Discourse properties of French clitic left dislocated NPs and their effect on prosody

Lisa Brunetti (Université Paris Diderot/CNRS)

and Mathieu Avanzi (Université Catholique de Louvain)

ABSTRACT. The present paper investigates the relationship between the variation in prosodic boundary strength at the right edge of a clitic left dislocated (CILDed) NP in French and its discourse properties. A CILDed element in French is known to be followed by a major prosodic break; however, scholars have noticed that the degree of boundary strength can greatly vary. There is an ongoing debate on what factors are responsible for this variation. The goal of our paper is to provide a quantitative as well as a qualitative analysis based on a pragmatic and a prosodic annotation of utterances collected from two spontaneous spoken corpora. In the pragmatic annotation, each CILDed constituent of our dataset is annotated according to its topic continuity, its discourse status, and its contrastive function. In the prosodic annotation, a score indicating the degree of prosodic boundary strength at the right edge of each constituent is assigned in a semi-automatic fashion, on the basis of the computation of 4 acoustic features. The main result of the interaction of the two annotations is that the boundary strength is significantly higher in dislocated elements that are discourse given but are not (not yet, or not anymore) active topics, namely, referents that need to be promoted to topic status. An effect of contrast is also found in a subset of the data.

1. INTRODUCTION

28 The present work is a quantitative and qualitative analysis of sentence topics

29 denoted by a clitic left dislocated NP in French. The dataset was obtained by

30 extracting utterances from a corpus of spontaneous speech. The work seeks to

31 contribute to the long-standing debate concerning the causes of the prosodic

32 variation that is observed at the right edge of dislocated phrases in French. Our

33 goal is to see whether certain discourse properties (topic continuity, discourse

34 status, contrast – by itself or in combination) play a role in determining this

35 variation.

36 After describing the syntactic, prosodic, and pragmatic properties of French

37 clitic left dislocated NPs (Section 1.1), we present the state of the art concerning

38 the causes of prosodic variation in this construction, as well as the way we intend

39 to tackle this issue in the present paper (Section 1.2).

40 1.1. CLITIC LEFT DISLOCATED NPs IN FRENCH

41 A clitic left dislocation (from now on, CILD) is a construction that is typical of

42 spoken French (Bally 1944, Ashby 1988, Lambrecht 1994). This construction is

43 characterized by the presence of a constituent in the left periphery of the clause

44 (Fradin 1990, Delais-Roussarie et al. 2004, De Cat 2007), where by 'left

45 periphery' we mean the position preceding the verb and (if the peripheral element

46 itself is not the subject) the preverbal subject (Rizzi 1997). Such a peripheral

47 constituent is considered as external to the clause and co-refers with a clitic

48 pronoun inside it.[1] Examples of clitic left dislocated constituents of different

---

[1] The construction is also present in other Romance languages (see Benincà et al. 2001 for

Italian, Zubizarreta 1998 for Spanish, Vallduví 1992, Villalba 2000 for Catalan), as well as non-

Romance languages such as Greek (Anagnostopoulou 1994, 1997, Le Gac and Yoo 2011),

49   categories are given in boldface in (1) (from Delais-Roussarie et al. 2004): the

50   NP *cet homme* 'this man' in (1a), the AP *heureuse* 'happy' in (1b), the VP *partir*

51   'to leave' in (1c), the CP *qu'il se soit trompé* 'that he made a mistake' in (1d), and

52   the personal pronoun *moi* 'me' in (1e).

53   (1) Delais-Roussarie et al. (2004: 505-6)

54   a. **Cet homme$_j$** , je ne le$_j$ connais pas.

55   this man   I not him know not

56   'This man, I don't know him'

57   b. **Heureuse$_j$** , elle ne l'$_j$ a jamais été.

58   happy she not it has ever been

59   'Happy, she has never been'

60   c. **Partir$_j$** ,   c'$_j$ est   mourir un peu.

61   to leave, that is to die   a bit

62   'To leave, it is to die a bit'

63   d. **Qu'il se soit trompé$_j$** ,   c'$_j$ est évident.

64   that he has made a mistake, that is clear

65   'That he has made a mistake, it is clear'

66   e. **Moi$_j$** , personne ne veut m'$_j$ aider.

67   me, nobody not wants me to help

68   'Me, nobody wants to help me'

---

some Balkan languages (Kallulli and Tasmowsky 2008), or Arabic (for Lebanese Arabic, see

Aoun and Benmamoun 1998).

69   For the present work, only dislocated NPs are taken into account.[2] Three

70   examples extracted from our corpus are given below, displaying NPs with

71   different grammatical functions: subject (*le mec* 'the guy in (2), co-referring with

72   the pronoun *il* 'he'), direct object (*maman* 'mum' in (3a), co-referring with the

73   accusative pronoun *la* 'her'), and oblique object (the locative argument *Belleville*

74   in (3b), co-referring with the locative pronoun *y* 'there').[3]

75   (2) **Le mec**$_j$, il$_j$ parle bien français   [CID, AP]

76   the guy   he speaks well French

77   'The guy, he speaks French very well'

78   (3) a. Puis **maman**$_j$, je la$_j$ laisserais pas.   [CFPP, Mo-f]

79   then mum   I her would leave not

80   'Then my mum, I wouldn't leave her for any reason'

81   b. **Belleville**$_j$, j'y$_j$ vais souvent.

82   Belleville I there go often

83   'Belleville, I often go there'

84   Notice that (3b), if we follow Cinque's (1983) classical distinction (for Italian)

85   between CILD and Hanging Topic Left Dislocation (HTLD), is technically an

86   example of the latter construction. As a matter of fact, according to Cinque, one

---

[2] This restriction is motivated by the fact that dislocated elements of other categories are very

rare in spoken corpora (Blasco-Dulbececo 1999 , AUTHOR$_2$ 2012). See Section 2 for further

restrictions on our dataset and their motivations.

[3] Our corpus comprises two French corpora of spontaneous speech, which will be presented in

Section 2: the CID corpus and the CFPP2000 corpus. In the code associated to our examples,

the first letters indicate the name of the corpus, while the letters/numbers after the comma identify

the speaker.

87   of the properties of HTLD (but not CILD) is that the dislocated element is of NP

88   category only, even when it co-refers with an oblique pronoun. Nevertheless, as

89   noted by various scholars (see e.g. Delais-Roussarie et al. 2004, Berrendonner

90   2015), the distinction between HTLD and CILD is less strong in French than in

91   Italian, and cases like (3b) are very frequent in this language, and much more

92   common than their counterparts with preposition. And importantly, no prosodic

93   differences seem to exist between HTLD and CILD in French (see Doetjes et al.

94   2002, Avanzi 2012). In our data, we therefore included all dislocated NPs,

95   regardless of whether they were actually CILD or HTLD, in Cinque's terms. They

96   could be resumed by clitics marking different functions, as illustrated in (2) and

97   (3) above, and as explained in greater detail in Section 2.

98   Note also that a HTLD, according to Cinque's classification, allows for a full

99   tonic pronoun or an NP epithet to co-refer with the dislocated NP. Such

100   constructions were excluded from our corpus: only left dislocations with *clitic*

101   resumption were considered. That explains our decision to use the term 'CILD'

102   throughout the paper.

103   A final remark must be made on the construction under study, when the

104   dislocated NP is the subject. There is a long-standing debate in the French

105   literature concerning the status of the subject clitic in a CILD. According to

106   several authors, in contemporary spoken French, its status as a pronoun is

107   controversial and the clitic is rather analyzed as a verbal affix (basically, an

108   agreement marker, see Sankoff 1982, Auger 1993, Berrendonner 1993, 2007,

109   2015, Zribi-Hertz 1994, Apothéloz 1997, Fonseca-Greber and Waugh 2003,

110   Culberston 2010, among many others). If the analysis proposed by these

111   scholars is correct, then a construction in which the subject co-refers with a

112 subject clitic cannot be a dislocation: the lexical subject occupies the same

113 position as the subject of a canonical SVO sentence. As for the clitic, its function

114 is not to signal that the lexical subject is dislocated: in fact, its presence is not

115 motivated by syntactic reasons (see Culbertson and Legendre 2008, Culbertson

116 2010 for a semantic explanation of its presence in the clause, or Berrendonner

117 2007, 2015 for an account in terms of free variation).

118   If one follows the so-called 'affix hypothesis' and assumes that subject NPs co-

119 referring with a clitic are *not* dislocated in French, then the traditional pragmatic

120 analysis of that construction is also debatable. Traditionally, the literature

121 considers a dislocation – in different languages – as a topic-marking device

122 (Vallduví 1992, Lambrecht 1994, Anagnostopoulou 1994, Zubizarreta 1998,

123 Benincà et al. 2001, etc.). The dislocated constituent represents the sentence

124 topic, defined as the entity about which the utterance conveys some information

125 or provides a comment. For French, the topic status of the dislocated element

126 has been argued for by Barnes (1985), Ashby (1988), Lambrecht (1994, 2001),

127 among many others. Given the link between dislocation and topic, it is not so

128 straightforward, within the affix hypothesis, that the lexical subject be a topic.

129 This issue is debated in Brunetti et al. (2012), where a pragmatic and prosodic

130 analysis is conducted of both 'clitic-doubled' and 'non clitic-doubled' subjects

131 extracted from a French spoken corpus. What we observe in such studies is that

132 in both types of constructions, preverbal subjects may have a topic function, but

133 with some differences in their discourse and prosodic properties. In other words,

134 the link between dislocation and topic does not seem to be so tight, as far as

135 French preverbal subjects are concerned. In the present work, we therefore

136 assume that, even if the affix hypothesis is correct, it does not prevent a

137 preverbal clitic-doubled subject from being a topic.

138   Of course, we do not deny that there are 'clitic-doubled' subjects that are not

139 topics. Scholars endorsing the affix hypothesis have found several corpus

140 examples of non-topical doubled subjects, such as indefinites or

141 existential/negative quantifiers (see for instance Berrendonner 2007, 2015). No

142 occurrences of this kind of subjects, however, were present in the corpus we

143 examined. Furthermore, Berrendonner (2007) presents corpus examples where

144 is the linguistic context calls into question the dislocated position (and therefore

145 the topical status, in his view) of the doubled subject. One of these contexts is

146 when the utterance containing the double subject is coordinated with an

147 utterance *without* doubling, as in (4).

148   (4) a. **Les chiens** ils   sont autour   Berrendonner (2007:87, [CTFP])

149   the dogs they are around

150   b. et   le sanglier est au   milieu.

151   and the wild boar is at the middle

152   'Dogs are around and the wild boar is in the middle'

153   Cases like this are present in our corpus. However, as mentioned above, if we

154 assume that the link between dislocation and topic is not tight for French

155 preverbal subjects, this is not necessarily a problem, as both the doubled subject

156 in the first utterance and the simple subject in the second utterance may be

157 topics.

158   Another context where, according to Berrendonner (2007), the doubled subject

159 is not a topic is when it continues an already established topic, as in (5) below.

160 For Berrendonner, the subject does not need to be marked as topic in these

161 cases.

162   (5) a. j'y suis allée une fois   Berrendonner (2007:87, [Corpaix])

163   b. mais <u>mon père</u> <u>lui</u> y tenait pas que j'aille à ces fêtes

164   'I went there once, but my father did not like me to go to these parties'

165   c. parce que **mon père** il était beaucoup rigide

166   because my   father he was very   strict

167   'because my father was very strict'

168   Nevertheless, that depends on the definition of topic one adopts.

169 Berrendonner's notion seems to be restricted to those referents that cause a shift

170 in the ongoing discourse. As we will see in Section 4, our notion of topic is larger

171 and includes those cases where the topic continues an already established topic.

172 In sum, in our dataset we neither find subjects whose semantic status could

173 challenge their topic status, nor whose context of occurrence is incompatible with

174 their topic status.

175   In the present paper, we remain neutral with respect to the affix hypothesis,

176 focusing on the actual pragmatic and prosodic properties of subjects and objects

177 co-referring with a clitic.[4]

178   1.2. PROSODIC PROPERTIES OF CILDED NPS AND THEIR INTERACTION WITH

179 PRAGMATIC PROPERTIES

---

[4] Of course, if the affix hypothesis is correct, it is possible that doubled subjects and doubled objects display some pragmatic and/or prosodic differences. What our results show, as we will report in the Discussion Section, is indeed a difference in their *discourse* properties. However, their potential prosodic difference could not be tested, as the limited number of objects did not allow us to include them in the database for the statistical analysis. For more details on this, see Section 5 below.

180   Concerning the prosodic properties of French CILDs, it has been claimed for a

181   long time that in French, a CILDed constituent is followed by a major prosodic

182   break (Bally 1944, Rossi 1999, Mertens 2008). The break is usually realized by a

183   major rising pitch movement associated with a significant syllabic lengthening,

184   and possibly followed by a silent pause (see Avanzi 2012 for a review of the

185   studies making this claim).[5] In more or less recent literature, it has been claimed

186   that this prototypical prosodic realization may vary, and that the NP can be

187   followed by a minor prosodic break (actualized by a high static tone and a shorter

188   syllabic duration), or not followed by a prosodic boundary at all (Fradin 1990,

189   Berrendonner 2007, Culbertson 2010, Avanzi 2012).[6]

190   The reasons that motivate the variation in prosodic strength of the right

191   boundary of a left-dislocated constituent are still mysterious, and various

192   proposals have been put forward in the literature. The weight of the CILDed

193   constituent, when evaluated in terms of number of syllables, has been claimed to

194   play a role. Avanzi (2012) showed that short constituents (2 or 3 syllables long)

195   were better candidates to be associated with a minor prosodic break or no break

196   than longer ones. Nevertheless, the author failed to find differences between the

197   prosodic boundaries of embedded CILDed elements and non-embedded ones,

198   and between the prosodic boundaries of CILDed objects and CILDed subjects.

199   Delais-Roussarie and Feldhausen (2014) tested the effects of sentence modality

---

[5] The claim is also made for other languages, e.g. Italian (Gili-Fivela 1999), Catalan (Feldhausen 2010), German (Féry 2011), and Hugarian (Surányi et al. 2012).

[6] For example, in AUTHOR$_2$'s (2012) corpus (around 440 CILD constructions extracted from spontaneous speech), 70% of the dislocated elements are not followed by a major prosodic break. In read speech, falling movements have been observed (see Rossi 1999 or Delais-Roussarie et al. 2004).

9

200 (assertion vs interrogation) and syntactic complexity of the CILDed phrase

201 (branching vs non-branching), but failed to draw clear conclusions, due to the

202 lack of statistical analysis.

203   Pragmatic factors have also been invoked to explain the prosodic variability of

204 the dislocated constituent (see Lacheret-Dujour and François 2004, Grobet and

205 Simon 2009, Likhacheva 2008). In particular, Lacheret-Dujour and François

206 (2004) argue that the discourse activation status of the topic referent denoted by

207 the CILDed NP is inversely proportional to the degree of strength of the prosodic

208 boundary: strongly salient referents are not followed by a strong prosodic break,

209 while non-salient or new referents are more often followed by a strong prosodic

210 break. In previous quantitative work on spontaneous speech, however, this

211 hypothesis could not be confirmed: the degree of strength of the prosodic

212 boundary of the dislocated NP and the degree of salience of its referent were not

213 significantly related (see Avanzi et al. 2012, Brunetti et al. 2012).

