *mais* than by right *mais*. Finally, we have compared these results with those concerning *pouvoir* and *devoir* in the press corpus. This comparison reveals that, like *probablement*, *devoir* is not attracted by right *mais*, but with regard to *pouvoir* and *peut-être* the comparison shows differences, since *pouvoir* is more attracted by left *mais* than by right *mais*. However, the LL value of *pouvoir* with right *mais* is well above 10.83 which still indicates an attraction to right *mais*.

5.2.   Broader perspectives

The quantitative research has put forward some data that merits further research. It reveals some similarities between the behavior of *peut-être* in Universalis and DDA and *pouvoir* in the press corpus. In these corpora, both forms show more attraction with left *mais* than with right *mais*, bringing out a similarity between *peut-être* and *pouvoir*. This could be an indication that the persistence principle of grammaticalization is still at stake for *peut-être*. The verb *pouvoir* would maintain some of its traits in *peut-être*, making the adverb behave similarly to *pouvoir* in less colloquial corpora (such as Universalis) or in a previous diachronic span (such as DDA). The results for *pouvoir* in both corpora shed light on this assumption:

**Diderot et d'Alembert
LL (span 10)**
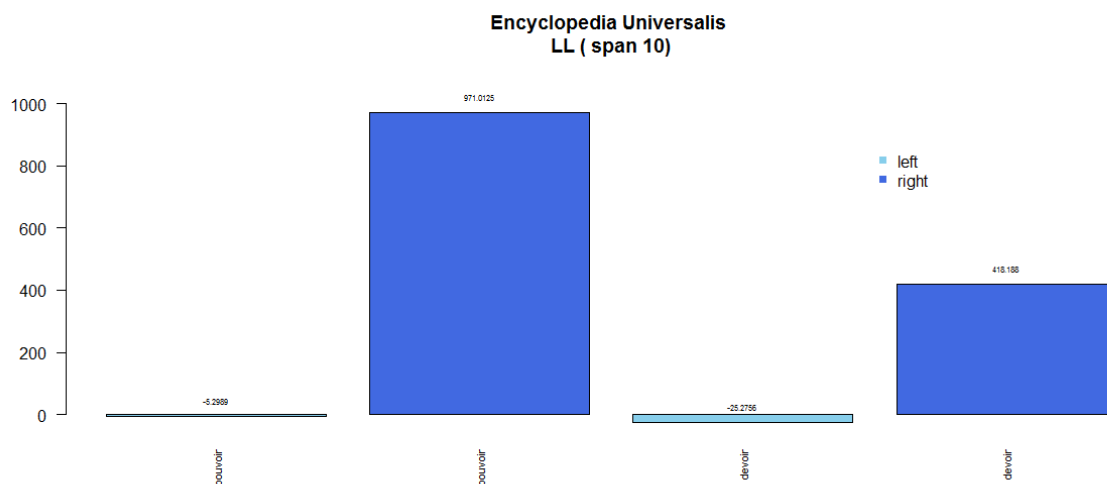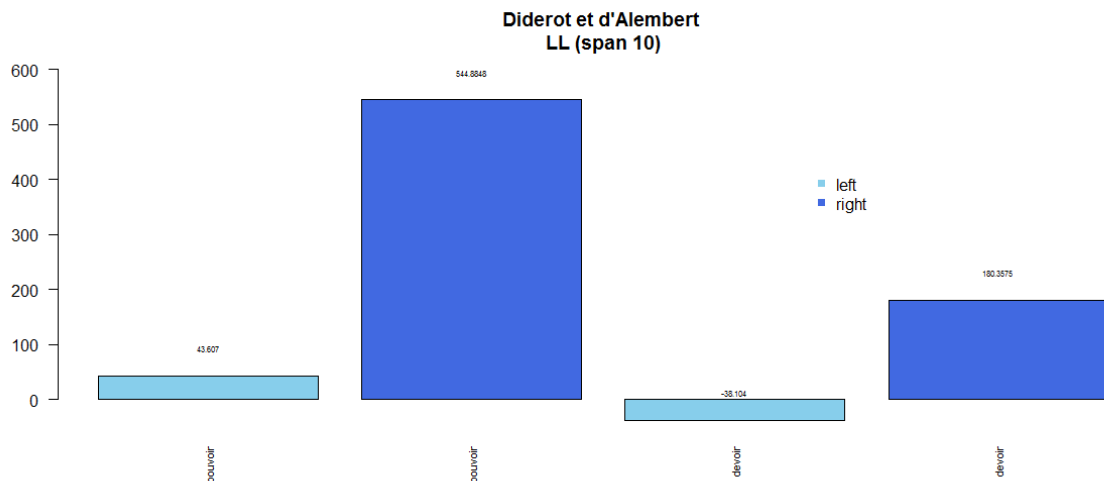
In both corpora, *pouvoir* shows a clear attraction to the right *mais* which is also at stake for *peut-être* in the same corpora. So, there are elements to consider that the balance between right and left *mais*, which we can observe for *peut-être* according to the three corpora selected, is the consequence of the persistence principle still upheld for *peut-être*. If this assumption holds, it is less surprising that in press corpora, which represent a less conservative mode of expression, the balance is inverted for *peut-être* (with attraction for the sequence *Adv…mais* instead of *mais…Adv*).

