Pragmatic orientation of syntax in spontaneous speech

A corpus-based comparison between Brazilian Portuguese and Italian adverbial clauses

Giulia Bossaglia

Universidade Federal de Minas Gerais

This paper reports on the syntax of finite adverbial clauses in spoken Brazilian Portuguese (BP) and Italian (IT), based on C-ORAL-BRASIL and Italian C-ORAL-ROM *corpora*. The analysis of adverbial clauses here is mainly focused on the syntax/pragmatics and syntax/information structure interfaces, aiming to show significant differences in the syntax of spoken *versus* written language.

Keywords: Language into Act Theory; syntax; spontaneous speech; Brazilian Portuguese; Italian

1.   A note on adverbial clauses

Adverbial clauses are traditionally defined within the domain of subordination as dependent clauses with an adverb-like function with respect to the predicate of the main clause, in opposition to complements, functioning as noun phrases saturating the valence of the main predicate, and relatives, modifying nouns in the main clause (Longacre & Thompson 1985, Kortmann 1997, Thompson, Longacre & Hwang 2007).

Adverbial clauses are typically introduced by specific morphemes (adverbial subordinators or adverbial conjunctions) carrying the lexical information specifying the kinds of semantic relation existing between the two clauses (Foley & Van Valin 1984, Givón 1994, Hengeveld & Wanders 2007)[1]. Such semantic re-

---

[1] Kortmann (1997: 56-ff.) points at some circularity problems related to definitions of adverbial clauses focusing on the presence of an explicit adverbial subordinator, since adverbial

lations are generally identified in temporal, conditional, cause/reason, concessive, means, manner (Matthiessen & Thompson 1988, cf. Longacre & Thompson 1985; see also Hengeveld 1993, 1998, Hengeveld & Wanders 2007 for a slightly different classification).

In other terms, two states of affairs can be linked in a way that one (the adverbial clause) represents the circumstances in which the other (the main clause) occurs (Cristofaro 2003: 155). The codification of such circumstances is described as optional, and adverbial clauses can be not only preposed or postponed to the main clause, but also omitted without affecting its grammaticality (Hengeveld 1998). This implies that the event or state of affairs encoded by an adverbial clause is conceived as semantically autonomous from that of its main clause, in a way that is significant in order to better understand the peculiar position of these clauses within the domain of subordination itself.

It is nowadays assumed that coordination and subordination are not parts of a dichotomy, but rather the endpoints of a *continuum* (Haiman & Thompson 1984, Lehmann 1988, Givón 1980, 1990; see also Foley & Van Valin 1984). Starting from this assumption, Lehmann (1988) points out that clause combining in subordination can be analyzed as well by means of a multi-dimensional set of related *continua* involving parameters such as: (1) downgrading of subordinate clause (parataxis vs. embedding; cf. the distinction between the features *dependent* and *embedded* in Foley & Van Valin 1984), (2) desententialization of dependent clause (finiteness vs. *nouniness*; cf. *balancing* vs. *deranking*: Stassen 1985, Cristofaro 2003), (3) interlacing of the clauses and (4) explicitness of linking between them (syndesis vs. asyndesis).

All these *continua* describe different morphosyntactic levels in which the degree of semantic autonomy or integration of the subordinate clause can be iconically represented, since "the closer the semantic relationship, the tighter the syntactic linkage" between two clauses (Foley & Van Valin 1984: 264; see Haiman 1983, Givón 1980, 1990).

Adverbial clauses typically codify autonomous events or states of affair with respect to their main clause, and they actually represent the less subordinate type of dependent clauses.[2] Actually, (finite) adverbial clauses are not embedded,[3]

---

relations between clauses are not directly linked to it. Nonetheless, this paper will not deepen this aspect since only a working definition of adverbial clause is needed.

[2] Actually, not all semantic types of adverbial clauses share the same degree of autonomy with respect to the main clause, as, for example, purpose clauses (see Cristofaro 2003: Chapter 6, Schulte 2007, Scmidtke-Bode 2009 for details). Compare, in a generative perspective, the distinction between *central* and *peripheral* adverbial clauses in Haegeman (2004, 2012).

their predicate is more predictably "balanced", they are rarely interlaced with the main clause, and the adverbial morphemes introducing them are explicit not only in terms of the syndetic relationship with the main clause, but also in specifying the semantics of such a relationship.

1.1   Adverbial clauses in spoken language

The behavior of finite adverbial clauses has received much attention in many studies on speech, which analyze a set of properties and functions in different languages and with slightly varied focuses. Studies that include intonation in their analyses are particularly interesting, since in our theoretical and methodological perspective (see section 2) prosody represents a crucial, then not negligible, dimension of spoken language.

The following properties of adverbial clauses in speech are important for the purposes of this paper: the mobility of the position of adverbial clauses with respect to the main one and the different functions those positions appear to carry out in interaction, and the type of prosodic linking between adverbial and main clauses.

The taxonomy of adverbial clauses proposed in Chafe (1984, 1988) constitutes the basic ground of many other works on this topic (Longacre & Thompson 1985, Mathiessen & Thompson 1988, Ford 1993, Couper-Kuhlen 1996, Ford & Thompson 1996, Thompson, Longacre & Hwang 2007, Hopper & Thompson 2008, among others).

Chafe's taxonomy is based on (i) position with respect to the main clause, (ii) ± boundedness in the same pitch contour (i.e. intonation unit) of the main clause, (iii) type of intonational profile ("comma" and "period" intonation corresponding respectively to non-final and final pitch movement) of the two clauses. Adverbial clauses, therefore, are classified as preposed and bound (see (a) below), postponed and bound (b), preposed and free (c), postponed and free (d; examples taken from Chafe 1984: 439):

(a)   *Because it has such a big memory* I decided to buy it**.**
          (one intonation unit segmented by "period" intonation)

---

[3] This is the defining feature of subordination according to Mathiessen & Thompson (1988), but in the slightly different view of Foley & Van Valin (1984: 264) a clause [+ dependent] and [-embedded] would be considered within the domain of *cosubordination*, since their notion of subordination requires embedding.