214   In the present work, we propose a more exhaustive classification of the topic

215 referents denoted by a dislocated NP with respect to their discourse properties,

216 in order to study the interaction of these properties with the strength of the

217 prosodic boundary at the right edge of the NP. The classification is presented in

218 detail in Section 4, but to anticipate a little, a brief definition of the three

219 parameters in the classification is given here, namely: topic continuity, discourse

220 givenness, and contrast. Concerning the first parameter, topics can be

221 'continuing' or 'shifted' (see Givón 1983, Lambrecht 1994, Butt and King 1997, de

222 Hoop and de Swart 2000, Brunetti 2009). A continuing topic is the topic of the

223 current discourse (see examples (11)-(13) in section 4.1), while a shifted topic is

224 new with respect to the ongoing discourse (see examples (14)-(19) and (21)-(24)

225 in section 4.2) (cf. Givón 1983, Herring 1990, Vallduví 1992, Lambrecht 1994,

226 Gundel and Fretheim 2004, de Hoop and de Swart 2000, Frascarelli and

227 Hinterhoelzl 2007). Shifted topics are often confused with or defined in the same

228 way as *contrastive* topics.[7] In the present paper, we consider contrast as a

229 separate property: a topic is contrastive if a set of alternative propositions is

230 explicitly given or at least easily recoverable in the discourse (Büring 1997,

231 Valduví and Vilkuna 1998). We further assume that what is asserted in the

232 proposition containing the topic is compared or opposed to what is asserted in

233 the alternatives (cf. the contrastive relations SIMILAR and OPPOSE in Repp 2016).[8]

234 Shifted topics may or may not have contrast. Cases of non-contrastive as well as

235 contrastive shifted topics will be given in Section 4.2. The same holds for

236 continuing topics: they can be contrastive or not, as we will show in Section 4.1.

237 In other words, the continuing/shifted dimension interacts orthogonally with the

238 contrastive dimension (cf. Vallduví and Vilkuna 1998, Neeleman and Vermeulen

---

[7] As already noted by Barnes (1985) and Ashby (1988). The reason is probably the fact that, when the contrastive topic is in the second utterance of a pair of contrasting utterances, then it obviously also shifts, because the preceding utterance contains a different topic. See for instance (27) (copied here for convenience), where the topic *les autres* is in contrast with *celui-là,* and it also shifts with respect to *celui-là.*

(27) **celui-là**, je l' connais, **les autres**, je les connais pas   [CFPP, MoF]

    that-one   I it know   the others I them know not

    'that one I know, the other ones I don't'

[8] Repp also discusses the relation CORR (for correction). Only two cases of corrective contrast were found in our dataset.

239   2012). The third dimension we consider - discourse givenness – concerns the

240   availability of a topic referent in previous discourse. Since continuing topics

241   cannot be discourse new by definition, this dimension only applies to shifted

242   topics, which can be either new or recoverable from previous discourse. The

243   differences between the latter two topics – which we call *new* and *resumptive*

244   shifted topics – are presented in Sections 4.2.1 and 4.2.2.

245   In previous literature, some scholars associated different prosodic profiles with

246   topics having (some of) the properties mentioned above. Frascarelli and

247   Hinterhoelzl (2007), within the AM framework, proposed that three different tones

248   are systematically associated, in Italian and German, with three topic types:

249   'shifting', 'contrastive', and 'familiar'.[9] Specific accents, marking a (contrastive)

250   topic shift have been claimed to exist in many languages, starting from

251   Jackendoff's (1972) analysis of the pragmatic properties of the English 'B-

252   accent', which was later developed by Büring (1997, 2003) for both English and

253   German (see also Brunetti et al. 2010 on Italian and Marandin et al. 2002 on

254   French). Marandin et al. (2002) propose the existence of a high tone on the first

255   syllable of the 'thematic shifter' (that is, simplifying a bit, the part of the XP that

256   triggers contrast), see *anglais* in (6a), or on the head of the topic phrase (see

257   *chanteurs* in (6b)), and in some cases, on both (see (6c)):

258   (6)   Que fumaient les chanteurs de rock?   Marandin et al. (2002:2)

259   a. Les chanteurs de rock ANglais fumaient des cigarettes.

260   b. Les CHANteurs de rock anglais fumaient des cigarettes.

261   c. Les CHANteurs de rock ANglais fumaient des cigarettes.

---

[9] For instance, for Italian, they propose that a complex L*+H tone is systematically associated

with shifting topics, a H* tone with contrastive topics, and a L* tone with familiar topics.

262   Doetjes et al. (2002) claim that the discourse referent status can explain the

263   distribution between two types of contours after a dislocated NP. When the

264   discourse referent is active, the contour would be H%, while when it is non-active

265   but accessible, the contour would take the shape of a H(L)%, i.e. of an

266   illocutionary tone corresponding to a confirmation request. Refining Doetjes et

267   al.'s proposal, Delais-Roussarie et al. (2004) suggest that the H(L)% boundary

268   tone is used "whenever an agreement concerning the choice of the topic is not

269   fully established between the speech participants" (Delais-Roussarie et al.

270   2004:519).

271   All the studies mentioned above are mainly based on constructed examples.

272   When examples are extracted from corpus data, either such data come from

273   read texts, or, if they come from a spontaneous speech corpus, they are only

274   subject to a qualitative analysis. No quantitative analysis of *spontaneous* speech

275   corpus data exists, to the best of our knowledge, on the relationship between

276   prosody and discourse functions of the sentence topic denoted by a ClLDed NP

277   in French. In the present work, we try to fill this empirical gap.

278   The rest of our paper is organized as follows. In Section 2, a description of our

279   dataset is given. A detailed account of the prosodic and pragmatic annotation is

280   provided in Sections 3 and 4 respectively. The results of the analysis of the

281   interaction between the two annotations are given in Section 5, and a discussion

282   of our results is proposed in Section 6. In Section 7, general conclusions are

283   drawn.

284   2. Dᴀᴛᴀ

285   Naturally occurring utterances were extracted from two corpora of spontaneous

286   spoken French: the Parisian French *Corpus de Français Parlé Parisien des*

287   *années 2000 CFPP2000* (Branca-Rosoff et al. 2009) and the Southern French

288   *Corpus of Interactional Data CID* (Bertrand et al. 2008).

289   The CFPP2000 comprises interviews that are about 1 hour and a half long and

290   involve one or more inhabitants from different districts in Paris, who were asked

291   to talk about their past and present life. We extracted utterances from 7

292   interviews out of the 31 in the corpus, for a total of about 9 hours of speech. The

293   extracted utterances were pronounced by the interviewees. The CID corpus

294   comprises 8 one-hour spontaneous dialogues between two acquaintances

295   (people working or studying at the 'Parole et Langage' laboratory in Aix-en-

296   Provence, France) in an acoustically isolated room. Participants were given a

297   general topic: either to talk about unusual episodes that happened in their life, or

298   about conflicting relationships at their workplace. Utterances were extracted from

299   the entire CID corpus, which is approximately 8 hours long. The total number of

300   speakers in our corpus is 28: 21 women and 7 men. Their age ranges from 15 to

301   70 years old. Speakers uttered a minimum of 1 and a maximum of 29 CILDs.

302   2.1. DATA SELECTION

303   In order to make our data as homogeneous as possible and to limit the

304   complexity of the analysis, we only included a subset of the dislocations, which

305   respected a number of prosodic, syntactic, and semantic criteria.

306   As we said above, only clitic left dislocated NPs were included in our corpus.

307   Also, following Avanzi (2012), we only took those NPs that are resumed by a 3rd

308   person singular or plural clitic pronoun: *il(s)*, *elle(s)* for dislocated subjects, *le(s),*

309   *la, en* for dislocated direct objects, and *lui, y, en* for dislocated oblique objects,

14

310　therefore excluding 1st and 2nd person dislocations (*moi, je..., toi, tu..., nous,*

311　*on...*), as well as dislocations resumed by the clitic *ce*.[10]

312　　For prosodic reasons, we excluded utterances whose dislocated element was

313　separated from the verb by some prosodically relevant linguistic material, such

314　as *dans le temps* in the example below. Note that we did not consider clitics as

315　prosodically relevant material, and that in the utterances with dislocated objects,

316　the subject was always a clitic, so it never intervened between the dislocated

317　object and the verb.

318　　(7) **Les chiens** dans le temps on en　　avait qui...　　[CFPP, 6X]

319　　　　the dogs　　in　the　time　we of-them had who

320　　　　　'We had dogs in the past, who…'.

321　　We also discarded those utterances whose dislocated element is not sentence

322　initial, more precisely whose dislocated element is preceded by another topic

---

[10] Contrary to 3rd person pronouns, preverbal 1st and 2nd person pronouns are obligatorily

resumed by a clitic:

(i) a. Moi/Toi *(je/tu) préfère(s) la voiture rouge.

　　me　you　I/you prefer　the　car　　red

　b. Paul/Lui (il) préfère la voiture rouge.

　　Paul/he　he　prefer　the car　red

Thus, they might not share the same prosodic, syntactic and pragmatic properties as 3rd person

NPs. As for resumption with the clitic *ce*, it triggers a generic interpretation of the NP. This means

that clitic resumption (and therefore dislocation) in this case is triggered by interpretative needs,

and may therefore be totally unrelated to topic-marking. As for the sequence *XP c'est X* (*c'est* 'it

is'), the construction is ambiguous between a ClLD and a pseudo-cleft (see Pekarek Doehler and

Mueller 2009).

323   phrase or by an adjunct phrase, typically a temporal or spatial sentential adjunct,

324   such as *après le repas* 'after the meal' in the example below:

325   (8) Après le repas **les gamins** tu les   envoies à la sieste   [CID, EB]

326   after the meal the kids   you them send   to the nap

327   'After lunch, the kids are sent to have a nap'

328   This choice was motivated by the fact that the rising contour may be less marked

329   in the case of a sequence of extrasentential elements as in this example.

330   Based on Avanzi (2012), we also excluded all utterances where the dislocated

331   element was followed by a disfluency, marked by a filled pause or a repetition,

332   such as in (9).

333   (9) **Les autres** euh ils sont un peu partout en France   [CID, MB]

334   The others ehm they are a little bit everywhere in France

335   'The others, they are scattered almost all over France'

336   This choice was motivated by the fact that disfluent elements disturb the prosodic

337   structure, and therefore would have created problems during the calculation of

338   the relative acoustic parameters (see infra, Section 3).

339   Also for reasons of prosodic uniformity, we only retained those dislocated NPs

340   that were between 2 and 6 syllables long. On the one hand, monosyllabic NPs

341   are too short for the prosodic boundary to be clearly detected. On the other, it

342   has been shown that dislocated elements comprising more than 6 syllables are

343   rarely phrased in a single prosodic phrase (Martin 1987, Delais-Roussarie 1995),

344   i.e. their behavior differs from the average number of NPs in our data. Their

345   exclusion therefore allowed us to limit the effect of the length parameter in the

346 realization of the final contour.[11] In addition to this, we carefully checked all the

347 NPs containing more than one lexical word.[12] Utterances such as (10) were

348 concerned.

349   (10)   [CFPP, 70]

350   **les mamans africaines** elles ont beaucoup de mal à se séparer de leur petit

351   the mums   African   they have a lot of difficulty to part from their little-one

352   'African mothers find it hard to part from their little ones'

353 In this example, *les mamans africaines* comprises two prosodic phrases.[13]

354 Indeed, the NP is made of two lexical words (*mamans* and *africaines*) which are

355 both associated with a prosodic boundary on their rightmost syllable. Dislocated

356 NPs that comprised two prosodic phrases were identified by visual inspection of

357 pitch curves and alignment (for duration). In total, we found 7 cases of NPs

358 phrased into two prosodic phrases, which we excluded.

359   Given that the prosody – as well as the information structure – of interrogative

360 and imperative sentences is presumably different from that of declarative ones

361 (Delais-Roussarie et al. 2015), we excluded such sentences from our corpus.

362 Also, in order to facilitate the pragmatic annotation, we excluded those

363 utterances that were part of quoted speech, as it is not very clear how the topic

---

[11] The number of syllables was obtained by calculating the number of vowels in the NP (including

extensions and determiners). The preceding or following material was not taken into account.

[12] According to Nespor and Vogel (1986) and Delais-Roussarie (1995), all lexical items belonging

to the paradigm of nouns, verbs, adjectives and adverbs, constitute a lexical word.

[13] By prosodic phrase, we mean a syntactic phrase that minimally contains a lexical word and all

the grammatical words associated with it, which is bounded on its rightmost syllable by a prosodic

boundary.

364 element is related to previous (non-quoted) speech (and to the referents

365 contained in it).[14]

366   As far as subordinate clauses are concerned, according to some scholars they

367 do not have an independent informational partition into topic and focus.

368 Berrendonner (2007), for instance, in support of the 'affix hypothesis' for subject

369 clitics (see Section 1.1), takes the presence of a CILDed subject in a subordinate

370 clause as evidence that the subject is not dislocated (and therefore not a topic).

371 Yet, many different root phenomena, and in particular information-structure

372 related ones, such as dislocations, have been observed in subordinate clauses in

373 different languages (see the seminal paper by Hooper and Thompson 1973, as

374 well as more recent studies such as Haegeman 2004, Sawada and Larson 2004,

375 Bianchi and Frascarelli 2010, Lahousse 2010, Sleeman 2012). Among the 36

376 subordinate clauses in our dataset, the availability of an independent information

377 structure was debatable for only three of them; these utterances were therefore

378 discarded. Seven of the remaining 33 are *que* declarative clauses introduced by

379 speaker-oriented verbs such as 'think' or 'realize'. Following Simons (2007),

380 Bentzen (2009) and others, we assume that such subordinate clauses express

381 the 'Main Point of the Utterance' (MPU). In other words, the most relevant

382 information provided by the utterance at that point in the discourse is contained in

383 the subordinate clause. This means that the clause has *an information structure*

384 *of its own*, and as a consequence, that the dislocated NP may legitimately be

385 identified as the sentence topic. Finally, 23 subordinate clauses in our corpus are

386 adjunct (or adverbial) clauses. Following Lahousse 2010, we take adjunct

387 clauses to have an information structure of their own. Adjunct clauses express

---

[14] Their prosody might presumably be different too (see Bertrand and Espesser 1998).

388   what is generally called "not-at-issue" content (Simons et al. 2010), namely

389   information that does *not* convey the utterance main point, or put it differently,

390   that does not answer the main Question Under Discussion (Roberts 1996,

391   Ginzburg 1996). Not-at-issue content, however, does answer a *secondary*

392   question. The conclusion is therefore the same as the one reached for *que*

393   declarative clauses introduced by speaker oriented verbs: these clauses have an

394   information structure of their own, and therefore their CILDed NP can be

395   considered as a topic. Among the remaining three subordinate clauses of our

396   corpus, two constitute the second part of a cleft introduced by *y a* (*y a des fois*

397   *où…* 'there are times when…') and one constitutes the second part of a pseudo-

398   cleft (*le problem c'est que…* 'the problem is that…'). We assume that both

399   constructions convey a single proposition and have a single information structure

400   (cf. Lambrecht 1994, 2001, Pekarek Doehler and Mueller 2009), so again,

401   nothing prevents the CILDed NP from being considered as a topic. While the

402   information structure of the subordinate clauses of our corpus is comparable to

403   that of root clauses, the prosody might be affected by the embedded context.

404   That is why we eventually excluded them from the dataset used for the statistical

405   analysis (see Section 5).