For further research, we could assume that the modal *pouvoir* has the potential to be used to indicate that the utterance in which it occurs is in the background (like the three modal sentence adverbs illustrated above) and thus attracted by left *mais*, but such a potential is linked to a grammaticalization process which is initiated for *peut-être* and not for *pouvoir*.

References

Barbet, Cécile (2014): *Sémantique et pragmatique des verbes modaux du français. Données synchroniques, diachroniques et expérimentales*. Unpublished PhD thesis, University of Neuchâtel.

De Saussure, Louis (2012): « Modalité épistémique, évidentialité et dépendance contextuelle ». In: *Langue française* 173: 131–143.

Diwersy, Sascha (2013): 'BTLC.Primestat'. <http://persan.rom.uni-koeln.de/btlsc/>

Guimier, Claude (1996): *Les adverbes du français. Le cas des adverbes en* -ment. Paris: Ophrys.

Le Querler, Nicole (2001): « La place du verbe modal *pouvoir* dans une typologie des modalités ». In: *Cahiers Chronos* 8: 33–46.

Hütsch, Annalena (2016): « Les valeurs modales du futur en français et en allemand ». presented at *Temps, aspect, modalité, évidentialité : perspectives comparative, cognitive, théorique, appliquée* (TAM-E). Paris.

Mari, Alda (2017): *Modalités et Temps: des modèles aux données*. Bern: Lang.

Molinier, Christian (1990): « Une classification des adverbes en –ment ». In: *Langue française* 88: 28–40.

Nølke, Henning (1993): « *Peut-être*: fonctions modale et discursive ». In: H. Nølke, *Le regard du locuteur. Pour une linguistique des traces énonciatives*. Paris, Kimé: 173–180.

Potts, Christopher (2012): "Conventional implicature and expressive content". In: Maienborn, Claudia/ von Heusinger, Klaus/ Portner, Paul (eds.): *Semantics: An International Handbook of Natural Language Meaning*. Volume 3. Berlin, Mouton de Gruyter: 2516–2536.

Ricci, Claudia (2017): *Les emplois modaux du futur et de l'imparfait. Analyse contrastive italien-français avec un regard sur la diachronie*. Thèse de doctorat de l'université de Neuchâtel.

Rossari, Corinne (2014): "How does a concessive value emerge?". In: Ghezzi, Chiara/Molinelli, Piera (eds.): *Discourse and Pragmatic Markers from Latin to the Romance Languages*. Oxford Studies in Diachronic and Historical Linguistics. Oxford, Oxford University Press: 237–259.

Rossari, Corinne et al. (2004): *Autour des connecteurs. Réflexions sur l'énonciation et la portée*. Bern: Lang.

Rossari, Corinne/Salsmann, Margot (2017): « Étude quantitative des propriétés dialogiques des adverbes épistémiques ». In: *Journées Internationales de Linguistique de Corpus (JLC2017)*. https://jlc2017.univ-grenoble-alpes.fr/Contenu/LivretJLC2017.pdf.

Rossari, Corinne/Hütsch, Annalena/Ricci, Claudia/Salsmann, Margot/Wandel, Dennis (2016): « Le pouvoir attracteur de mais sur le paradigme des adverbes épistémiques: du quantitatif au qualitatif ». In: *Actes de Colloques JADT 13èmes Journées internationales d'Analyse statistique des Données Textuelles*. Nice.

Vetters, Carl (2004): « Les verbes modaux *pouvoir* et *devoir* en français ». In: *Langues et littératures modernes – Moderne taal en litterkunde, Revue belge de philologie et d'histoire* vol. 82/3: 657–671.