(b)   I decided to buy it *because it has a big memory*.
        (one intonation unit segmented by "period" intonation)

(c)   *Because it has such a big memory*,
        I decided to buy it.
        (two intonation units separated by *comma* intonation; the adverbial
        clause is performed with non-final contour)

(d)   I decided to buy it,
        because it has such a big memory.
        (two intonation units separated by *comma* intonation; the main clause is
        performed with non-final contour)

In Chafe's view, free preposed adverbial clauses (like c) cannot display period intonation, while postponed ones can display both types. This follows from the assumption that, in general, preposed adverbial clauses function as "guideposts" to the information flow, providing background to the information in the following main clause, while postponed ones fulfill a more "coordinate" function, commenting a specific condition relevant to the preceding main clause. When free (i.e. in a separate intonation unit, with final intonation) postponed adverbial clauses are used in interaction as afterthoughts.[4]

In a similar perspective, Couper-Kuhlen (1996) analyses *because*-clauses in spoken English, finding in the presence or absence of prosodic reset of f_0 (fundamental frequency) declination between main and postponed dependent clauses the defining cue of two possible functions of *because*-clauses: with the prosodic reset, the dependent clause represents an intonation unit of its own, serving as a turn construction means ("intonational coordination": 402); without the declination reset, the *because*-clause belongs to the prior clause ("intonational subordination": *ibidem*). All these studies challenge the traditional notion of syntactic dependency, making it clear that, by virtue of specific *pragmatic* functions of these (formally) dependent clauses in the construction of speech and interaction, adverbial clauses are not interpretable as instances of proper subordination in various of the contexts they occur.[5]

---

[4] Ford, Fox & Thompson (2002) argue that such a use does not correspond to afterthought, but rather to an increment strategy in the construction of the discourse in interaction, typically involving adverbial clauses (see also Couper-Kuhlen 1996, Thompson & Couper-Kuhlen 2005).
[5] Hopper & Thompson (2008) present some interesting data on conditional *wenn*-clauses (equivalent to *if*-clauses) in spoken German. Here, such a challenge to the notion of syntactic dependency has as well a concrete syntactic manifestation in the fact that the (standard) ex-

Our analysis of adverbial clauses in spoken Brazilian Portuguese and Italian considers a similar set of parameters, as well.

2.   Theoretical framework: *Language into Act Theory*

This study on adverbial clauses in spoken BP and IT is founded on the *Language into Act Theory* (from now on, L-AcT; Cresti 1995, 2000), an inductive theory constructed throughout decades of empirical work on *LABLITA* (*Linguistics Laboratory of the Italian Department*, Florence University) spoken corpora (Moneglia 2005). Based on L-AcT, the C-ORAL-ROM (Cresti & Moneglia 2005) and C-ORAL-BRASIL (Raso & Mello 2012a) projects designed the C-ORAL-ROM (Italian, French, Spanish, European Portuguese) and C-ORAL-BRASIL (Brazilian Portuguese) comparable spoken corpora.

L-AcT represents an innovative approach to the study of spoken language, since it emphasizes the importance of the prosodic dimension of speech. Actually, many speech studies have been based solely on speech transcriptions for a long time, which means they have taken *written* texts as data source, thus failing to account for an adequate analysis of the richness and peculiarities of spoken language, starting from syntax. Moreover, L-AcT is innovative in the way the importance of prosody in speech is conceived even with respect to approaches that do recognize and include intonation in their analysis (cf. section 1), as we will show throughout this section.

L-Act is an extension of Austin's Speech Act Theory (Austin 1962) and assumes that a Speech Act is composed by the following simultaneous acts:

a.   locutionary act: the transmitted linguistic content;
b.   illocutionary act: the (type of) action performed by means of the transmitted linguistic content;
c.   perlocutionary act: the affective stimulus which makes the speaker perform the speech act.[6]

---

pected verb inversion in the apodosis in some cases does not occur, manifesting the lack of syntactic integration with the preceding protasis. These cases are precisely considered examples of apparent biclausal structures, understood as grammatical formats accomplishing discourse oriented functions.

[6] Austin's perlocutionary act corresponds, instead, to the speaker's purpose in performing the speech act.

Based on prosodic criteria, L-AcT identifies the unit of reference of spoken language in the utterance, understood as the minimal speech segment prosodically and *pragmatically* interpretable as autonomous, that is, as a speech act. Prosody can provide pragmatic autonomy to *any* linguistic content, even if it lacks syntactic or semantic completeness, as it is shown in the examples below (from Italian C-ORAL-ROM and C-ORAL-BRASIL corpora):[7]

(1)   *CAR: muito feio // Ex. bfamcv03 [252]
        '*CAR: very bad //'

(2)   *REN: ahn // Ex. bfamdl01 [522]
        '*REN: uhm //'

(3)   *MAX: a Recco // Ex. ifamcv01_taglio [58]
        '*MAX: to Recco //'

(4)   *ELA: a Recco // Ex. ifamcv01_taglio [59]
        '*ELA: to Recco //'

(5)   *VAL: hanno pubblicato la graduatoria / Ex. ifamcv18 [185]
        '*VAL: they published the ranking list /'

In (1)-(4) an adjectival phrase, an interjection and two prepositional phrases, respectively, display prosodic and pragmatic autonomy, i.e., they are utterances. Additionally, (3) and (4) show how different prosodic profiles convey different speech acts (illocutions), independently from the locutive content, which is the same in both (*a Recco*).

On the other hand, listening to (5), a syntactic and semantically complete sentence, it is possible to perceive that such a tonal unit is not pragmatically autonomous, as if something were missing. In the terms of Chafe (1988), (5) displays a "comma intonation", signaling that there is "more to come". Actually, (5) represents only a sub-part of a bigger utterance, reported in (6) below:

---

[7] Starred abbreviations stand for the speakers' pseudonyms in the *corpora*; the name of the source texts of the examples are indicated in the audio link: *b* = Brazilian, *i* = Italian, *fam* = familiar/private context, *pub* = public context, *mn* = monologue, *dl* = dialogue, *cv* = conversation (then: *bfamcv03* stands for familiar/private conversation n. 03 of Brazilian Portuguese C-ORAL *corpus*, and so on).

(6)   *VAL: hanno pubblicato la graduatoria / il nove di maggio //
        Ex. ifamcv18 [185]
        '*VAL: they published the ranking list / on May the ninth //'

(6a)   / il nove di maggio // Ex. ifamcv18 [185]
        '/ on May the ninth //'

The utterance in (6) is composed by two units, of which only the second is pragmatically autonomous (listen to (6a)), for it carries the prosodic nucleus of the illocution. The unit carrying the illocution is the Comment (COM), and it is the only one which is necessary, in speech, for an utterance to be pragmatically autonomous. *Simple* utterances are formed only by this single, necessary unit, like (1)-(4), while *compound* utterances are formed by the COM, plus one or more units.

The prosodic/pragmatic criteria adopted by L-AcT manage to identify a consistent unit of reference for speech in the utterance, since the adoption of a syntactic one fails to account for a large amount of spoken language data, which hardly correspond to a clause or a sentence, but still do function properly from a communicative point of view. [8]

The segmentation of speech flow into tonal units is based on prosodic criteria as well: prosodic breaks perceived as non-terminal are marked with "/", while terminal breaks, determining utterance boundaries, with "//" in the transcription.[9] Thus, a Prosodic Interface assumed by L-AcT (Cresti 2000, 2014; Moneglia 2011) exists in order to (i) provide pragmatic autonomy to linguistic contents, (ii) define the type of action performed by means of the utterance, and (iii) segment the speech flow in utterances and utterance sub-units.