406   As for the intonational properties of the NPs, the H(L)% pattern observed in

407   Doetjes et al. (2002) was not found in our corpus. In our corpus, all the contours

408   are actualized by a continuative contour (Hcont). Because of this lack of

409   variation, we could not address Doetjes et al.'s (2002) claim concerning the link

410   between a H(L)% contour and the non-active status of the referent. As for

411   Marandin et al.'s (2002) claim about the presence of an internal H* tone when the

412   referent of the dislocated NP is a thematic shifter, the low number of utterances

413 with an internal prominence (only 5) did not allow us to test such a claim either

414 (see Figure 5 below for an illustration).[15] Note, furthermore, that the context of

415 the topic in Marandin et al.'s example in (6) is a question-answer pair with a

416 *partial* answer (which informs about *British* rock stars rather than *all* rock stars).

417 As shown by Büring (1997), a prosodic marking is obligatory in that context,

418 because it legitimates the partial answer by signaling the existence of

419 alternatives. Contrastive topics in our dataset, however, never occur in partial

420 answers; therefore, we do not expect such a prosodic marking to be present.

421   2.2. CONCLUSIONS ON DATA SELECTION

422 To conclude this section concerning the construction of our dataset, it should be

423 mentioned that about 40 utterances were eventually excluded because of their

424 poor sound quality. We therefore ended up with a total of 241 utterances: 204

425 subjects and 37 objects. They are distributed in the two corpora in the following

426 fashion:

		CFPP	CID	Total
function	Subject	89	111	204
	Object	22	15	37
embedded	yes	11	22	33
	no	99	109	207
number of syllables	2 syll.	56	72	128
	3 syll.	29	31	60
	4 syll.	13	14	27
	5 syll.	8	11	19
	6 syll.	5	2	7

427   Table 1. Characteristics of the utterances according to their structural properties as a function
428 of the corpus.
429
430
431
432

---

15 This little number can be explained by the fact that we retained for the analysis only short NPs,

whose prosodic structure is usually very simple.

433   3.   PROSODIC ANNOTATION: DEGREE OF BOUNDARY STRENGTH

434   It is commonly assumed that a sentence is composed of different prosodic

435   phrases, which are hierarchically organized (Selkirk 1984, Nespor and Vogel

436   1986). In the most recent AM accounts of French prosodic structure, it is

437   generally agreed that three levels of prosodic units need to be distinguished: the

438   'Accentual Phrase' (AP), the 'Intermediate phrase' (ip) and the 'Intonational

439   Phrase' (IP) (Michelas and D'Imperio 2010; Delais-Roussarie et al. 2015). The

440   AP boundary is perceptually less salient than the ip boundary, which is also less

441   salient than the IP boundary. Theoretically, all these levels are considered to

442   have a distinct realization, mostly regarding the way their boundaries are realized

443   (D'Imperio et al. 2007; Michelas and D'Imperio 2010; Delais et al. 2015). In

444   practice however, it is very difficult for the analyst to decide at which level of the

445   prosodic hierarchy a boundary is by relying on acoustic cues only, especially in

446   the case of spontaneous speech. Since no automatic system to annotate such

447   categories exists yet, we decided to use the Analor software to categorize the

448   strength of the prosodic break associated with the last syllable of each dislocated

449   NP. Some equivalences between the scores provided by Analor and the units

450   defined in the AM theory have been proposed in the literature (see Avanzi 2012;

451   Delais-Roussarie & Feldhausen 2014), but such correspondences rely on the

452   intuition of the authors, and have not been tested perceptually. In this context, we

453   thought it was wiser not to establish any *a priori* equivalence based on the score

454   and the AM hierarchy, and therefore to consider the scores provided by Analor

455   as continuous variables. We will return to this point in the discussion.

456   Analor is a tool developed by Avanzi et al. (2011) for the automatic

457   identification of prominent syllables in French. It is based on the hypothesis that

458 prosodic prominences and prosodic boundaries are tightly intertwined in this

459 language. Since it takes into account duration phenomena, it is more accurate

460 than an exclusively tonal approach in terms of pitch accent shape, such as the

461 one proposed for instance in Frascarelli and Hinterhoelzl (2007) for Italian and

462 German or in Le Gac and Yoo (2011) for Greek.[16] We therefore used the Analor

463 system to estimate the degree of strength of the prosodic boundary ending the

464 dislocated constituent.

465   The use of this tool presents numerous advantages:

466   •   It makes it possible to account for the fact that the perception of prosodic

467 boundaries and their realization in speech is continuous, not categorical (Collier

468 et al. 1993, Cole et al. 2010);

469   •   It provides an objective annotation, which can be reproduced, and thus

470 constitutes a good alternative to the variation typically occurring among experts

471 when labeling prosodic boundaries manually (Ladd 1996, Lacheret-Dujour et al.

472 2010).

473   •   It helps in taking into account the fact that the realization of the prosodic

474 boundary in French involves both F0 and duration variations, with *compensation*

475 effects between the two (Delattre 1938, Hart et al. 1990). Indeed, if one of the

476 parameters involved in the perception of a boundary presents a low value and

---

[16] The decision to rely on prominence score instead of pitch contour shape was also dictated by

the fact that in spontaneous speech, the pitch contour associated with the last syllable of the NP

is generally rising or flat. In our data, all the pitch contours were flat or rising, depending on their

duration (long syllables being more often associated to rising tones than shorter ones, which are

preferentially realized by a flat tone, see Section 2.1. above).

477   another one a high value, the software calculates a global medium score, thus

478   reflecting the actual perception of a medially strong prosodic break.

479   The estimation of boundary strength in Analor is performed in three steps.

480   First, all sentences are orthographically transcribed in Praat (Boersma and

481   Weenink 2014). Next, all sentences are processed with the EasyAlign script

482   (Goldman 2011). This tool provides a 3-layer automatic segmentation: phone

483   string, syllable string, and word string. All the alignments are checked and

484   corrected by hand on the basis of the inspection of both wave curves and

485   spectrograms. Third, an estimation of prosodic boundary strength is obtained by

486   calculating four prosodic parameters:

487   •   syllabic duration;

488   •   syllabic height average (following House 1990, only the F0 points of the

489   vowels are taken into account);

490   •   amplitude of the rising tone on vocalic nucleus;

491   •   presence of a silent pause not connected with a hesitation or a false-start.

492   In detail, the algorithm calculates for the current syllable S_0 (see Figure 1): its

493   relative average height and duration compared with the f0 and averages of the

494   three preceding syllables (S_{-3}; S_{-2} and S_{-1}) and the three following ones (S_{+1}; S_{+2}

495   and S_{+3}); the presence of a rise if there is a positive movement of f0 on the

496   syllabic nucleus (in lighter color in Figure 1); and the presence of an adjacent

497   silent pause. F0 measures are given in semi-tones, while duration measures are

498   calculated without any unit. Let us note that in its current version, the Analor

499   algorithm does not deal with the intensity parameter. The argument put forward

500   by one of the designers of the software (Avanzi 2012) is that intensity in French

501   is not an acoustic parameter involved in accentuation and phrasing (Delattre

502   1938 : 142; Astésano 1999), and that this parameter is instead often correlated

503   with F0 modulations (Lacheret-Dujour and Beaugendre 1999).[17]

504   For each of these three acoustic parameters (relative duration, relative height

505   and rising movement) a score between 0 and 10 is then attributed. This score is

506   determined according to the difference with an optimized threshold fixed during

507   the corpus-based learning procedure (see Avanzi et al. 2011 and Avanzi 2012 for

508   more details). The results of the calculation are provided in a dedicated tier: a

509   score between 0/10 and 10/10 (from the least prominent to the most prominent)

510   indicates the degree of strength of the break (rounded to the nearest unit). They

511   are then retrieved in a spreadsheet for further statistical analysis.

512   In the example in Figure 1, the rightmost syllable of the CILDed constituent *le*

513   *premier rang* is associated to a prosodic boundary of 6/10. This score results

514   from a rising movement on the nucleus of 6.3 semi-tones, a difference in height

515   of 2.2 semi-tones with respect to the six surrounding syllables, and a lengthening

516   of approximately 42% regarding the average duration of the three preceding and

517   three following syllables.

---

[17] Another reason for which we did not take intensity into account is that the conditions within which the recordings were made were not controlled. Intensity is very sensitive to the recording conditions.

518

Figure 1. Analor screenshot of the utterance: *le premier rang il est vraiment collé à l'écran* 'the front row is really stuck to the screen' [CID]. In the abscissa, temporal values are given in milliseconds; in the ordinate, the values of F0 in a logarithmic scale can be seen on the left, and the intensity values in Db can be seen on the right. Duration labels are given in milliseconds. Annotation tiers are, from top to bottom: phones, syllables (both in SAMPA) and graphemic words.
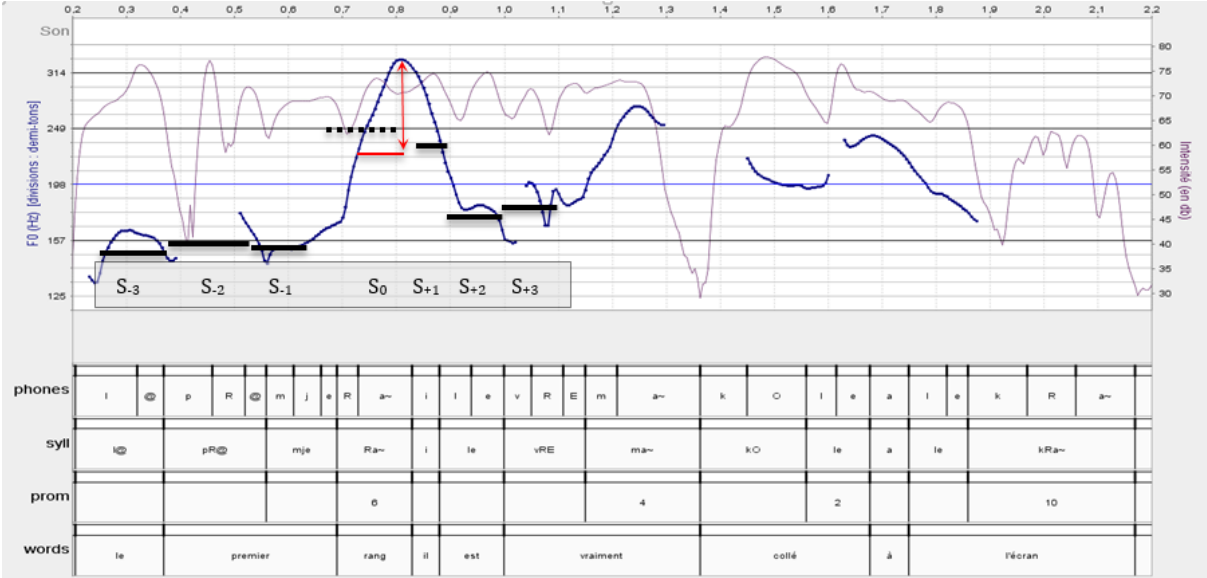
In the example in Figure 2, a prosodic boundary is detected after the last syllable of the dislocated NP *les femmes*. As can be seen, the F0 points on the nucleus of the syllable are just a little bit higher than the F0 points of the surrounding syllables (3.2 semi-tones), but the syllable is not really lengthened when compared with the directly preceding and following syllables.

Figure 2. Analor screenshot of the utterance: *euh les femmes elles arrivaient effectivement*
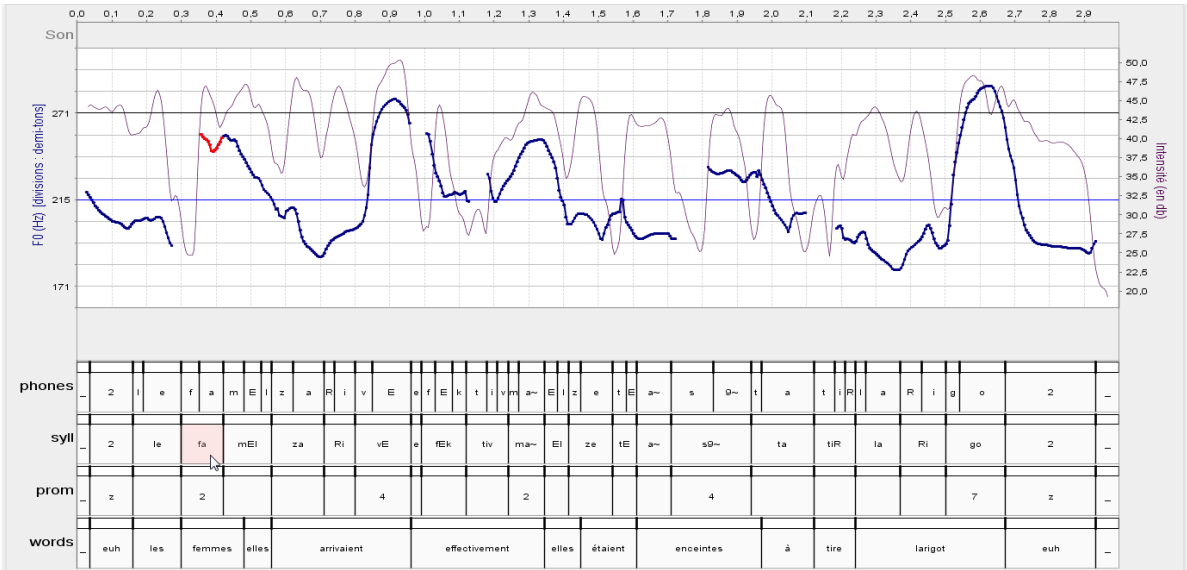
536 *elles étaient enceintes à tire larigot* 'euh women used to come, in fact they were constantly
537 pregnant'" [CFPP].
538

539   It is important to note that the 6-syllable window is a maximum context for

540 calculation; in fact, the context used for the calculation of the relative parameters

541 is often smaller. The context for calculation of the relativized acoustic

542 measurements can be reduced in two specific situations. First, when one of the

543 three adjacent syllables preceding the target syllable is associated with a

544 disfluency, as is the case in the example in Figure 2. Due to the fact that the

545 second syllable preceding the target syllable (*euh*) is disfluent, the relativization

546 is made on a restrained context (one syllable before, three after)[18].

547   The second case involves utterances within which the CLIDed element is

548 preceded by some material that contains a major prosodic break on its rightmost

549 syllable (i.e. a word or group of words whose last syllable is associated with a

550 major pitch movement). For example, in Figure 3, the material preceding the NP

551 (*retour à l'envoyeur*) is associated with a major prosodic boundary on its

552 rightmost syllables (the final syllable of the word *envoyeur* is associated with a

553 pitch rise of 4.3 semi-tones). In this context, in order not to weaken the estimation

554 of the score of the target syllable (see Avanzi 2012), it was decided that this

555 syllable would not be taken into account in the calculation of the prosodic

556 parameters. In other words, in this example, the calculation is made only on the

557 syllable preceding the target syllable and the three following it[19].

---

[18] When the target syllable was directly followed by a disfluency, the utterance was excluded from

the analysis.