Tonal units other than COM can carry prosodically conveyed information values as well: following the IPO model (*Institute for Perception Research Eindhoven*; see t'Hart, Collier & Cohen 1990), L-AcT assumes that f_0 exhibits several movements within the utterance. Some of them are voluntary,[10] though unconscious, and have perceptual relevance, conveying determined information values, in accordance with the *Information Patterning Theory* (Cresti 1995;

---

[8] See Miller & Weinert (1998) for a discussion of the relevant literature on the problem of defining a consistent unit of reference for speech.

[9] Details on transcription criteria and norms in Mello *et al.* (2012).

[10] Involuntary movements, instead, are due to micro-melodical phenomena and have no perceptual relevance. In this perspective, the f_0 declination reset pointed out by Couper-Kuhlen (1996) in performing a specific type of *because*-clause is a voluntary movement carrying relevant information.

Moneglia & Cresti 2006; Cresti & Moneglia 2010).[11] In (6) above, for example, the first tonal unit conveys a Topic (TOP) information unit, with the value of defining the domain of application of the illocution in COM, by virtue of a relation of pragmatic *aboutness* existing between the two. In this way, based on prosodic and pragmatic criteria, L-AcT manages to widen the range of information units relevant to the analysis of speech, traditionally limited to Topic and Comment (Chafe 1976; Halliday 1989; Krifka 2007). Actually, several information units were identified, characterized by specific (i) functions, and (ii) prosodic profiles implying their (iii) distribution within the utterance. Following the IPO model (t'Hart, Collier & Cohen 1990, Cresti & Moneglia 2010), the possible prosodic profiles are: *root*, which defines the type of illocution and identifies the Comment unit only; *prefix*, which precedes a *root* unit and is typical of Topic units; *suffix*, which follows a *root* unit and is proper of Appendix units, for example (see Table 1 below for the explanation of information units functions); and *postfix*, which can precede or follow a *root* unit, but cannot occur at the beginning of the utterance (it is the Parenthesis unit prosodic profile).[12]

Information units are divided into *textual* units, which form the text of the utterance or are directed to the interpretation of it (see Table 1), and *dialogic* units, which regulate the interaction (Table 2; tables adapted from Moneglia & Raso 2014: 490).

Units without information value are found in the speech flow as well, as illustrated in Table 3 (adapted from Moneglia & Raso 2014: 491).

Such an overview on the information units identified by L-AcT is fundamental in order to understand how this theory conceives the syntax of spoken language.

2.1 L-AcT: properties of spoken language syntax

As introduced in the previous section, L-AcT identifies the unit of reference of speech in the utterance, determined through a pragmatic/prosodic criterion, therefore overtly excluding a syntactic one. Actually, in the still ongoing debate on which would be the proper unit of reference for speech, the inadequacy of the sole syntactic criterion has emerged, since a great amount of linguistic content in

---

[11] Besides f_0 movements, other variables such duration, intensity and syllabic alignment to the f_0 curve are taken into account.

[12] *Root* and *prefix* are special profiles for they host a specific prosodic nucleus, while the other types do not. This is consistent with the relevance of Comment and Topic units, the pattern Topic-Comment representing spoken language's main one (Moneglia & Cresti 2010, Mittmann 2012).

Table 1. Information Units according to L-AcT: Textual Units

Type	Tag	Function
Comment	COM	It accomplishes the illocutionary force of the utterance.
Topic	TOP	It establishes the domain of application of the illocution expressed by the Comment.
Appendix of Comment	APC	It integrates the text of the Comment and concludes the utterance.
Appendix of Topic	APT	It gives a delayed integration of the information given in Topic.
Parenthesis	PAR	It provides instructions about how the utterance, or a part of it, has to be interpreted; it has a *backward* or *forward* scope.
Locutive Introducer	INT	It signals that the subsequent locutive space has to be interpreted by means of different coordinates (generally, it introduces metaillocutions).
Multiple Comments	CMM	They constitute a chain of Comments forming an illocutionary prosodic pattern, which is interpreted holistically.
Bound Comment	COB	It constitutes a chain of Comments, produced by progressive adjunctions which fol-low the flow of thought. Chains of COB are called Stanzas.
List of Topic	TPL	List of two or more TOP units semantically and syntactically connected, forming one prosodically marked major unit.
List of Parenthesis	PRL	List of two or more PAR units semantically and syntactically connected, forming one prosodically marked major unit.

Table 2. Information Units according to L-AcT: Dialogic Units

Type	Tag	Function
Incipit	INP	It opes the communicative channel in order to start a dialogic turn or an utterance, bearing contrastive value.
Conative	CNT	It pushes the interlocutor to participate or to stop a non-collaborative behaviour in a dialogue.
Phatic	PHA	It controls the communicative channel, contributing to its maintenance and stimulating the interlocutor towards social cohesion.
Allocutive	ALL	It specifies the addressee of the message, keeping his attention and having a social cohesive function.
Expressive	EXP	It provides emotional support to the speech act, marking social cohesion.
Discourse Connector	DCT	It connects different parts of the discourse, marking continuity.

Table 3. Units without information value

Type	Tag	Function
Scanning Unit	SCA	It is a tonal part of a bigger information unit, whose locutive content's needs to be scanned, normally for expressive or speech incompetence reasons.
Interrupted Unit	i-[TAG]	It is part of an information unit which is interrupted by a PAR or dialogic unit.
Empty Unit	EMP	Unit whose locutive content is not to be considered as part of the utterance, as it happens in case of 1) retracting; 2) interrupted last unit of an utterance.
Time taking Unit	TMT	It corresponds to the so-called filled pause.
Unclassified	UNC	Unit to which it is not possible to attribute another tag for some reason.

speech does not correspond to complete syntactic units, but rather to phrases, interjections, and fragments. Therefore, the sentence has been discarded as the unit of reference for speech, and most authors identify it in the clause (Chafe 1984, 1988, Quirk et al. 1985, Halliday 1985, Voghera 1992, Miller & Weinert 1998, Thompson & Couper-Kuhlen 2005). Additional problems follow from the fact that syntactic boundaries frequently do not correspond to prosodic ones in the speech flow, making it necessary for many authors to postulate the existence of an underlying "crystalline deep structure with clear clause boundaries" (Heath 1985: 103), attempting to reduce to some abstract syntactically well-formed structure the strongly fragmented structures of spoken language.