[19] For each example, the identification of a major prosodic break in the vicinity of CILD was

performed manually, on the basis of the perception of the second author, and the examination of
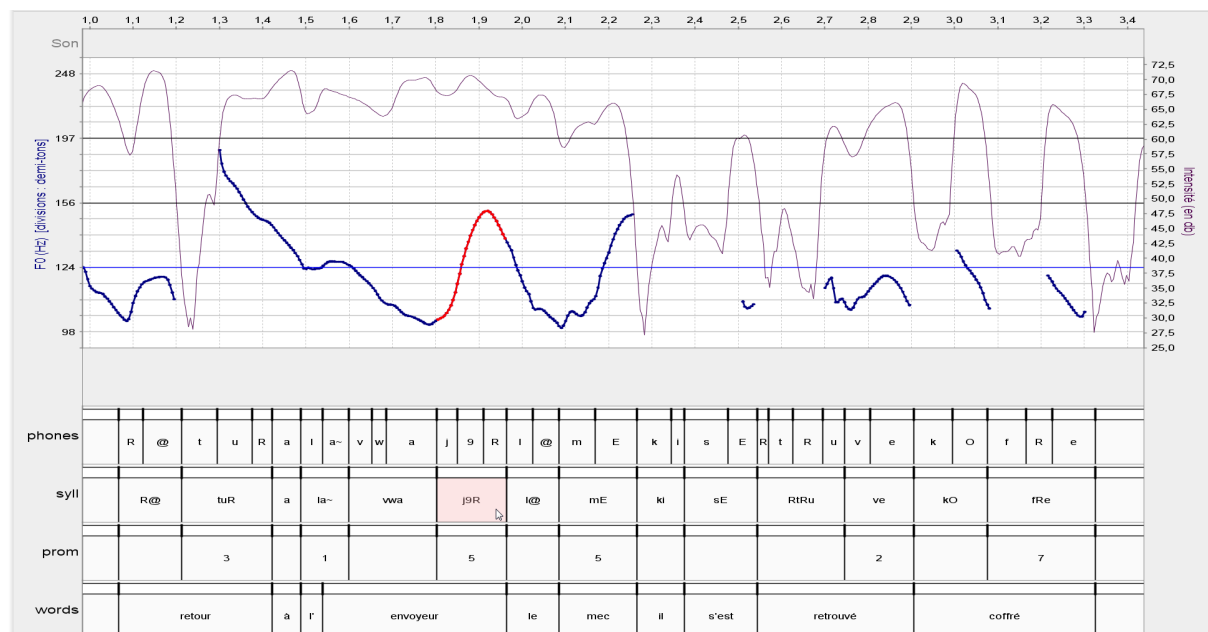
Figure 3. Analor screenshot of the utterance: *retour à l'envoyeur le mec il s'est retrouvé coffré* 'return to sender: the guy found himself locked away' [CID].

In contrast, when the last syllable of the words or group of words that preceded the dislocated element did not stand out from its environment, it was then considered as part of the context, and therefore used for the calculation of the relativized acoustic parameters. The example presented in Figure 4 exemplifies this:

---

pitch tracks. Following Mertens (2008), we considered that a melodic movement was major if its amplitude was higher than 2dt.

568
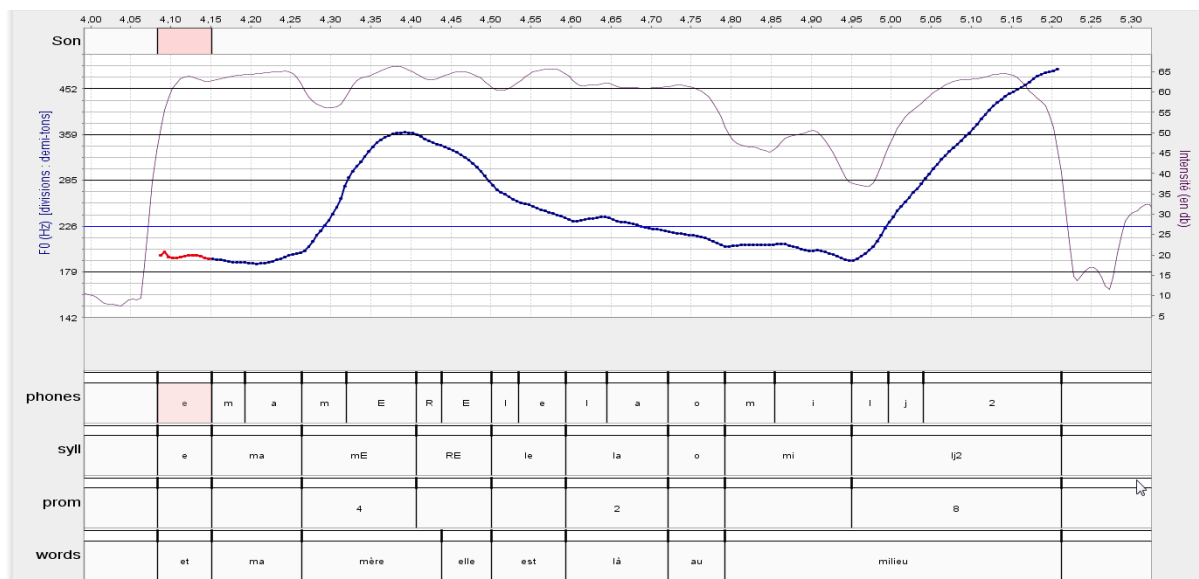
Figure 4. Analor screenshot of the utterance: *et ma mère elle est là au mileu* "and my mother, she is there in the middle" [CFPP].

In this example, as the coordinator *et* ('and') was not coded as prominent, it was

considered as a valid context for prosodic relativization.

Finally, it should be noted that some examples in the corpus presented

internal prominence within their NP, such as the word *tous* in Figure 5:
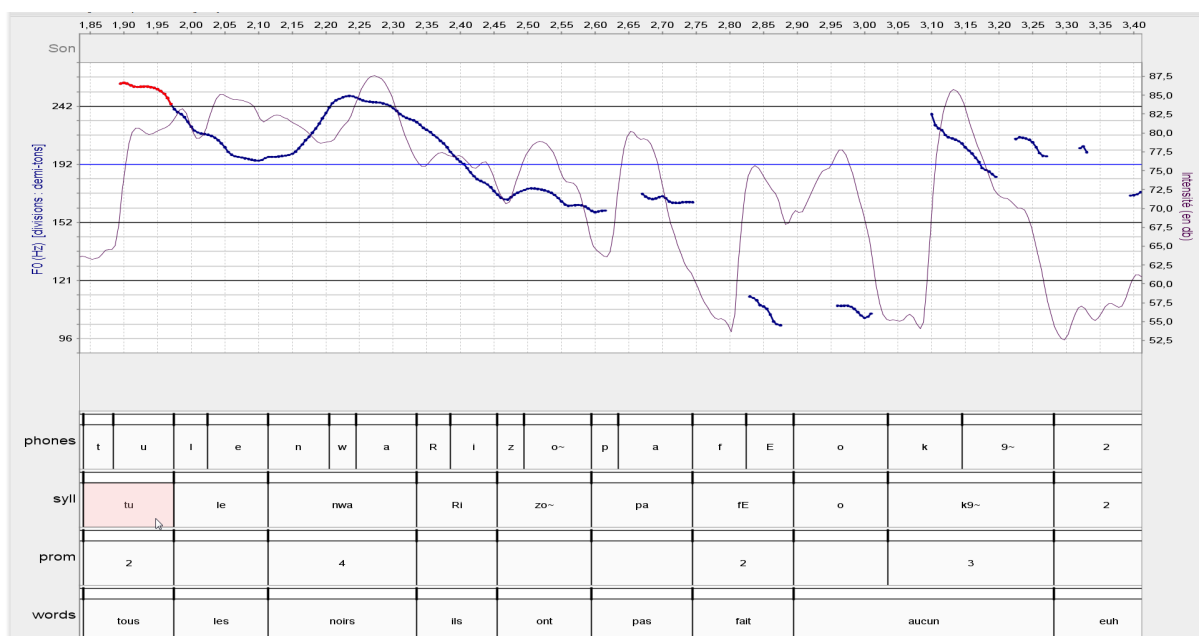

576

Figure 5. Analor screenshot of the utterance: *Tous les Noirs ils ont pas fait aucun euh aucune étude* 'Blacks didn't get any ehm any education' [CFPP].

579

28

580 Even if these cases are pretty rare in our corpus (only 5 utterances were found to

581 present a potential internal prominent syllable), we decided to apply the same

582 procedure as the one used for the disfluencies and major breaks discussed

583 above, and for the same reasons explained above. In other words, for these

584 examples, the syllable associated with an initial rise was excluded from the

585 context of relativization.

586   4. PRAGMATIC ANNOTATION

587 For the pragmatic annotation,[20] CILDed NPs were first divided into continuing

588 topics and shifted topics, as mentioned above. Two further properties were then

589 considered: discourse givenness and contrast. A continuing topic can only be

590 discourse given, so its sole margin of variation concerns the presence of

591 contrasting alternatives. A shifted topic, on the contrary, can be discourse given

592 or discourse new, and it can be contrastive or not. Three different possibilities

593 then result from the combination of topic continuity with discourse givenness: a

594 continuing topic, a shifted topic that is new in the discourse, and a shifted topic

595 that is given in the discourse, which we call 'resumptive'. The latter has either

596 been the topic of a preceding segment of discourse, but not the current one, or it

597 is present in the preceding discourse, but not as topic (for more details on this,

598 see Section 4.2 below). All three topics can additionally evoke contrasting

599 alternatives. Table 2 summarizes these possibilities.

600

601

602

---

[20] The pragmatic annotation was done by the first author and by another trained annotator. The

two annotations were then compared and the ambiguous or divergent cases discussed.

	Properties		
Topic	topic continuity	givenness	contrast
Continuing	YES	YES	YES/NO
Shifted (new)	NO	NO	YES/NO
Shifted (resumptive)	NO	YES	YES/NO

Table 2. Characteristics of continuing and shifted topics with respect to discourse givenness and the presence of contrast.

In the following sub-sections, we are going to detail the characteristics of each topic type.

4.1. *Continuing topic*

An example of continuing topic is the NP *le mec* 'the man' in (11b). Speaker AP is talking about a person he met at his own place during a party. AP found this person weird, but then he admits that he had misjudged him, realizing that he is just a withdrawn person. The referent of *le mec* 'the guy' in (11b) is discourse given (its antecedents are underlined in (11a)) and it is the topic of the immediately preceding discourse: the speaker has been talking about this man just before uttering (11b). The dislocated NP therefore represents a continuing topic.[21]

(11) [CID, AP]

a. Maintenant je <u>le</u> connais hum assez: je dis pas que le courant passe super bien tu vois, mais parce que c'est un peu un mec un peu renfermé; mais en fait euh je me suis planté dans mon jugement quoi; en fait <u>le type</u> il était mal euh <u>il</u> était chez moi, <u>il</u> m'avait jamais vu, euh <u>il</u> savait pas trop quoi me dire tu vois enfin bon

---

[21] Disfluences of spontaneous speech such as repetitions, false-starts, lengthenings, fillers, etc. have been discarded in the examples.

623   'Now I know him: I don't say that we get along extremely well, you know,

624   but it's just because he is a rather withdrawn person; but before, I was mistaken

625   in my judgment; in fact, the guy was not at ease, he was at my place, he had

626   never seen me before... he didn't know what to say, you know, so...'

627   b. et   bon tu vois   **le mec**, il   s'est renfermé   sur lui-même

628   and so   you see the man he has withdrawn on himself

629   'and so, you see, the guy withdrew into himself'

630   The first thing we noticed in our corpus is that continuing topics are few in

631   number (11 in total). This is not unexpected. A continuing topic maintains topic

632   continuity, so it corresponds to what speakers most expect to be the topic of the

633   upcoming utterance. It is therefore natural that such topics be generally realized

634   in a pronominal form, instead of a full (dislocated) NP. A pronominal realization of

635   a continuing topic is indeed frequently observed in the literature (see Reinhart

636   1981, Lambrecht 1988, Butt and King 1997, de Hoop and de Swart 2000,

637   Brunetti 2009, among many others). The puzzling fact may rather be the

638   opposite, namely that some continuing topics are realized as dislocated full NPs.

639   For our data, various explanations are possible. In (12), the reason for uttering

640   *Rémi* in (12f), which is the topic of the preceding utterances (where, as expected,

641   it is realized as a clitic) can be EB's intention to emphasize, or express his

642   agreement for, SR's comment in (12e).[22]

643   (12)   a. SR: avec Rémi on s'engueule putain   [CID, EB-SR]

644   'Rémi and I, we argue with each other all the time, fuck'

---

[22] In her corpus study on French left dislocations, Barnes (1985:67) had already mentioned

emphasis as a possible explanation for the presence of a left dislocated continuing topic (in her

terms, a referent that is given and is already the topic of the discourse).

645   b. EB : ouais avec <u>Rémi</u> vous vous engueulez ouais

646   'Yeah, Rémi and you, you argue all the time, that's true'

647   c. SR : on s'engueule; mais c'est parce que <u>il</u> dit n'importe quoi

648   'We argue; but it's because he talks nonsense!'

649   d. EB : ouais c'est vrai ; c'est marrant <u>il</u> dit la même chose de toi; ah

650 tu vois, des fois il fait euh : "C'est quand même moi qui suis informaticien quoi

651 alors bon je sais de quoi je parle"

652   'Yes, it's true; funny: he says the same thing about you! You know,

653 sometimes he says: "*I* am the computer engineer! So I know what I'm talking

654 about'

655   e. SR : <u>il</u> est pas diplomate

656   'He is not diplomatic'

657   f. EB : et ouais ouais non et **Rémi**, **i**l   est pas diplomate

658   And yeah, yeah, no, and Rémi he$_{cl}$ is not diplomatic

659   'Right, no, Rémi is not diplomatic at all'

660   g. SR: il est pas diplomate euh

661   'he is not diplomatic'

662   h. EB : ouais il sait pas dire les choses

663   'yeah, he doesn't know how to say things'

664   Four other continuing topics in our corpus were pronounced in the AP-LJ (CID)

665 dialogue (three by the same speaker, AP). For these cases, we note that the NP

666 is the semantically poor expression *le mec* 'the guy', which is close to a pronoun.

667 We also note that the two speakers are close friends and use a very informal and

668 colloquial register. It may be the case that the use of *le mec* instead of a clitic is

669 just a way to enhance the expressivity of the narration.

670   The other occurrences of continuing topics of our corpus (five out of eleven)

671   involve contrast. Following the proposal made by Brunetti (2009) for Italian,

672   contrast might indeed explain the presence of a dislocated full NP to express a

673   continuing topic in these utterances.[23] As we said before, we consider sentence

674   topics as contrastive whenever the proposition denoted by the utterance

675   containing the topic is compared or opposed to one or more alternative

676   propositions explicitly given in the discourse or implicitly recoverable from it. An

677   example of contrastive continuing topic is given in (13). Spk2 is a janitor in a

678   primary school and in a junior high school. Spk1 asks her whether kids ever go

679   visit her after class. Spk2 answers that young kids sometimes do, but junior high

680   school ones don't, as they are very proud.

681   (13)   [CFPP, 1260]

---

[23] AUTHOR$_i$ (2009) argues for Italian that *whenever* a continuing topic is realized as a full NP, the topic evokes an alternative set. In (i), the dislocated Prepositional Phrase *A Dante* 'to Dante' in the answer - which is a continuing topic, because the preceding question is already about *Dante* - triggers the implicature that an alternative set exists, namely that there are other relevant people in the context that may or may not receive a present.

    (i) A: A Dante, che cosa (gli) regalerai?   AUTHOR$_i$ (2009)

        'What will you give to Dante (as a present)?'

      B: **A Dante**, (gli)   regalerò   un libro.

        to Dante   to-him I-will-give a book

        'To Dante I'll give a book'

While continuing topics in the present study are not always contrastive, it is interesting to observe that contrast appears in almost *half* of the occurrences, a proportion that is much higher than that found in other topic types (see Table 3 in Section 5).

682   a. spk1 : après <u>ils</u> reviennent vous voir quelque fois <u>les gens</u>, <u>ils</u> vous

683 saluent?

684   'and later, do people come visit you sometimes? Do they say hi to you?'

685   spk2 : oui y a <u>des enfants</u> mais pas trop

686   'yes, some kids do, but not many'

687   spk1: <u>les collégiens</u>?

688   'junior high school kids?'

689   spk2 : ah non <u>les collégiens</u> ah non on vient pas voir la gardienne de

690 l'école

691   'oh no, not junior-high school kids; they don't go to visit the school janitor!'

692   spk1 : non?

693   'they don't?'

694   b. spk2 : non non **les collégiens**,   ils   sont assez fiers   d'

695 eux-mêmes

696   no no the junior high school students they are rather proud of

697 themselves

698   'No, no, junior high school students are too proud'

699 The dislocated NP *les collégiens* 'junior high school students' in (13b)

700 corresponds to the topic of the current discourse: hence, it is a continuing topic.

701 At the same time, the dislocation evokes a set of (two) opposing alternatives,

702 namely the assertion that junior high school students are proud (and therefore do

703 not go and visit the janitor) and the implicit assertion that primary school students

704 are *not* so proud (and therefore do sometimes go and visit the janitor, as

705 mentioned in (13a)).

706 4.2. *Shifted topic*

707 Left dislocation is typically associated in the literature with topic shift (see

708 Vallduvi 1992, Lambrecht 1994, de Hoop and de Swart 2000, Benincà et al.

709 2001). De Hoop and de Swart (2000), for instance, argue for English that:

710 In producing a left-dislocated construction, the speaker introduces or re-
711 introduces a discourse referent, which is made the center of attention. Such
712 introducing topics combine new information with the definition of an NP as the
713 topic of the sentence. They are called shifted topics (de Hoop and de Swart
714 2000:111).
715
716 Analogously, Lambrecht (1994) considers dislocation as a device to "establish

717 a new topic" (Lambrecht 1994:183), where 'new' "does not designate 'a topic

718 expression with a new referent' but rather the 'new coding of an active or

719 accessible referent as a topic expression' (Lambrecht 1994:353)". A confirmation

720 of this link between dislocation and topic shift comes from our data, where shifted

721 topics cover almost all of the dataset (see Table 3 in Section 5 below).

722 Two possible shifted topics are identified, depending on their status as

723 discourse referents (the variable 'discourse givenness' in Table 2): *new* shifted

724 topics and *resumptive* shifted topics. They are presented respectively in Sections

725 4.2.1 and 4.2.2.

726 4.2.1.   *New shifted topics*

727 New shifted topics are not, of course, completely new discourse referents, as

728 topics cannot by definition be completely new (cf. Barnes 1985, Ashby 1988,

729 Lambrecht 1994:353, fn 40).[24] A topic referent will always be retrievable in one

730 way or another, either from shared knowledge or by inference from the linguistic

731 context. In particular, the topic referent may present various kinds of lexical or

732 semantic relationship with a referent in the preceding context, such as kinship,

733 set-membership or other. Example (14) illustrates an instance of kinship

---

[24] They cannot be 'brand new' in Prince's (1992) sense.

relationship (father-daughter) between the topic referent and a referent in the

preceding discourse.[25] Spk4 is asked whether, among her former schoolmates,

she is still in contact with anyone who is not Caucasian (the interviewer has in

mind somebody who belongs to a family of immigrant origins). Spk4 recalls that

she is still in contact with Laurence, who is of mixed-race, but she specifies that

Laurence's father (denoted by the CILDed NP in (14b)) comes from the French

Overseas Departments and Territories (therefore, implicitly, Laurence's family is

not of immigrant origins).

  (14) [CFPP, 25-60]

     a. spk4: c'est vrai que finalement mes copines que j'ai revues,

       y en avait une qu'était… (…) elle était métisse mais (…) y a <u>Laurence</u>

qui était métisse mais

      'it's true that among those schoolmates I've kept contact with, there

was one who was (…) she was of mixed-race but (…) there's Laurence who was

of mixed-race but…'

     spk3 (spk4's mother) : elle était quand même bien métisse attends

      'hang on, she was of mixed-race alright!'

    b. spk4: oui puis attends surtout   **son père** il venait des   DOM,

c'est

       yes but wait   specifically   her father he came from the ODT

that's autre chose

another thing

       'Yes, but wait, her father was from the Overseas Departments

[25] In this and in the following examples of this section, we underline the parts of the previous

context from which the topic referent can be inferred.

757 and Territories, that's different'

758 The following example illustrates a case of shifted topic whose referent is in a

759 set-member relationship with an antecedent. The speaker is talking about her

760 negative experience with the metro system in Paris. Then she focuses on a

761 specific metro line, line 7. [26]

762 (15) [CFPP, 25]

763 a. les rares fois où j' devais prendre le métro pour X ou X raisons là c'était

764 très très chiant hein

765 'The few times I had to take the metro for this or that reason, it was a real

766 pain in the neck'

767 b. **la ligne sept** elle est atroce

768 the line seven it is horrible

769 'Line 7 is horrible'

770 Finally, example (16) is a case of a cause-consequence relationship,

771 specifically between a rain event and the presence of rain water (*la flotte*).

772 Speaker MB complains about her neighbor's behavior and then she talks about a

---

[26] Lambrecht 1987 and Barnes 1985 (following a distinction originally formulated by Hopper and Thompson 1980), distinguish between backgrounded and foregrounded utterances, namely utterances with a more parenthetical function, which assist or comment on the speaker's goal, and utterances which supply the main points of the discourse and whose topic is therefore generally maintained in the following stretch of discourse. In (15), (15b) has a backgrounded function. This is confirmed by the fact that the topic (line 7) is not maintained in the following discourse. In our study, any (sentence) topic that changes with respect to the topic of the preceding utterance is shifted, no matter whether it has a parenthetical or a main function in the discourse. It may be interesting to note, however, that most of the *new* shifted topics of our corpus have a backgrounded function.

particular episode, in which her neighbor, who lives in the apartment above hers,

swept his balcony after it rained, so that the rain water fell onto her own balcony.

(16)   [CID, MB-AC]

a. MB : ben nous le voisin en fait il nous emmerde avec le linge; il met
ses draps sur son étendage tu sais (…)

'Well, our neighbor annoys us with his laundry; he puts his sheets on
the drying rack you know (…)'

b. AC : faut lui piquer du linge comme ça après il te le fait plus

'You should pinch his laundry, so he won't do it anymore'

c. MB : mais non mais il s'en fout

'but no, he doesn't care'

d.   et ce matin   comme <u>il a plu</u>   cette nuit, il balaye   le
balcon

and this morning since it has rained this night he sweeps the
balcony

'And this morning, since it rained last night, he swept the balcony'

e.   et   tu sais **la flotte**,   elle tombe dessous

and you know the rain-water it   falls   down

'and you know, obviously the rain-water fell down'

The topic antecedent may also just have different referential properties to the

topic. In (17), the speaker is describing the inflexible attitude of school staff

supervising kids during meals in the school canteen. The antecedent *les gosses*

'the kids' in (17a) is generic, while the dislocated NP *un gamin* 'a child' in (17c)

(which basically reformulates (17b)) refers to an instantiation of the generic

referent.

798 (17) a. y en a une qu'a un sifflet et qui siffle <u>les gosses</u> ;   [CID, IM]

799   'One of them has a whistle and whistles to the kids'

800   b. dès qu'<u>ils</u> parlent [...] elles <u>les</u> mettent au coin tout le repas ;

801   'whenever they speak [...] they make them stand in the corner during

802 the whole meal'

803   c.   parfois   **un gamin**, il est puni   au coin

804   sometimes a kid   he is punished to-the corner

805   'Sometimes, a kid is sent to the corner'

806   As we will see in Table 3 below, most of the topics in our dataset are shifted.

807 Among shifted topics, most of them are new (134) and among new shifted topics,

808 only a few (22) are contrastive. A typical example is (18) below. The dislocated

809 NP *ma mère* 'my mother' is a new shifted topic, as it has not been mentioned

810 before and is inferable from the context by means of the NP *vos parents* 'your

811 parents'. The NP is also contrastive, because utterances (18b) and (18c) denote

812 two alternative propositions: the former – which contains the dislocated NP –

813 talks about the speaker's mother's profession; the latter – a canonical subject-

814 predicate construction, talks about the speaker's father's profession.[27]

815 (18) a. Spk1: C'est c'que faisaient <u>vos parents</u> ?   [CFPP, MoH]

816   'Was it what your parents did?'

817   b. Spk2: Oui **ma mère**, elle travaillait dans l' vingtième rue d'la Réunion,

818   yes my mother she worked in the twentieth street of the

---

[27] Note that (18) is different from (15). In the latter example, the topic referent (the metro line 7) is a member of a set (the set of all Paris metro lines), but the topic is not contrastive, because there is neither explicit nor implicit comparison/opposition in the discourse between the proposition denoted by (15b) and other propositions predicating about other metro lines.

819   Reunion

820   'Yes: my mother worked in the 20th district, in Reunion street

821   c.   [...] et mon père était artisan

822   ' [...] and my father was a craftsman'

823   Another example of contrast is (19), where the speaker – a schoolteacher –

824   compares her age and that of her pupils when she had her first teaching

825   experience.

826   (19) a. Je suis partie directement en fin d'études   [CFPP, 6X]

827   'I left [=started teaching] right after my studies'

828   b. j'avais dix-huit ans

829   'I was eighteen'

830   c. **les gosses**, elles en   avaient quatorze

831   the kids   they of-them had   fourteen

832   'and the kids were fourteen'

833   The current topic is the speaker (see (19a,b)), but a topic shift occurs in (19c),

834   where *les gosses* 'the kids' – whose reference is inferable from previous

835   discourse and denotes the pupils of the speaker's first teaching experience –

836   becomes the sentence topic. The newly shifted topic is also contrastive, because

837   utterances (19b) and (19c) form a set of (two) alternative propositions: what is

838   asserted about the speaker's age is compared with what is asserted about the

839   pupils' age.

840   4.2.2.   *Resumptive shifted topic*

841   Other occurrences of topic shift involve a *given* referent: either its antecedent is

842   the topic of the preceding discourse, but not of the immediately preceding one, or

843   it is the focus or part of the focus of the preceding utterance. There are 96

844   resumptive shifted topics in our dataset, among which 19 are contrastive.

845   In the first case of topic resumption, the referent is promoted to topic status

846   after topic continuity has been interrupted by a topic shift. Left dislocation seems

847   to be a typical device to express this kind of function in various languages (Givón

848   1983, Lambrecht 1994). Lambrecht (1994) calls a dislocation a 'topic promoting

849   construction' (Lambrecht 1994:177) and illustrates this property with the English

850   example below (taken from Givón 1976). The referent of *a wizard* is introduced

851   for the first time in the first sentence and it is the topic of the subsequent

852   sentences ((20b) and (20c)). Then topic continuity is interrupted by the

853   introduction of other topics (see (20d-f)). The wizard is eventually promoted again

854   to topic status with the dislocation in (20g).

855   (20)   a. Once there was <u>a wizard</u>.   Lambrecht (1994 :177)

856   b. <u>He</u> was very wise, rich,

857   c. and was married to a beautiful witch.

858   d. They had two sons. The first was tall and brooding, he spent his days

859   in the forest hunting snails,

860   e. and his mother was afraid of him.

861   f. The second was short and vivacious, a bit crazy but always game.

862   g. Now **the wizard**, he lived in Africa.

863   An example from our corpus is given in (21). The speaker is telling a story

864   involving a conflict between her cat and her neighbor: her neighbor scared her

865   cat away from his door by hitting the cat with a broom. The speaker then keeps

866   talking about her neighbor's bad habits, and eventually resumes the story about

867   the cat in (21b). About 13 exchanges separate the last mention of the cat in (21a)

868   and its 'resumption' as topic in (21b).[28]

869   (21) [CID, MB]

870   a. Donc je lui ai dit: "Ben très bien tu lui as fait peur un bon coup; le chat

871   il est pas idiot il remontera plus de chez toi de sitôt tu vois euh", mais bon… Il

872   laisse sa porte ouverte… qu'est-ce qu'il s'en fout, ah ouais, et non mais parce

873   que lui il habite au premier mais tout l'escalier il a décrété que c'était à lui (...)

874   'So I told him *[the neighbor, ndr]*: "Very well, you really scared him *[=the*

875   *cat]* quite a lot: the cat is not stupid, he won't climb up again to your place so

876   soon, you know", but well… He leaves his door open…He really doesn't care! Oh

877   yeah, because he lives on the first floor, but he has decided that the whole

878   staircase belongs to him (…)'

879   b. et donc **le chat**, il était dans l'escalier, il   a dû   le choper

880   and so the cat   he was on the stairs   he has must him to catch

881   'and so the cat was on the stairs, he must have caught him'

882   An example where the resumptive shifted topic is also contrastive is given in

883   (22) below. Spk2 is talking about different shops that there used to be in her

884   neighborhood when she was a child. In particular, she remembers a married

---

[28] For cases like (21), it is clear that the speaker is resuming a topic that was abandoned and

then recovered *within the same discourse segment*. When a speaker talks about a referent that

was already mentioned, but in a completely different point of the text, we did not consider it as a

case of topic resumption. For instance, in one interview of the CFPP corpus, a speaker talks

several times about her father, remembering different episodes of her family life. Such different

episodes are considered as separate discourse segments; therefore, the father is not counted as

a resumptive topic each time that it is mentioned, but only when it is resumed *within* the same

discourse segment.

885 couple of newspaper sellers, Mr and Mrs Luca, whom she calls *le père Luca*

886 'father Luca' and *la mère Luca* 'mother Luca'. She goes on mentioning other

887 shops and then, at Spk1's request, she gives a description of Mrs Luca,

888 explaining why she found her nasty. Eventually, she mentions Mr Luca again

889 (topic resumption) by means of the CILD in (22b). The topic *le père Luca* 'father

890 Luca' is contrastive, because the assertion that Mr Luca was nice is the opposite

891 of what was previously asserted about Mrs Luca, namely that she was not nice.

892 (22) [CFPP, Mo-67]

893 a. Spk2: y avait un marchand d' journaux c'était chez Luca; la mère

894 Luca elle était pas gentille, monsieur Luca qu'était très gentil; et après y avait l'

895 boulanger c'était chez Berlère et tout au bout y avait une vieille épicerie (…)

896 'There was a newsagent's, Luca's; mother Luca was not nice, and

897 Mr Luca who was very nice; and then there was the baker, Berlère's, and at the

898 far end there was an old grocery (…)

899 spk1 : la mère Luca qu'est-ce qu'elle avait de méchant ?

900 'mother Luca, what was nasty about her?'

901 spk2 : ben elle était méchante: elle avait des grands sourcils (…) avec des

902 yeux mauvais; elle faisait un peu la trouille (…)

903 'Well, she was nasty: she had bushy eyebrows (…) with mean eyes; she

904 was a bit scary (…)'

905 b. mais par contre **le père Luca** il était gentil

906 but on the contrary the father Luca he was kind

907 'but on the contrary father Luca was nice'

908 In the second case of topic resumption mentioned at the beginning of this

909 section, an (identical) antecedent is given in the previous discourse – generally,

910 in the immediately preceding utterance – that does not have a topic function yet.

911 Consider (23). The speaker – who works in a primary school – is talking about a

912 first-grade kid who behaves badly and likes to provoke older kids. The

913 antecedent of the topic *les grands* 'the older kids' in (23b) is in the preceding

914 utterance ((23a)): there, the referent is introduced as part of the *focus*, which

915 means that it is not the topic of the current discourse. It is only in the left

916 dislocated position in (23b) that the referent gets promoted to topic status.