Assuming that intonation represents a paramount part of speech grammar, L-AcT proposes that the analysis of the syntactic properties of spoken language take into account how the utterance package is constructed, that is, how the communicative content is pragmatically/informationally built by the speaker. Thus, the syntax of spoken language is analyzed according to whether the utterance is a simple or a compound one. Therefore, according to L-AcT, each information unit represents a syntactic and semantic island, and only in this specific domain is it possible to analyze proper syntactic relations:

(7)   *PAU: esse tipo de muro /=TOP= se ficar baixo demais ele fica feio
        //=COM= Ex. bpubdl01 [71]
        '*PAU: this kind of wall /=TOP= if (it) is (built) too short it is ugly
        //=COM='

(8)   *LAU: departamento /=TOP= Artes Plásticas //=COM=
         Ex. bfamdl03 [148]
         '*LAU: department /=TOP= Plastic Arts //=COM='

In (7), the noun phrase in TOP cannot be analyzed as syntactic subject of the (although co-referent) predicate in COM, because this role is already fulfilled by the pronoun *ele*. The relation between the two units, in fact, is a pragmatic rather than a syntactic one: TOP unit defines the domain of application of the illocution of COM unit, by virtue of a relationship of pragmatic *aboutness* existing between the two. In (8) this is even clearer, since the two information units, corresponding to two different nominal phrases, are informationally linked together by prosody, without the presence of a predication being necessary for the utterance to be well-formed. Therefore, there is no syntactic compositionality between different information units, but rather a pragmatically oriented combination of the local clauses, phrases, or fragments constituting their locutive content.[13]

Based on these assumptions, L-AcT distinguishes between: *linearized syntax*, proper subordination and coordination structures performed within the same information unit; and *patterned constructions*, subordination and coordination structures performed across more than one information unit or utterance, therefore not analyzable in terms of syntactic compositionality (Cresti 2014: 374):[14]

(9)   *GCM: se tu lo ritieni necessario tu contestualizzi l'autore //=COM=
         Ex. ipubdl05 [134]
         '*GCM: if you think it's necessary you contextualize the author
         //=COM='

(10)   *SAR: se c'hai il costo /=TOP= lo vedi subito costo e prezzo //=COM=
         Ex. ifammn17 [123]
         '*SAR: if you have the cost /=TOP= you immediately see cost and price
         //=COM='

Examples (9) and (10) show two differently performed conditional sentences: in the former, protasis and apodosis stay together within the same island (cf. the *bound* adverbial clauses in Chafe 1984, 1988), which means that they are syn-

---

[13] See Cresti (2014, section 4.2) for a detailed discussion on syntactic compositionality based on Italian C-ORAL-ROM corpus.
[14] Cf. *integrated* vs. *unintegrated/fragmented syntax*: Miller & Weinert (1998) and *micro-* vs. *macro-syntaxe*: Blanche-Benveniste (2002).

tactically and semantically compositional. In the latter, protasis and apodosis are performed separately in two different islands, a TOP and a COM unit respectively (within the same utterance, though: in this sense the protasis in the patterned construction is not comparable with Chafe's *free* adverbial clause). In this case, the conditional sentence is performed according to a specific pragmatic organization, the protasis serving as pragmatic background for the apodosis in COM. Even if in conditionals it is still possible to observe certain degree of congruence between their semantic value and the pragmatic functions they appear to have in speech (together with other types of adverbial clauses, see *infra*), in cases such as (11) below it becomes clearer that the prevailing hierarchy is a pragmatic and not a syntactic one:

(11)   *SAB: e insomma /=PHA= ti dicevo che io e questa mia amica /=TOP=
        si voleva farci il tatuaggio //=COM= Ex. ifamdl09 [138]
        '*SAB: and so /=PHA= I was telling you that me and this friend of mine
        /=TOP= we wanted to get the tattoo //=COM='

The complement clause in the utterance above is performed via a patterned construction, where the main clause, which traditionally would be considered the head of regency, is performed in the TOP unit, while the dependent in COM. The informational relationship existing between the two units is, therefore, opposite to the syntactic one: from the point of view of pragmatics, which organizes the packaging of information in speech, in (11) the dependent clause is the most relevant, for it carries the illocution, while the main clause serves only to provide a background for it. In this sense, the relationship between main and complement clause is shaped by the speaker in a way that the complement "overrides" the main clause, which is rather used in order to posit a stance (Thompson 2002: 134). For such reasons, then, our analysis of adverbial clauses in spoken BP and IT constitutes as well a challenge to the traditional view of subordination and syntactic dependency in general.

3.   The syntax of adverbial clauses in spoken BP and IT

3.1   Methodology

The data analyzed in this study were retrieved from two informationally tagged *minicorpora* of C-ORAL-BRASIL (Raso & Mello 2012a) and Italian C-ORAL-ROM (Cresti & Moneglia 2005).

3.1.1 *The corpora*

Built following the principles of L-AcT, the Italian C-ORAL-ROM (Cresti & Moneglia 2005) and the C-ORAL-BRASIL (Raso & Mello 2012a, 2012b) *corpora* represent adequate resources for the analysis of spoken language, since they comprise:

- audio recordings;
- transcriptions including the annotation of prosodic breaks;
- speech-to-text alignment (through the WinPitch software, Martin 2004);
- morphosyntactic annotation (via the parser PALAVRAS, Bick 2000);
- metadata for each recording session, providing sociolinguistic information about the participants and the interactional context.

These corpora were designed in order to achieve the highest degree of diaphasic variation, and therefore to succeed in being representative of the great variety of spoken communicative situations (Raso & Mello 2012a, Mello 2014). Three interactional types are represented, determined on the basis of the number of participants actively involved in the interaction: dialogue (two participants), conversation (more than two participants), and monologue (one active participant). Interaction contexts are divided into familiar/private and public.

   C-ORAL-ROM *corpora* present both a Formal and an Informal speech sections, while the C-ORAL-BRASIL project has published the Informal one so far.

3.1.2 *The Italian and Brazilian Portuguese C-ORAL minicorpora*

Two comparable *minicorpora* of Italian C-ORAL-ROM and C-ORAL-BRASIL are available online at the *IPIC DataBase for Information Patterning Interlinguistic Comparison* (http://lablita.dit.unifi.it/app/dbipic/; Panunzi & Gregori 2012, Panunzi & Mittmann 2014).

   The two *minicorpora* were designed by choosing a collection of texts from Informal sections of their reference *corpora*, maintaining the same architecture in terms of interactional typologies proportions. Thus, the two *minicorpora* are actually representative of Informal Italian C-ORAL-ROM and C-ORAL-BRASIL. Besides all the listed characteristics of the *corpora* (cf. 3.1.1), IPIC *minicorpora* received informational unit manual annotation (see the tags in Tables 1 and 2 above). General data on the architecture and size of both *minicorpora* are summarized in Table 4 below (adapted from Panunzi & Mittmann 2014: 140):

Table 4. IPIC Brazilian Portuguese and Italian minicorpora

Corpus section		Brazilian			Italian		
Communicative context	Interaction types	sessions	words	TSs	sessions	words	TSs
Family/Private	mn	6	8,635	856	6	8,750	1,086
	dl	5	8,360	1,877	5	9,306	1,771
	cv	4	6,421	1,407	3	5,152	1,283
Public	mn	1	1,616	143	2	2,927	265
	dl	2	3,011	584	2	3,129	555
	cv	2	3,422	645	2	8,136	703
Total		**20**	**31,465**	**5,512**	**20**	**37,355**	**5,663**

The comparability of the two *minicorpora* is clear, in both architecture and size. Actually, despite the difference in number of words (31,465 vs. 37,335), the number of reference units (terminated sequences, i.e. Utterances and Stanzas) is nearly the same (5,512 vs. 5,663).

Since the online IPIC database provides a multi-filter query interface, finite adverbial clauses were retrieved by searching the occurrences of adverbial sub-ordinators or phrasal subordinating expressions. In Table 5, they are listed with their translations and correspondent adverbial values.

3.2   Data analysis

In the Brazilian and Italian *minicorpora* were gathered 321 (5.8% on the total number of terminated sequences) and 457 (8%) utterances containing adverbial clauses, respectively. This difference is not that surprising, since based on the IPIC *corpora*, Panunzi & Mittmann (2014) have already observed that Italian terminated sequences display a greater textual complexity, while BP shows a preference for simple utterances (single COM units) and, in compound utterances, a more frequent use of dialogic, instead of textual, units. Since it is in textual units that the most part of locutive content appears (dialogic units are used as interactional regulators, see Table 2 above), it is more likely that a greater number of complex syntactic structures be used in IT than in BP.

Table 5. Adverbial subordinators in BP and IT minicorpora

Brazilian Portuguese		Italian		Adverbial value
porque	'because'	*perché*	'because'	**Cause/**
como	'since'	*dato che*	'since' (lit. 'given that')	**Reason**
já que	'since'	*visto che*	'since' (lit. 'having seen that')	
		dal momento che	'since' (lit. 'from the moment that')	
		siccome	'since'	
se	'if'	*se*	'if'	**Conditional**
caso	'if' (lit. 'in the case that')			
quando	'when'	*quando*	'when'	**Time**
enquanto que	'while'	*finché*	'until'	
depois que	'after'	*mentre*	'while'	
se bem que	'although'	*anche se*	'even if'	**Concessive**
apesar que	'although'			

3.2.1 *Adverbial values*

As Table 5 above is showing, there is an interesting correspondence between the found values of finite adverbial clauses in both languages: Cause, Condition, Time, Concessive and Manner. In Table 6 below, it is possible to observe that BP and IT display similar frequencies for the different types of adverbial clauses:

Table 6. Adverbial values in BP and IT minicorpora

Adverbial Value	Brazilian Portuguese	Italian
Cause	44%	49%
Condition	29%	28.5%
Time	23%	18.5%
Concessive	2%	2%
Manner	2%	2%

Among the adverbial values verified in the two *minicorpora*, Cause, Condition and Time appear to be the most relevant in both languages, while very few occurrences of Concessive and Manner clauses were retrieved. For this reason our analysis will be mainly focused on the first three types of adverbial clauses.

Another fact of particular interest is that the great majority of adverbial clauses found in the two *minicorpora* occur in patterned constructions. Actually, on the total of terminated sequences hosting adverbial clauses, linearized ones (i.e. proper adverbial clauses, performed within the same information unit of the main clause, see (12)-(15) below) represent only 6% of BP data, and 4% of IT.

(12)   *LUZ: porque /=DCT= eu só soube que eu nũ [/6]=EMP= eu tive certeza
        absoluta que eu nũ era daqui quando eu saí //=COM=
        Ex. bfamdl03 [1]
        '*LUZ: because /=DCT= I only knew that I don't [/6]=EMP= I felt
        absolutely sure that I didn't pertain here when I left //=COM='

(13)   *CAR: não falo porque acho muito pesado //=COM=
        Ex. bfammn05 [58]
        '*CAR: I don't say it because I think it's very heavy //=COM='

(14)   *CLA: io quanno me [/3]=EMP= me fanno ride quanno dice questa
        fratellanza //=COM= Ex. ifammn02_taglio [195]
        '*CLA: I when me [/3]=EMP= they make me laugh when they say "this
        brotherhood" //=COM='

(15)   *OTT: perché io &esa [/3]=EMP= io costruisco la tre F /=SCA=
        esattamente perché /=SCA= sono militante marxista rivoluzionario
        //=COM= Ex. ipubcv01_taglio [95]
        '*OTT: because I &exa [/3]=EMP= I build the "Tre F" /=SCA= exactly
        because /=SCA= I'm a revolutionary marxist militant //=COM='

The opposite trend has been observed with respect to finite complement clauses: based on the same *minicorpora* of Italian C-ORAL-ROM and C-ORAL-BRASIL, Bossaglia (2014) found that in both languages the great majority of complements occur in linearized configuration (83.9% in BP, 67.3% in IT; see also Cresti 2014: 407 on Italian).

Therefore, it seems that a certain degree of iconicity exists between the semantic and syntactic integration of adverbial and complement clauses and the way the speakers perform them in speech. Actually, adverbial clauses and com-

plements represent opposite degrees of integration with their main clause, as it has been extensively shown in previous sections.

3.3   Time adverbial clauses

The main subordinator for Time adverbial clauses is *quando* "when" in both BP and IT. This type of adverbial clause is almost exclusively used in the TOP/COM pattern (90% on the total of patterned constructions in BP, 80% in IT), the dependent clause being hosted in the Topic unit, therefore always preposed to the main one in COM, as it is shown in the examples below:

(16)   *BEL: quando eu cheguei aqui /=TOP= todas as minhas calças tinham ficado lá hhh //=COM= Ex. bfamdl02 [243]
'*BEL: when I arrived here /=TOP= all my trousers had remained there hhh //=COM='

(17)   *GIL: <principalmente depois que eles tão> fazendo um campeonato e não nos chamaram /=TOP= eles <tão fora> //=COM=
Ex. bfamcv01 [243]
'*GIL: mainly after they have been organizing a tournament and they haven't convoked us /=TOP= they're out //=COM='

(18)   *ALE: cioè /=PHA= ha visto /=PHA= quando ti sposi /=TOP= un po' devi cambiare /=SCA= anche /=SCA= abitudini /=COM= < sai >
//=PHA= Ex. ifamcv15 [285]
'*ALE: I mean /=PHA= you see /=PHA= when you get married /=TOP= you have to change a little bit /=SCA= also /=SCA= your habits /=COM= you know //=PHA='

(19)   *NIC: dal momento che /=INT= non avendo i rotoli della carta /=PAR= finché la carta è dentro /=TOP= puoi far tutto quello che vuoi //=COM=
Ex. ifamdl17 [194]
'*NIC: since /=INT= having any paper roll /=PAR= until there's paper inside /=TOP= you can do anything you want //=COM='

Since the TOP unit has the function of defining/delimiting the domain of application of the illocution in COM, it is likely that adverbial clauses defining a specific temporal circumstance be extensively used in TOP. In this sense, the semantics of such adverbial clauses is consistent with TOP's pragmatic function.