917   (23)   [CID, AC]

918   a. …un CP six ans tu vois (…) tu ne peux pas le lâcher de l'œil, il fait

919 que des conneries et il agresse <u>les grands</u> (…)

920   '…a first grader, six-years old, you know (…) you have to constantly

921 keep your eyes on him, he does stupid things all the time and he attacks the

922 older ones (…)'

923   b. Et comme il est petit, **les grands** ils   le poussent un peu tu vois   ou

924 ils sont plus patients

925   and since he is little the older ones they him push   a bit   you see or

926 they are more patient

927   'and since he is little, the older kids just push him a bit, you know, they are

928 more patient'

929   Below, another example of this kind of topic is given, where the topic

930 additionally has a contrastive interpretation. Spk1 has asked Spk3 whether the

931 behavior of dogs in her neighborhood annoys her. Spk3 answers by describing

932 the different kinds of dogs that she may encounter in her neighborhood.

933   (24)   [CFPP, 6X]

934   a. Spk3 En plus les gens qu'ont des chiens agressifs ils sont agressifs

935 aussi

936   'And besides, those who own aggressive dogs are aggressive too'

937   Spk1: et vous avez ça ?

938   'And do you have that?' ([= aggressive dogs in the neighborhood,

939 *ndr*]

940   Spk3: ah oui

941   'Oh yes'

942   Spk1: vous est arrivé d'avoir peur…

943   'Have you ever been scared…'

944   Spk3 : ah oui oui. Y a des chiens euh dans la banlieue là qui sont…

945 oui, qui

946 sont pas très sympas oui, ou dans le métro même, y en a hein quelques-uns …

947 et puis y a <u>beaucoup de petits chiens</u> là;

948   'Oh yes, sure. There are dogs in the suburbs that are… yes, that are not

949 very nice yes, or even in the metro, there are some… and then there are many

950 small dogs;'

951   b. **ceux-là** ils aboient ils font   un peu d'bruit mais...

952   those   they bark   they make a little of noise but

953   'Those, they bark, make some noise, but…'

954   c. (…) ça m'fait pas peur, mais ça n'est pas effrayant

955   'I'm not scared, that's not scary'

956 The referent of the dislocated pronoun *ceux-là* 'those' in (24b) is introduced as a

957 new referent in the immediately preceding presentational *y a* construction (*y a*

958 *beaucoup de petits chiens* 'there are many small dogs') and is then used as topic

959   in the dislocation construction.[29] The topic is contrastive because the assertion

960   that these small dogs bark (but are harmless) is opposed to what was previously

961   asserted about other dogs, namely that they are aggressive and can be scary.

962   As we said, the distance between the dislocated NP and its antecedent, in this

963   particular case of resumption, can be quite short (usually, the antecedent is in the

964   preceding utterance). Still, since the topic referent is not the current discourse

965   topic, the interlocutor does not expect it to be the topic of the upcoming discourse.

966   That's why we cannot consider this as a case of topic continuity, but as another

967   kind of topic shift.

968   4.3. Summary of the pragmatic annotation

969   The scheme in Figure 6 summarizes the pragmatic annotation. The scheme

970   illustrates the classification obtained by combining the different pragmatic

971   properties that can be associated with a ClLDed topic.
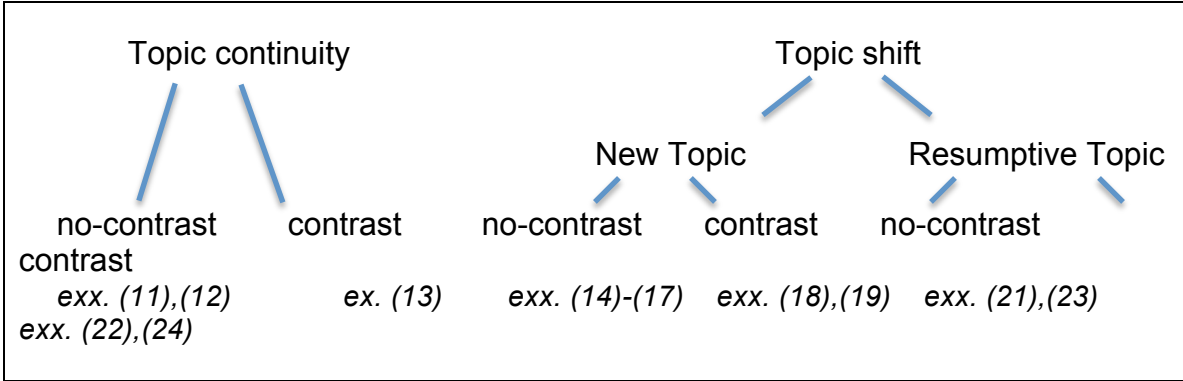
972   Figure 6. Classification resulting from the annotation of topics with respect to topic continuity,
973   discourse givenness, and contrast. For each possible combination of these properties, the
974   numbers of the relevant examples are given.
975
976   In the following section, after reporting the number and frequency of topics per

977   category, we will present the results of the statistical analysis concerning the

978   relationship between topics and prosody.

---

[29] According to Lambrecht (1994:177) presentational constructions are "topic promoting" devices,

namely they introduce a new referent so that it can be used as topic in subsequent discourse.

979   5.   ANALYSIS AND RESULTS

980   Table 3 presents the number of contrastive and non-contrastive continuing, new

981   shifted and resumptive shifted topics in our corpus, as well as their relative

982   frequencies.

	No-contrast	(%)	Contrast	(%)	Total	(%)
Continuing	6	(2.49)	5	(2.07)	11	(4.56)
Shifted (new)	112	(46.47)	22	(9.13)	134	(55.6)
Shifted (res)	77	(31.95)	19	(7.88)	96	(39.83)
Total	195	(80.9)	46	(19.1)	241	(100)

983   Table 3. Number of items of continuing and shifted topics associated or not to contrasting
984   alternatives. Relative frequencies are given in brackets.
985
986   Interestingly, this distribution is similar to the one obtained by Ashby in his

987   work on the pragmatics of French left and right dislocations (Ashby 1988), in

988   which the data were also taken from a spontaneous speech corpus. Ashby's

989   classification is partially different from ours, so some adjustments must be made

990   in order to compare the two datasets. If we sum Ashby's subject and object

991   CILDed NPs (see Ashby 1988: 212, Table 3a), we obtain 12 occurrences (7.4%)

992   of strictly given referents, 99 (60.7%) of new (inferred) referents, and 50 (33.7%)

993   of textually given referents. 'Strictly given' referents are "those where an

994   anaphoric pronoun apparently would have sufficed to identify the referent"

995   (Ashby 1988:211), so they clearly correspond to our continuing topics. As for

996   'new (inferred)' referents, Ashby clarifies the meaning of 'new' : "it is generally

997   acknowledged that neither left- nor right- dislocations can be used with referents

998   that are 'brand-new' […] or 'unidentifiable' […]. That is, there must be a 'semantic

999   link' […] between the proposition containing the dislocation and the preceding

1000   discourse" (Ashby 1988:215). This is precisely the case of our own new shifted

1001   topics: their referent, although new in the discourse, is 'inferentially accessible',

1002   as Ashby would put it. Finally, 'textually given' referents correspond to what we

1003   call resumptive shifted topics, because Ashby defines them precisely as referents

1004   that "may have been mentioned earlier in the discourse, but are not the

1005   immediate discourse topic".Ashby (1988:217, Table 5a) also counts 79 (21%)

1006   cases of contrast (that is, where a set of alternatives is recoverable in the

1007   discourse, and of which the topic itself is a member), and 279 (73%) occurrences

1008   of topic shift. We see in Table 3 that contrastive topics represent, similarly, about

1009   19% of our dataset. If we sum new and resumptive shifted topics that are not

1010   contrastive, we obtain a similar frequency to Ashby's for shifted topics too, that is

1011   78% (189 utterances).

1012   Concerning the relationship between these topic types and prosody, as

1013   explained in the introduction, the goal of our paper was to assess the existence

1014   of an effect of such topics on the degree of the prosodic boundary strength. In

1015   order to pursue this goal, Linear Models were carried out.[30] For computational

1016   reasons, it was not possible to enter in the model all the predictors coded in

1017   Table 1. In order to limit potential variation due to syntactic reasons (the

1018   grammatical function of the dislocated NP and the root/subordinate clause

1019   distinction) the analysis was made only on subject NPs and on non-embedded

1020   sentences. Finally, we excluded continuing topics (11 utterances). In total, the

1021   analysis was therefore made on 159 utterances.

1022   Prominence score (as a linear factor) was entered as the outcome, the

1023   number of syllables (also as a linear factor), topic type (new shifted

1024   topic/resumptive shifted topic), contrast (contrastive/non contrastive) and the

1025   interaction between topic type and contrast were entered as predictors. Corpus

---

[30] Statistical calculations were conducted with the software R, v.3.1.2. (R Development Core
Team, 2015).

1026 (CID/CFPP), items (the head noun of the NP) and speakers were entered as

1027 random effects. P-values, set at a level of 0.05, were obtained by likelihood ratio

1028 tests of the model against the model without the fixed effects in question.

1029   The results of the model revealed that the interaction between topic shift and

1030 contrast was significant (x^2(1)= 12.2870, p<0.05). Post-hoc tests showed that

1031 new shifted topics differ significantly according to the presence or not of

1032 contrasting alternatives (p < 0.05): the boundary score is higher in the contrastive

1033 context than in the non-contrastive one. Also, a significant difference was found

1034 between a resumptive and a new shifted topic in the non-contrastive context (p <

1035 0.001), the former topic type being more prominent than the latter. Figure 7

1036 illustrates these differences:
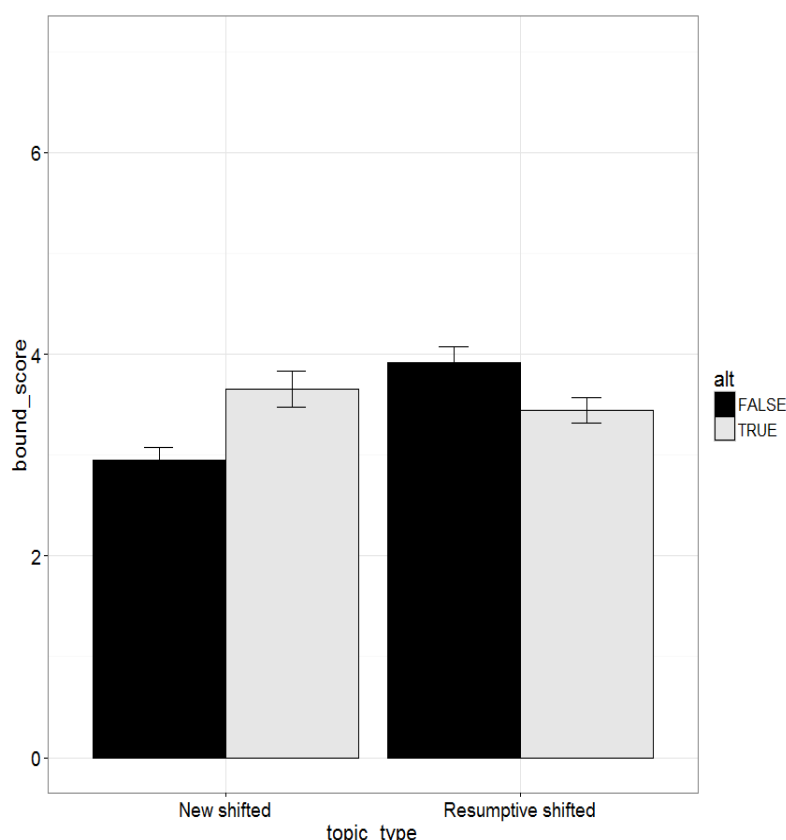
1037
1038   Figure 7. Effect of the discourse properties of the topic on prosodic boundary score. The
1039 columns show the average score for new shifted topics and resumptive shifted topics as a
1040 function of the presence of contrastive alternatives (alt). Error bars represent the standard error
1041 from the mean.
1042
1043

1044 6. DISCUSSION

1045 Summarizing the results presented in the preceding section, our statistical
1046 analysis reveals:
1047
1048 - a significant effect of contrast on new shifted topics, contrastive ones being

1049 more prominent than non-contrastive ones;

1050 - a significant difference with respect to the pragmatic properties of the topic,

1051 resumptive shifted topics having a stronger prosodic boundary than new shifted

1052 ones; this difference, however, only appears to be significant among *non-*

1053 *contrastive* shifted topics.

1054 Given these results, the following remarks can be made concerning the

1055 prosody/pragmatic interface. As for the given/new dimension, namely the mere

1056 presence of an identical topic antecedent in previous discourse, the interaction

1057 with prosody *does* occur, but in an unpredictable way with respect to previous

1058 assumptions (see discussion in Section 1.2). According to our results, the

1059 dislocated element displays a stronger prosodic break when the referent is *given*,

1060 not new. Note however that this effect is only visible, in our data, with shifted

1061 topics, as we eventually excluded continuous topics from the analysis. Therefore,

1062 we cannot be certain that the effect is due to givenness by itself. Recall that in

1063 resumptive shifted topics the shift occurs either by resuming a topic referent from

1064 a previous discourse segment and making it topic again, or by taking a non-

1065 topical referent from the discourse and promoting it to topic.[31] We may suspect

---

[31] In a previous stage of this work, a model was tested where these two kinds of resumptive

topics were considered as separate factors. Each of them displayed a significantly stronger

prosodic boundary than new shifted topics. Given such results, and given that the two topics

share the same values with respect to the features of Table 1 (both are *shifted* but *given*), we

1066 that the interpretation of our results is then that prosody helps to promote a

1067 referent to topic status whenever the referent, although present in previous

1068 discourse, is not expected to be the upcoming topic.

1069   As for the fact that the difference between new and resumptive shifted topics

1070 only concerns *non*-contrastive topics, this can be related to the other result of our

1071 analysis, namely that there is a significant difference between contrastive and

1072 non-contrastive new shifted topics. The boundary score is higher with contrastive

1073 new shifted topics: such a higher score presumably blurs the difference between

1074 new and resumptive topics, and in particular the higher score of the latter.

1075   We may wonder why the effect of contrast on prosody should only appear

1076 within new shifted topics and not with resumptive ones. If we look more carefully

1077 at the data, we see that within resumptive shifted topics, in most cases, the

1078 contrasting members of the set have already been introduced in the discourse.

1079 For instance, in (25), the contrasting alternatives (the speaker and the speaker's

1080 children) are already introduced in (25c), hence before contrast is expressed in

1081 (25d) and (25e).

1082   (25)   [CFPP, 70]

1083   a. spk1 : le passage de la de la maison à l'extérieur avant le mariage [...]

1084 ça s'est fait comment

1085   'the passage from home to outside, before marriage [...] how is it

1086 done?'

1087   b. spk2 : ah pour mes enfants

1088   'for my children?'

1089   c. spk1 : <u>pour vous et pour eux</u> ?

---

opted for a model where they belong to the same class.

1090   'for you and for them?'

1091   d. spk2: <u>pour moi</u> c'était hors de question bien évidemment [...]

1092   'for me, it was out of question, of course [...]'

1093   e. ben **mes enfants** ils se sont installés chez eux quand ils m'ont dit

1094 qu'ils voulaient s'installer chez eux

1095   '(as for) my children, they went living by themselves when they told me

1096 that they wanted to go living by themselves'a.