A few Time clauses were found in the IT *minicorpus* in illocutionary patterns as well, in both CMM (Multiple Comments) and COB (Bound Comments).

Here, the position of the adverbial clause appears to be freer, as the Time clauses appear postponed as well:

(20)   *EDO: < io mordo > /=CMM= quando c' è la crema /=CMM= la crema la mordo /=CMM= e cade subito //=CMM= Ex. ifamcv15 [178]
'*EDO: I bite /=CMM= when there's the cream /=CMM= I bite the cream /=CMM= and it immediately drops //=CMM='

(21)   *CLA: ricatti /=COB= pressioni /=COB= quello che è successo alla FIAT /=COB= quando buttarono fuori migliaia di compagni /=COB= eccetera eccetera //=COM= Ex. ifammn02 [170]
'*CLA: blackmails /=COB= pressures /=COB= what happened at FIAT /=COB= when they kicked out thousands of comrades /=COB= etcetera etcetera //=COM='

Listening to (20) it is clear that the Time clause *quando c'è la crema* is postponed to the first CMM and not preposed to *la crema la mordo*. In these information patterns, adverbial and main clause are both hosted by units that are equal with respect to the pragmatic organization of the speech flow, since they all are illocutionary. The higher degree of freedom in the distribution of the adverbial clause may be related to this fact. Additionally, listening to (20) and (21) it is clear that in such patterns the adverbial clauses are adding information to the utterance, developing the speakers' accounts, and not providing any sort of background information, as they would do in a TOP/COM pattern.

3.4   Conditional adverbial clauses

The TOP/COM pattern is the most frequent in hosting conditional sentences as well (87% in both languages), the Conditional clause (mostly introduced by *se* "if" in BP and IT) always appearing in TOP. Some textually more complex patterns are also attested, namely TOP/APT/COM and TOP/INT/COM, in BP and IT.

(22)   *SIL: se for vinho importado /=TOP= eu tomo //=COM=
Ex. bfamdl04 [163]
'*SIL: if it's imported wine /=TOP= I'll drink it //=COM='

(23)   *BRU: isso aqui /=TOP= se ocê faz /=APT= é a primeira letra //=COM=
Ex. bfamcv04 [237]
'*BRU: this one here /=TOP= if you do it /=APT= it's the first letter //=COM='

(24)   *ALD: se uno riesce a comunicare con facilità con le persone /=TOP=
        riesce a fare il rappresentante //=COM= Ex. ifammn14 [15]
        '*ALD: if someone can communicate easily with people /=TOP= he can
        be a salesman //=COM='

(25)   *ANT: oh /=CNT= se mi vedi coll' ape /=TOP= 'un dire /=INT= guarda
        qui' bischeraccio va a picchiare agl' altri //=COM_r=
        Ex. ipubcv05 [35]
        '*ANT: hey /=CNT= if you see me with the Apecar /=TOP= don't say
        /=INT= look at that fool that goes rear-ending other people //=COM_r='

Once again, the semantic relation between the protasis in TOP(/APT) and the
apodosis in (INT/)COM is consistent with the pragmatic relation existing be-
tween TOP and COM. As it was observed for Time adverbial clauses, then, such
a consistency between the semantic and pragmatic dimensions could explain
why such adverbial clauses occur in this information pattern in the vast majority
of cases.

        A few occurrences of conditional sentences appear also in illocutionary pat-
terns, but mostly in chains of CMM. The only occurrence in COB was found in
BP (see (26)) and shows a postponed protasis:

(26)   *RUT: Nossa /=EXP= eu &co (/2)=EMP= eu adoeço /=COB= se me
        chamar //=COM= Ex. bfamcv02 [224]
        '*RUT: wow /=EXP= I &co (/2)=EMP= I fall ill /=COB= if he calls me
        //=COM='

(27)   *PAO: se t' ha lasciato i' bigliettino /=CMM= 'un' è assicurato //=CMM=
        Ex. ipubcv05 [60]
        '*PAO: if he left you his note /=CMM= he's not insured //=CMM='

Finally, IT data provide 8 occurrences of protases in PAR, as in (28), while BP
data present only one occurrence, showed in (29):

(28)   *ART: che [/1]=EMP= che /=SCA= lucida /=COM= la pelle /=APC= <
        se è troppo grinzosa > /=PAR= eccetera //=APC= Ex. ifamdl04 [105]
        '*ART: that [/1]=EMP= that /=SCA= polishes /=COM= the leather
        /=APC= if it's too wrinkled /=PAR= etcetera //=APC='

(29)　*EUG: ela regula aqui /=CMM= se cê quiser alargar /=PAR= <e aqui>
　　　　tem elástico //=CMM= Ex. bpubdl02 [119]
　　　　'*EUG: it can be regulated here /=CMM= if you want to stretch it
　　　　/=PAR= and here there's an elastic //=CMM=

PAR units perform a metalinguistic function, providing instructions about the interpretation of the utterance or part of it (see Table 1 in section 2); in this sense, by means of PAR the speakers provide a metalinguistic evaluation of what they are saying. In (28) and (29) *ART and *EUG are specifying additional information reinforcing what they are saying: it is because there can occur the eventuality of a too wrinkled leather that *ART is explaining the existence of a polishing machine; it is because *EUG imagines that the interlocutor could want to stretch the shoe tie that he is mentioning the possibility of regulating it.

Then, protases used in PAR do not display consistency with their semantic value in providing background context for the illocution (as it happens in TOP units), but rather in making reference to *eventual* circumstances that the speaker imagines and includes in the utterance in order to support what he has said from a metalinguistic point of view.

3.5　Cause adverbial clauses

Cause adverbial clauses display quite a different behavior from Time and Conditional.

First of all, in many cases the main subordinator *porque*/*perché* "because" does not introduce a dependent cause clause, but rather is implied by speakers to begin an utterance or a dialogic turn (cf. the notion of postponed *because*-clause as turn construction unit in Couper-Kuhlen 1996). Cresti (2005: 242-243) claims that coordinative conjunctions at the beginning of the utterance function as "pragmatic" coordinators. Raso & Mittmann (2012: 209-210) point out that both coordinative and subordinative conjunctions, when at the beginning of the utterance, display pragmatic functions (turn-taking, linking between autonomous utterances/speech acts, etc.).

Besides, it has been pointed out that *because*-clauses can project different types of causal linkage according to the different domains of the utterance on which they are operating (Sweetser 1990, Couper-Kuhlen 1996, Dancygier & Sweetser 2005; see also the distinction between *causal*, *explicative* and *focalizing* value of Italian *perché* in Acciardi 2010). Then, *direct* causal relations operate at the propositional level (SoA_2 takes place because of SoA_1), while *indirect* ones operate at the epistemic domain (the speaker accounts for why he has said

something in a previous clause) or at the speech act domain (the speaker accounts for why he performed a previous act of speech).

Actually, direct causal relation is found in the few *linearized* occurrences of *porque*/*perché*-clauses:

(30)   *ALO: agora /=INT_r= cê tá pagando porque ocê quer //=COM_r=
            Ex. bfammn03 [139]
            '*ALO: well /=INT_r= you are paying because you want to do so
            //=COM_r='

(31)   *TAM: e questa dirigente /=TOP= questa fra l' altro che conosce la realtà
            poggibonsese /=SCA= perché abita a Poggibonsi /=SCA= e ha due figlie
            adolescenti anche lei /=PAR= &he /=TMT= proponeva /=i-COM=
            appunto /=PHA= questo centro giovani e una libreria //=COM=
            Ex. ipubmn04 [21]
            '*TAM: and this manager /=TOP= this one that by the way knows
            Poggibonsi reality /=SCA= because she lives in Poggibonsi /=SCA= and
            she also has two adolescent daughters /=PAR= uhm /=TMT= she was
            proposing (to create) this youth center and one book shop //=COM='

A significant number of occurrences of the mentioned adverbial subordinator are used, alone, in dialogic units such as Discourse Connector (DCT: this is the most frequent dialogic unit where *porque*/*perché* appear), Incipit (INP), Phatic (PHA), Conative (CNT). Occurrences were found at the beginning of the utterance (see (32) and (33)) or within it (with the exception of INP, obviously; see (34) and (35)):

(32)   *SHE: não /=PHA= de conversar /=COB= de saber por que /=COB= que
            eles nũ [/1]=SCA= nũ [/1]=EMP= nũ gostam /=COB= ou o que que eles
            gostariam de aprender /=COB= o que que tá errado na [/1]=SCA= que
            eles vêem na minha aula /=COB= então eu já tô [/2]=SCA= já tô assim
            abrindo bem a minha mente em relação a isso //=COM=
            porque /=DCT= é lutar contra a maré //=COM=
            Ex. bpubmn01 [19.20]
            '*SHE: no /=PHA= about talking /=COB= about knowing why /=COB=
            they don't [/1]=SCA= don't [/1]=EMP= don't like it /=COB= or what
            they would like to learn /=COB= what is wrong in [/1]=SCA= that they
            see in may class /=COB= so I'm already [/2]=SCA= I'm already opening
            my mind with respect to it //=COM=
            because /=DCT= it's like fighting against the tide //=COM='

(33)   *PAO: allora /=INP= voglio dire /=PAR= se lei giustamente è &de
       [/1]=SCA= animalista /=TPL= lei è hhh buddista /=TPL= io sono
       /=SCA= &he /=TMT= cattolica /=TPL= e gli dico /=INT= guarda
       /=CNT_r= io /=TOP= sono della tre F /=CMM_r= però poi /=TOP=
       insomma /=PHA= dovresti convertirti //=CMM_r=
       perché /=DCT_r= insomma /=PHA= < i musulmani /=CMM_r= 'un
       hanno qui /=CMM_r= 'un hanno là > /=CMM_r= 'un hanno là
       //=CMM_r= Ex. ipubcv01_taglio [283.284]
       '*PAO: so/=INP= I mean /=PAR= if you are rightly of [/1]=SCA=
       animalist /=TPL= she is hhh buddhist /=TPL= I am /=SCA= ehm
       /=TMT= catholic /=TPL= and I tell him /=INT= look /=CNT= I /=TOP=
       come from the "Tre F" /=CMM_r= but eventually /=TOP= well /=PHA=
       you should convert yourself //=CMM_r=
       because /=DCT_r= well /=PHA = the Muslims /=CMM_r= they don't
       have this /=CMM_r= they don't have that /=CMM_r= they don't have
       that //=CMM_r='

(34)   *CAR: é uma história muito triste /=CMM= mas uma história muito
       bonita /=CMM= ela nũ [/2]=EMP= ela nũ me tirou /=SCA= o direito de
       ser mãe /=COB= eu também nũ direi [/2]=EMP= nũ tirei o direito dela
       ser mãe /=COB= até [/1]=EMP= então /= DCT= se ela quiser /=SCA=
       ligar aqui pra casa e falar se eu posso levar /=COB= eu levo /=COB= nũ
       tiro /=COB= o [/1]=EMP= o direito /=COB= mas /=DCT= nũ faço muita
       questão não //=COM=
       sabe /=PHA= eu nũ faço não /=COB= porque /=DCT= nós amamos
       /=COB= o nosso &f [/3]=EMP= nossos [/1]=EMP= nossa filha /=COB=
       eu e meu esposo /=COB= nós amamos < demais > +
       Ex. bfammn05 [100.101]
       '*CAR: it's a really sad story /=CMM= but a very nice story /=CMM=
       she didn't [/2]=EMP= she didn't take out of me /=SCA= the right to be
       mother /=COB= I also didn't dake [/]=EMP= I didn't take out her right
       to be mother /=COB= even [/1]=EMP= so /=DCT= if she wants /=SCA=
       to give us a call here at home and ask if I can take (the girl to hers)
       /=CMM= I take her /=COB= I don't take out /=COB= the [/1]=EMP=
       her right /=COB= but /=DCT= I don't insist in it so much //=COM=
       you know /=PHA= I don't insist in it /=COB= because /=DCT= we love
       /=COB= our &f [/3]=EMP= our [/1]=EMP= our daughter /=COB= me
       and my husband /=COB= we love her so much +'

(35)   *CLA: ci s' impauriva /=COB= perché /=DCT= cominciava a perde' il
           posto di lavoro //=COM= Ex. ifammn02 [102]
           '*CLA: people were scared /=COB= because /=DCT= they started losing
           their jobs //=COM='

When the DCT is used within the utterances, it always links COB units (see (34)
and (35)), that form Stanzas and are thus typical of monologic texts. The COB,
as we previously explained, are juxtaposed illocutionary units by which the
speaker builds his speech. The *because* subordinator in DCT, in these contexts,
seems to keep a certain degree of its direct cause semantics in the way he links
the COB: it is because *CAR and her husband love their adoptive daughter so
much that she doesn't insist in favoring contacts with her biological mother
(*CAR had just finished telling about an attempt that woman made to sell the
girl in order to buy drugs); it is because workers started to lose their job that they
were scared. In these cases, then, a relation of *direct* cause is still detectable.