1097   On the contrary, in the case of new shifted topics, it is often unexpected that

1098 the newly introduced topic is in contrast with some other referent. Such

1099 unexpectedness could motivate a stronger prosodic marking. Consider (26). The

1100 speaker is talking about her childhood and the fact that she used to go to do

1101 grocery shopping for her mother. She says that she loved to do that, and then

1102 she introduces her sister (for the first time in this discourse segment) and

1103 contrasts her attitude toward grocery shopping with her own attitude. The very

1104 presence of an alternative set only becomes evident once the second sentence

1105 is uttered.[32]

1106   (26) a. <u>Moi</u>, j'adorais ça   [CFPP, 1453]

1107   '(As for) me, I loved that'

1108   b. ah **ma sœur**, elle était plus timide plus réservée

1109   ah my sister she was more shy more reserved

1110   'my sister was more shy, more reserved'

---

[32] Notice that in this example the contrast is inferred: the utterance states that the speaker's sister was more shy, which implies that she did not like to interact with sellers, and therefore to do the shopping.

1111   Speaking in more general terms, contrastive contexts are extremely varied in

1112   our data. First, alternatives may be comparing or opposing ones. In (18) or (19),

1113   for instance, the target utterance, which is a member of an alternative pair, is

1114   simply compared to its alternative, while in (25) or (26), as well as (27) below, the

1115   alternatives are opposed. This is particularly clear in (27), where the same verb is

1116   used in both utterances, but with opposite polarity (cf. Myhill and Xing 1996).

1117   (27) **celui-là**, je l' connais **les autres**, je les connais pas   [CFPP, MoF]

1118   that-one I it know the others I them know not

1119   'that one I know, but the other ones I don't'

1120   The second source of variation is given by the different degrees of explicitness

1121   of the contrastive interpretation in our data. In most contrastive utterances where

1122   the alternatives are opposed, the opposition is implicit, namely the proposition

1123   that negates the assertion of its alternative must be inferred from the discourse.

1124   We have seen this in (26), where the opposition between the speaker and the

1125   speaker's sister (she liked to do grocery shopping, her sister didn't) is inferable

1126   by what is predicated about the sister (she was shy), as explained in footnote 32.

1127   The level of explicitness of contrast also varies depending on the presence or not

1128   of a lexical contrastive marker (such as *plus* 'more' in (26) or *par contre* 'on the

1129   contrary' in (22)). Finally, it may depend on the position of the contrasting

1130   alternative, which may precede or follow the utterance containing the topic: when

1131   the alternative utterance *precedes* the one containing the topic (which is

1132   generally the case in our data), the opposition of the latter with the former is

1133   obviously clearer. We may hypothesize that the strength of the prosodic

1134   boundary of the dislocated element varies depending on these factors, that is on

1135   the presence or not of an opposition between alternatives, and the degree of

1136 explicitness of such an opposition. More precisely, the relationship between

1137 contrast and prosody might be gradual: the less explicit the contrast is, the

1138 stronger the prosodic marking.[33] In order to check this hypothesis, our dataset

1139 should contain many more contrastive utterances than it actually does. In fact, a

1140 specific study focusing on contrast alone should be pursued, where different

1141 degrees of transparency of contrast are explicitly taken into account. This,

1142 however, goes beyond the scope of the present work and will be left to future

1143 research.

1144   If we consider the pragmatic annotation alone, independently from its prosodic

1145 correlates, a qualitative analysis of our data may contribute to a better

1146 understanding of the pragmatic function of a CILDed NP in French. The fact that

1147 most of our topics are shifted and only few are continuing seems to be consistent

1148 with what is said in the literature (see references and discussion in Sections 4.1

1149 and 4.2), namely on the one hand, that left dislocation is mainly used as a topic

1150 shifting device, and on the other, that continuing topics are not generally realized

1151 as CILDed elements.[34] The data are also consistent with results obtained in

---

[33] See for instance Braun and Ladd (2003) for results showing a gradual prosodic marking of

contrastive topics in German. See Repp (2016:6) (and references quoted therein) for a discussion

of contrast as a gradable phenomenon.

[34] Of course, while we can safely assume that most of the remaining continuing topics in our text

are clitics (see discussion in Section 4.1 and references cited therein), the proportion of

continuing topics that are full NPs but *not* left dislocated (that is, either right dislocated or not

dislocated at all) is not known. Without such information, three explanations for the low number of

continuing topics among CILDs are conceivable: either they are few because *left* dislocated, or

1152   previous work on CILDs in spontaneous spoken French, such as Ashby (1988)

1153   (see Section 5). As for contrast, we have seen that topics belonging to an

1154   alternative set account for 19.1% of the total (see Table 3). This percentage

1155   seems rather small, given the association that is commonly made between

1156   contrast and left dislocation (see e.g. Chafe 1976, Lambrecht 1994, Arregi 2003,

1157   Molnár 2002). A possible explanation for this mismatch is the following. The

1158   constructions studied in previous literature (in different languages) are generally

1159   *object* dislocations, while in our corpus, dislocated subjects constitute the largest

1160   set. Contrast could be a function that is more often associated with object

1161   dislocation. This is what Brunetti (2009) claims for Italian and Spanish. Brunetti's

1162   argument is that an object is not a prototypical topic in most cases, as it often

1163   lacks agentivity, animacy, and other prototypical topic properties. Therefore, an

1164   object dislocation (which promotes an object to topic position) only occurs in

1165   specific contexts that precisely demand that the object be the topic. One of these

1166   contexts is the contrastive one, hence the frequent contrastive interpretation of

1167   dislocated objects. A partial confirmation of Brunetti's claim comes from the

1168   difference, in our data, between the number of object topics and subject topics

1169   associated with contrast: contrastive object topics constitute about 35% of all

1170   objects (13 out of 37 items), while contrastive subject topics are only 16.6% of all

---

they are few simply because they are *dislocated*, or, finally, they are few because they are full

NPs (and not clitics).

1171 subjects (34 out of 204 items). The number of items is not very large, but these

1172 results can be a starting point for further and more accurate research.[35]

1173   What we think is important to highlight here, to conclude, is that an analysis of

1174 a spontaneous speech corpus – which allows for a study of the *frequency* of a

1175 construction as well as its different *discourse functions* – can be crucial for a

1176 better understanding of the construction itself.

1177   7.   CONCLUSIONS

1178 The goal of our paper was to test, by means of a quantitative as well as

1179 qualitative analysis of data from French spontaneous speech, the relationship

1180 between the prosodic boundary strength at the right edge of a dislocated NP and

1181 the different discourse properties of the NP itself. Our results allow us to suggest

1182 that prosodic boundary strength variations are used to signal the 'promotion' of a

1183 referent to topic status, when such a referent is already present in the discourse

1184 but not yet, or not any more, the topic. In other words, prosody does not signal

1185 that a topic is new, but rather that a known referent is *newly* promoted to topic.

1186 This conclusion, if correct, suggests that the role of prosodic boundary strength

1187 variations is related to discourse structuring: prosody seems to help the hearer to

1188 retrieve the speaker's intentions/strategy concerning the structuring of his/her

1189 discourse.

1190   As for contrast, our results show the existence of a stronger prosodic

1191 boundary in the presence of contrasting alternatives, but only with new shifted

1192 topics. We have suggested an explanation of this result based on the fact that, in

---

[35] In a recent work based on the same CFPP2000 corpus, Riou and Hemforth (2015) confirm these results by finding a significantly higher frequency of contrastive topics among object dislocations than among subject ones.

1193 our data, contrast of a newly introduced topic with some other referent is

1194 generally unexpected, while this is not the case with a resumptive topic. We must

1195 further add that our work is only concerned with the strength of the prosodic

1196 boundary. It is possible that by taking into account the intonational patterns of the

1197 NP, our results on contrast would be more evident. Indeed, as we have

1198 mentioned at the beginning of this paper, it is often claimed that contrast is

1199 encoded by a specific tonal realization (different pitch accents and tonal

1200 alignment differences, see for instance Braun 2006, among others). As for

1201 French, it has recently been shown that contrast may be encoded by a late dip

1202 (see Turco & Delais-Roussarie 2014).

1203 From the mere count of the contrastive utterances in our dataset, we can

1204 finally point out that the presence of contrasting alternatives is not an inherent

1205 property of CILDed NPs in French. Indeed, only a small percentage of dislocated

1206 NPs (especially subjects) involves contrast.

1207 This study deserves to be completed in several directions. Concerning the

1208 prosodic analysis, future work should take at least two paths. First, other

1209 prosodic cues should be tested: we did not take into account intensity values or

1210 post-boundary syllabic lengthening, which have been considered as significant in

1211 previous studies dealing with syntactic aspects of clitic left dislocations (Guilbault

1212 1995, Astésano et al. 2008).[36] Secondly, perception tests should be conducted,

1213 in order to assess the relevancy of the differences in boundary strength found

1214 between new and resumptive shifted topics, and more precisely to check whether

1215 the discourse property related to resumptive shifted topics – the combination of

---

[36] We admit that this is a limit of a study based on spontaneous speech: more controlled data are

needed to take these elements into account.

1216 givenness and topic shift – is phonologically coded in the system of French

1217 speakers.

1218   As for the pragmatic analysis, the small amount of left dislocated continuing

1219 topics forced us to exclude them from the statistical model. A larger dataset

1220 would allow us to verify whether the prosodic effect that we find in resumptive

1221 shifted topics is indeed due, as we suspect, to the combination of givenness and

1222 shift, and not to givenness alone. Finally, our pragmatic annotation only focuses

1223 on the properties of the topic referent in the *preceding* discourse: whether it is

1224 given or not and whether it is the topic or not of the preceding discourse.

1225 However, the pragmatic function of a left dislocation has also been linked to topic

1226 persistence in the *succeeding* discourse (see Lambrect 1984, Barnes 1985,

1227 Gregory and Michaelis 2001). What remains to be tested is therefore whether

1228 such a property may be related in some way to the prosodic realization of the

1229 dislocated NP.

1230

1240

REFERENCES

Anagnostopoulou, Elena 1994. *Clitic Dependencies in Modern Greek*. PhD
dissertation, University of Salzburg.

Anagnostopoulou, Elena 1997. 'Clitic Left Dislocation and Contrastive Left
Dislocation'. In E. Anagnostopoulou, H. Van Riemsdijk, and F. Zwarts (eds),
Materials on Left Dislocation, Amsterdam: Benjamins.

Aoun, Joseph E. and Elabbas Benmamoun 1998. 'Minimality, Reconstruction,
and PF Movement', *Linguistic Inquiry*, 29-4, 569-597.

Apothéloz, Denis 1997. 'Les dislocations à gauche et à droite dans la
construction des schématisations'. In Miéville and Berrendonner (eds)
Logique, discours et pensée. Mélanges offerts à Jean-Blaise Grize. Berne:
Peter Lang, 183-218.

Arregi, Karlos 2003. 'Clitic left dislocation is contrastive topicalization'. In
*Proceedings of the 26th Annual Penn Linguistics Colloquium, Penn Working
Papers in Linguistics* 9.1:31-44. Philadelphia: Penn Linguistics Club.

Ashby, William J. 1988. 'The syntax, pragmatics, and sociolinguistics of left- and
right dislocations in French'. Lingua, 75, 203-229.

Astésano, Corinne 1999. *Rythme et discours : invariance et sources de
variabilité des phénomènes accentuels en français*, PhD dissertation,
Université d'Aix-Marseille I.

Astésano, Corine, Espesser, Robert, Rossi-Gensane, Nathalie, 2008. Quelques
cas particuliers de détachement à gauche - ou la prosodie à l'aide de la
syntaxe _. Proceedings of Journées d'Etude sur la Parole (JEP), Jun 2008,
Avignon, France. pp.109-112.

Bally, Charles 1944. *Linguistique générale et linguistique française*. Bern:

1266   Francker.

1267 Barnes, Betsy K.1985. *The Pragmatics of Left Detachment in Spoken Standard*

1268   *French*, Amsterdam: Benjamins.

1269 Benincà, Paola, Frison, Lorenza, and Giampaolo Salvi 2001. 'L'ordine delle

1270   parole e le costruzioni marcate'. In L. Renzi, G. Salvi and A. Cardinaletti

1271   (eds), *Grande grammatica italiana di consultazione*, Bologna: Il Mulino, Vol.

1272   1, 115-225.

1273 Bentzen, Kristine 2009. 'Embedded root phenomena, assertion, presupposition,

1274   and main point of utterance' paper presented at the workshop Root

1275   Phenomena, Centre for General Linguistics (ZAS), Berlin, Germany,

1276   September 2-4, 2009.

1277 Berrendonner, Alain 1993. 'Sujets zéro'. In *Complétude et incomplétude dans les*

1278   *langues romanes et slaves. Actes du VI colloque international de Cracovie*.

1279   Cracovie: WSP, 17-45.

1280 Berrendonner, Alain 2007. 'Dislocation et conjugaison en français'. *Cahiers de*

1281   *praxématique*, *28*, 85-110.

1282 Berrendonner, Alain 2015. 'Constructions disloquées'. *Encyclopédie*

1283   *Grammaticale du Français* (http://encyclogram.fr).

1284 Bertrand, Roxane and Robert Espesser 1998.Prosodie et discours rapporté: La

1285   mise en scène des voix. In J. Verschueren (ed.), *Pragmatics in 1998:*

1286   *Selected papers from the 6th International Pragmatics Conference*. Antwerp:

1287   Int. Pragmatics Association, vol.2, 45-56.

1288 Bertrand, Roxane, Blache, Philippe, Espesser, Robert, Ferré, Gaêlle, Meunier,

1289   Christine, Priego-Valverde, Béatrice and Stéphane Rauzy 2008. 'Le CID –

1290 Corpus of Interactional Data - Annotation et Exploitation Multimodale de

1291 Parole Conversationnelle'. *Traitement Automatique des Langues, 49/3*, 1-30.

1292 Bianchi, Valentina and Mara Frascarelli 2010. Is topic a root phenomenon?.

1293 *Iberia* 2 . 43-88.

1294 Blanche-Benveniste, Claire 2005. 'Detachment constructions'. In K. Brown (ed.),

1295 *2nd Encyclopedia of Language and Linguistics*, Vol. 3: 477-485, Elsevier.

1296 Blasco-Dulbecco, Mylène 1999. *Les dislocations en français contemporain.*

1297 *Étude syntaxique*. Paris: Champion.

1298 Boersma, Paul and David Weenink, 2014. 'Praat: doing phonetics by computer

1299 (v. 5.3)'. Amsterdam: Department of Language and Literature, University of

1300 Amsterdam, http://www.praat.org/.

1301 Branca-Rosoff, Sonia, Fleury, Serge, Lefeuvre, Florence and Mat Pires. 2009.

1302 *Discours sur la ville. Corpus de Français Parlé Parisien des années 2000*

1303 *(CFPP2000). http://ed268.univ-paris3.fr/CFPP2000/.*

1304 Braun, Bettina and Ladd, D. Robert 2003. 'Prosodic correlates of contrastive and

1305 non-contrastive themes in German'. In *Proceedings of the 8th European*

1306 *Conference on Speech Communication and Technology*. Geneva,

1307 Switzerland. 789-792.

1308 Büring, Daniel 1997. *The Meaning of Topic and Focus - The 59th Street Bridge*

1309 *Accent*, London: Routledge.

1310 Büring, Daniel 2003. 'On D-Trees, Beans, and B-Accents', *Linguistics and*

1311 *Philosophy*, 26 (5):511-545.

1312 Butt, Miriam and Tracy H. King, 1997. 'Null Elements in Discourse Structure', in

1313 Karumuri Venkata Subbarao (ed.), *Papers from the NULLS Seminar*.

1314 Chafe, Wallace L. 1976. 'Givenness, contrastiveness, definiteness, subjects,

1315   topics and point of view'. In Charles N. Li (ed), *Subject and Topic*. 27–55.

1316   New York: Academic Press.

1317 Cinque, Guglielmo 1983. ''Topic' Constructions in Some European Languages

1318   and 'Connectedness''. In K. Ehlich and H. van Riemsdijk (eds)

1319   *Connectedness in Sentence, Discourse and Text*, Tilburg: KBU.

1320 Cole, Jennifer, Yoonsook, Mo, and Baek Soondo 2010. 'The role of syntactic

1321   structure in guiding prosody perception with ordinary listeners and everyday

1322   speech'. *Language and cognitive processes*, 25/7: 1141-1177.

1323 Collier, Renk, de Pijper, Jan Roelof, and Angelien Sanderman, 1993. 'Perceived

1324   prosodic boundaries and their phonetic correlates', *Human language*

1325   *technology: Proceedings of a Workshop held at Plainsboro, NJ, USA*, March

1326   21-24, 1993, San Francisco: Morgan Kaufmann Publishers, 341-345.

1327 Culbertson, Jennifer 2010. 'Convergent evidence for categorical change in

1328   French: from subject clitic to agreement marker'. *Language*, 86(1), 85-132.

1329 Culbertson, Jennifer and Geraldine Legendre 2008. 'Qu'en est-il des clitiques

1330   sujet en français oral contemporain?' In *Actes du 1er Congrès mondial de*

1331   *linguistique française*. Paris, France.

1332 De Cat, Cécile 2007. *French Dislocation. Interpretation, Syntax, Acquisition*,

1333   Oxford: Oxford University Press.

1334 Delattre, Pierre 1938. 'L'accent final en français : accent d'intensité, accent de

1335   hauteur, accent de durée'. *French Review*, 12: 141-145.

1336 Delais-Roussarie, Elisabeth 1995. *Pour une approche probabiliste de la structure*

1337   *prosodique, étude de l'organisation prosodique et rythmique de la phrase*

1338   *française*. PhD dissertation, Université Toulouse-le-Mirail.

Delais-Roussarie, Elisabeth, Doetjes, Jenny, and Peter Sleeman 2004. 'Dislocation', in F. Corblin and H. de Swart (eds.), *Handbook of French Semantic*, Stanford, 501-528.

Delais-Roussarie, Elisabeth and Ingo Feldhausen 2014. 'Variation in Prosodic Boundary Strength: from a study of dislocated XP in French'. *Online Proceedings of Speech Prosody*. Dublin, Ireland.

Delais-Roussarie, Elisabeth, Post, Brechtje, Avanzi, Mathieu, Buthke, Caroline, Di Cristo, Albert, Feldhausen, Ingo, Jun, Sun-Ah, Martin, Philippe, Meisenburg, Trudel, Rialland, Annie, Sichel-Bazin, Rafèu and Hi-Yon Yoo, 2015. 'Developing a ToBI System for French'. In Sonia Frota and Pilar Prieto (eds), *Intonational Variation in Romance*, Oxford: Oxford University Press.

De Hoop, Helen and Henriette de Swart 2000. 'Topic and Focus'. In L. Cheng and R. Sybesma (eds), *The First Glot International State-of-the-Article Book,* Berlin/New York: Mouton de Gruyter.

D'Imperio, Mariapaola, Bertrand, Roxane, Di Cristo, Albert and Cristel Portes 2007. 'Investigating Phrasing Levels in French: Is there a Difference between Nuclear and Prenuclear Accents?' In *Selected Papers from the 36th Linguistic Symposium on Romance Languages*. Amsterdam/New York, 97-110.

Doetjes, Jenny, Delais-Roussarie, Elisabeth, and Peter Sleeman 2002. 'The Prosody of Left Detached Constituents in French'. *Proceedings of the Speech Prosody 2002 conference*, Aix-en-Provence.

Feldhausen, Ingo 2010. *Sentential form and prosodic structure of Catalan*. Amsterdam: Benjamins.

1363 Féry, Caroline 2011. 'German sentence accents and embedded prosodic

1364   phrases'. *Lingua* 121/13: 1906-1922.

1365 Fonseca-Greber, Bonnie and Linda R. Waugh 2003. 'The subject clitics of

1366   Conversational European French'. In Núñez-Cedeño, R., López, L. and

1367   Cameron, R. (eds.), *A Romance Perspective on Language Knowledge and*

1368   *Use, Selected papers from the 31st Linguistic Symposium on Romance*

1369   *Languages (LSRL)*, Chicago.

1370 Fradin, Bernard 1990. 'Approche des constructions à détachement : inventaire'.

1371   *Revue Romane*. 3-34.

1372 Frascarelli, Mara and Roland Hinterhölzl 2007. 'Types of Topics in German and

1373   Italian', in Winkler, S. and Schwabe, K. (eds), *On Information Structure,*

1374   *Meaning and Form*, Amsterdam: Benjamins. 87-116.

1375 Gili-Fivela, Barbara 1999. 'The prosody of left-dislocated topic constituents in

1376   italian read speech'. *Proceeding of: Sixth European Conference on Speech*

1377   *Communication and Technology, Eurospeech 1999*, September 5-9, 1999,

1378   Budapest, Hungary.

1379 Ginzburg, Jonathan 1996. 'Interrogatives: Questions, facts, and dialogue'. In

1380   Shalom, L. (ed), *Handbook of Contemporary Semantic Theory*. Oxford:

1381   Blackwell, 385-422.

1382 Givón, Talmi 1976. "Topic, pronoun and grammatical agreement." In Li, C.,

1383   editor, Subject and Topic. Academic Press: New York. 149–211.

1384 Givón, Talmi 1983. 'Topic continuity in discourse: the functional domain of switch

1385   reference'. In J. Haiman and P. Munro (eds), *Switch Reference, typological*

1386   *studies in language*. Vol. 2., Amsterdam: Benjamins.

1387   Goldman, Jean-Philippe. 2011. 'EasyAlign : an Automatic Phonetic Alignment

1388   Tool under Praat'. *Proceedings of Interspeech*, Firenze, Italy, 3233-3236.

1389   Grobet, A. & Simon, A. C. 2009. 'Constructions à détachement à gauche : les

1390   fonctions de la prosodie'. In Apothéloz, D., Combettes, B. & Neveu, F. (eds.),

1391   *Les linguistiques du détachement*. Berne, Peter Lang, 289-303.

1392   Guilbault, Christian 1995. *La dislocation à gauche en français québécois*

1393   *spontané : analyse instrumentale*. Master Thesis, Université de Laval.

1394   Gundel, Jeannette, T. Fretheim, 'Topic and focus', L.R. Horn, G. Ward (eds.),

1395   The Handbook of Pragmatics, Blackwell, Oxford (2004), pp. 175–196

1396   Haegeman, Liliane 2004. Topicalization, CLLD and the left periphery. In ZAS

1397   Papers in Linguistics 35: 157–192.

1398   Hart, Johan't, Collier, René and Antonie Cohen 1990. *A Perceptual Study of*

1399   *Intonation: an Experimental-Phonetic Approach*. Cambridge: Cambridge

1400   University Press.

1401   Herring Susan. 1990. Information structure as a consequence of word order type.

1402   *BLS,* 163-174.

1403   Hooper, Joan and Sandra Thompson 1973, 'On the applicability of root

1404   transformations', *Linguistic Inquiry* 4(4). 465–497.

1405   House, David 1990. *Tonal Perception in Speech*. Lund: Lund University Press.

1406   Jackendoff, Ray 1972. *Semantic interpretation in generative grammar*, MIT

1407   Press.

1408   Kallulli, Dalina and Liliane Tasmowski 2008. *Clitic Doubling in the Balkan*

1409   *Languages*, Amsterdam: Benjamins.

1410   Ladd, Robert 1996. *Intonational Phonology*, Cambridge: Cambridge University

1411   Press.

1412 Lacheret-Dujour, Anne and Beaugendre, Frédéric. La prosodie du français.

1413   Paris : Editions du CNRS.

1414 Lacheret-Dujour, Anne and Jacques François 2004. 'De la notion de

1415   détachement topical à celle de constituant thématique extrapropositionnel'.

1416   *Cahiers de praxématique*, 40, 167-198.

1417 Lacheret-Dujour, Anne, Obin, Nicolas and Mathieu Avanzi 2010. 'Design and

1418   evaluation of shared prosodic annotation for French spontaneous speech:

1419   from experts knowledge to non-experts' annotations'. *Proceedings of the 4th*

1420   *Linguistic Annotation Workshop*, 265-273.

1421 Lahousse, Karen 2010. 'Information Structure and epistemic modality in

1422   adverbial clauses'. *Studies in Language*.

1423 Lambrecht, Knut 1988. 'Presentational cleft constructions in spoken French'. In J.

1424   Haiman and S. Thompson (eds) *Clause combining in grammar and*

1425   *discourse*. Amsterdam: Benjamins.135-179.

1426 Lambrecht, Knut 1994. *Information structure and sentence form.* Cambridge:

1427   Cambridge University Press.

1428 Lambrecht, Knut 2001. 'A framework for the analysis of cleft constructions',

1429   *Linguistics* 01/2001; 39(3):463-516.

1430 Le Gac, David and Yoo, Hiyon 2011. 'Intonation of left dislocated topics in

1431   Modern Greek', *Proceedings of Interspeech*, Firenze, Italy, 1361-1364.

1432 Likhacheva, L. 2008. *L'articulation pragmatique du discours. Le cas des*

1433   *constructions détachées à gauche*. PhD Dissertation, Université de

1434   Toulouse.

1435 Marandin, Jean-Marie, Beyssade, Claire, Delais-Roussarie, Elisabeth, Rialland,

1436   Annie 2002. 'Discourse Marking in French: C Accents and Discourse Moves'.

1437   *Proceedings of Speech Prosody 2002*. Aix-en-Provence.

1438 Martin, Philippe 1987. 'Prosodic and Rhythmic Structures in French', *Linguistics*,

1439   25, 925-949.

1440 Mertens, Piet 2008. 'Syntaxe, prosodie et structure informationnelle: une

1441   approche prédictive pour l'analyse de l'intonation dans le discours'. *Travaux*

1442   *de linguistique*, 56, 97-124.

1443 Michelas, Amandine and Mariapaola D'Imperio 2010. 'Durational Cues and

1444   Prosodic Phrasing in French: Evidence for the Intermediate Phrase',

1445   *Proceedings of Speech Prosody*. Chicago, USA.

1446 Molnár, Valeria 2002. 'Contrast - from a contrastive perspective'. In Hallelgard,

1447   H., Johansson, S., Behrens, B., Fabricius-Hansen, C. (eds), *Information*

1448   *Structure in a Cross-Linguistic Perspective* , Amsterdam/New York: Rodopi,

1449   147–161.

1450 Myhill, John and Zhiqun Xing. 1996, 'Towards an operational definition of

1451   discourse contrast'. *Studies in Language*, 20: 303–60.

1452 Neeleman, Ad and Reiko Vermeulen (eds.) 2012. *The Syntax of Topic, Focus,*

1453   *and Contrast*. Berlin: Mouton De Gruyter.

1454 Nespor, Marina and Irene Vogel 1986. *Prosodic Phonology*. Dordrecht: Foris.

1455 Pekarek Doehler, Simona and Gabriele M. Mueller 2009. 'Le problème c'est de

1456   les distinguer: disloquée à gauche et pseudo-clivée dans la conversation',

1457   Proceedings of Apothéloz, D., Combettes, B., Neveu, F. (eds) *Les*

1458   *linguistiques du détachement. Actes du colloque international de Nancy (7-9*

1459   *juin 2006)*, Bern: Peter Lang.

1460 Prince, Ellen 1992. 'The ZPG letter: subjects, definiteness, and information-

1461   status'. In William C. Mann and Sandra A. Thompson (eds), *Discourse*

Description: Diverse Linguistic Analyses of a Fund-Raising Text, 295–326. Amsterdam: Benjamins.

R Development Core Team 2015. 'R: A language and environment for statistical computing', R Foundation for Statistical Computing, Vienna, Austria, available at http://www.R-project.org.

Reinhart, Tanya 1981. 'Pragmatics and Linguistics: An Analysis of Sentence Topics', *Philosophica, 27*, 53-94.

Repp, Sophie 2016. 'Contrast: Dissecting an elusive information-structural notion and its role in grammar'. In C. Féry & S. Ishihara (eds), *OUP Handbook of Information Structure*. Oxford: Oxford University Press.

Riou, Étienne and Barbara Hemforth 2015, 'Dislocation clitique de l'objet à gauche en français écrit', *Discours*, 16 | 2015, online since September 2015. URL : http://discours.revues.org/9037 ; DOI : 10.4000/discours.9037

Rizzi, Luigi 1997. The Fine Structure of the Left Periphery. In L. Haegeman (ed.), *Elements of Grammar*. Dordrecht: Kluwer Publications, 281-337.

Roberts, Craige 1996. 'Information Structure: Towards an integrated formal theory of pragmatics'. In J.H. Yoon and A. Kathol (eds), *OSUWPL* Vol. 49: *Papers in Semantics*, The Ohio State University, Dept. of Linguistics.

Rossi, Mario 1999. *L'intonation du français. Description et modélisation*. Paris/Gap: Ophrys.

Sawada, M. and R.K. Larson 2004. 'Presupposition and root transforms in adjunct clauses'. In M. Wolf and K. Moulton (eds.) *Proceedings of NELS 34*. (pp. 517-528) UMASS: GLSA.

Selkirk, Elisabeth 1984. *Phonology and Syntax: The relation between Sound and Structure.* Cambridge, MA: MIT Press.

Simons, Mandy 2007 "Observations on embedding verbs, evidentiality, and

    presupposition", *Lingua*, *117/6*, 1034-1056

Simons, Mandy, Judith Tonhauser, David Beaver, and Craige Roberts. 2010.

    'What projects and why'. In *Semantics and Linguistic Theory (SALT) 21*,

    309–327. Ithaca, NY: CLC Publications.

Sleeman, Petra 2012. 'Quantifier preposing in French and Italian as a root

    phenomenon: a syntactic or a pragmatic approach?*' January 2012 Bucharest*

    working papers in linguistics; vol. 14 issue 1, p5.

Surányi, Balázs, Ishihara, Shinichiro and Fabian Schubö 2012. 'Syntax-prosody

    mapping, topic-comment structure and stress-focus correspondence in

    Hungarian'. In Elordieta, G. and P. Prieto (eds) *Prosody and Meaning*. Berlin:

    Walter de Gruyter, 35-71.

Vallduví, Enric 1992. *The Informational Component*. New York: Garland.

Vallduví, Enric and Maria Vilkuna 1998, 'On rheme and kontrast. In Cullicover

    and McNally (eds.), *Syntax and semantics 29: The limits of syntax ,* San

    Diego: Academic Press, 79-108.

Villalba, Xavier, 2000. *The syntax of sentence periphery*. Ph.D. Thesis,

    Universitat Autònoma de Barcelona.

Zribi-Hertz, Anne 1994. The Syntax of nominative clitics in Standard and

    Advanced French. In Cinque, G., Koster, J. Pollock, J.-Y., Rizzi, L. and

    Zanuttini, R. (eds). Paths Towards Universal Grammar. Washington, D.C.:

    Georgetown University Press, 453-472.

Zubizarreta, María Luísa 1998. *Topic, Focus, and Word Order. Linguistic Inquiry*

    Monographs. Cambridge MA: MIT Press.