        When the DCT is at the beginning of the utterance, on the other hand, such
causal relation with the preceding utterance is not so straightforward, and the
function of "pragmatic" connector (then, frequently, *indirect* cause) prevails. In
the first utterance of (32), *SHE says that she has already started to open her
mind with respect to talking with her students in order to understand their neces-
sities and opinions about her classes. Then, in the second utterance (starting with
a *porque* in DCT) she adds a justification/explanation of why *she thinks* (causal
linkage in epistemic domain) such an attitude is fundamental for her job, which
is that teaching in that context is "as hard as fighting against the tide".

        In (33) we can see that a reported ([TAG]_r) speech (in this case, an imagi-
nary conversation performed by means of a reported speech) is performed across
two different utterances, the second one starting with a *perché* in DCT. In this
case a relation of causality in the epistemic domain (*indirect* cause) is recog-
nizable between the two utterances: the (imaginary) Muslim interlocutor of
*PAO should convert himself to Catholicism because *she thinks* that Muslims
lack a series of characteristics she (as a catholic) believes are important.

        The second aspect in which Cause adverbial clauses display a different be-
havior from Time and Conditional is that the TOP/COM pattern is next to never
attested: two occurrences in BP, three in IT. Moreover, in this pattern the adver-
bial clause is always performed in COM:

(36)   *ALO: e eu sei que ea devia /=TOP= **porque** /=SCA= &he /=TMT= foi
       [/1]=EMP= foi &q [/1]=SCA= nas véspera d' eu vim embora //=COM=
       Ex. bfammn03 [108]
       '*ALO: and I know that she had to /=TOP= **because** /=SCA= uhm
       /=TMT= it was [/1]=EMP= it was &q [/1]=SCA= the night before I left
       //=COM='

(37)   *CLA: vogliono sterminare questa gente /=TOP= **perché** /=SCA= è sul
       confine //=COM= Ex. ifammn03 [12]
       '*CLA: they want to exterminate those people /=TOP= **because** /=SCA=
       they are on the border //=COM='

The most common information pattern for Cause clauses is in chains of CMM or
COB, and here the adverbial clause is always postponed as well, since
porque/*perché* introduce postponed causal clauses:

(38)   *CEL: mas cê nunca vai adivinhar nenhuma minha /=CMM= **porque** eu
       nũ tenho a mínima noção //=CMM= Ex. bfamcv04 [164]
       '*CEL: but you'll never guess any of mine /=CMM= because I'm really
       incompetent //=CMM='

(39)   *CLA: si mettan a parlare /=COB= **perché** c' era rimasto l' autista
       /=COM= con noi //=APC= Ex. ifammn03 [147]
       '*CLA: they start talking /=COB= **because** the chauffeur had remained
       /=COM= with us //=APC='

This fact is of particular interest. As Diessel (2005: 464) points out, causal ad-
verbial clauses (in both spoken and written corpora) often do not display iconici-
ty in their realization: SoA_2 is caused by a *preceding* SoA_1, but SoA_1 (represent-
ed therefore, by the *because*-clause) consistently occurs *after* SoA_2 (the main
clause). Interestingly, other causal subordinators such as *since* and *as* in English
(*como*, *já que* in BP, *visto che*, *dato che*, *dal momento che*, *siccome* in IT) intro-
duce preposed causal clauses, but their frequency in both spoken and written
corpora analyzed by Diessel (2005) is significantly lower than that of *because*
(*porque*/*perché*), as our data confirm as well.[15] Dancygier & Sweetser (2005:
134-135) observe that *because*-clauses differ from *since*- ones in providing, tipi-
cally, new information (see the difficulty of *porque*/*perché*-clauses in occurring

---

[15] Diessel (2005: 465) notes that in scientific articles the higher incidence of preposed causal
clauses is due to the fact that in this specific *genre* they are used in the expression of conclu-
sions or inferences: in this sense, their codification would be more iconic ($SoA_1 > SoA_2$).

in TOP),[16] while *since*-clauses are used to encode *known* causes (see also Ford 1993, who reports on the prosodic correlates – "comma" or "period" intonation, following Chafe 1988 - of these properties in the realization of such clauses in speech).

After this first survey on the information patterns in which adverbial clauses of Time, Condition and Cause are used in spoken BP and IT, it emerged that TOP/COM and chains of CMM or COB represent the most common ones. However, another interesting use of adverbial clauses emerged from our analysis, namely a sort of pragmatic insubordination.

3.6   Insubordination phenomena

Insubordination is defined as "the conventionalized main-clause use of what, on *prima facie* grounds, appear to be formally subordinate clauses" (Evans 2007: 367).

Insubordination describes a widely attested use of adverbial clauses in spoken BP and IT (approximately 30% on the total occurrences of adverbial clauses in the two IPIC corpora), that means, formally finite adverbial clauses occurring without their main clause in the same utterance, but forming nonetheless completely autonomous units.

To be more specific, it is important to distinguish between two types of insubordinated occurrences of adverbial clauses in the BP and IT corpora: "proper" insubordination phenomena, when it is not possible to retrieve the 'missing' main clause in the adjacent or close utterances, and "pragmatic" insubordination, when it is still possible to identify the main clause in the close context, but the adverbial clause constitutes, anyway, a pragmatically autonomous unit within the speech flow. Actually, with the exception of Conditionals, adverbial clauses in speech display mostly this second insubordinated use.

We have already observed that a textual cohesion exists across different utterances in the way discourse is constructed by the speakers, and that the adverbial subordinator *porque/perché* can be used in DCT at the beginning of a terminated sequence, in order to mark causal linkages between utterances, even if not at the propositional/syntactic level (see (32) and (33) above). It seems that Cause and Time adverbial clauses forming autonomous utterances are used mostly in this way (but Conditionals can as well, see (41) below), adding infor-

---

[16] This doesn't mean, anyway, that new information cannot occur in TOP unit, but rather that it is more *common* for given information to be used as background for the illocution.

mation which is linked at some level (epistemic, speech act) with the utterance where the main clause occurs. See the examples below:

(40)   *FLA: é /=CMM= quer dizer /=PAR= **a culpa é do arroz** /=CMM= né //=PHA=
       porque e' tá novo //=COM=
       *REN: **culpa é do arroz** //=COM=
       *FLA: é //=COM=
       quando fica ruim //=COM= Ex. bfamdl01 [562.566];
       Ex. bfamdl01 [563]; Ex. bfamdl01 [566]
       '*FLA: yeah /=CMM= I mean /=PAR= **it's rice's fault** /=CMM= isn't it //=PHA=
       because it's young //=COM=
       *REN: **it's rice's fault** //=COM=
       *FLA: yeah //=COM=
       when it gets bad //=COM='

(41)   *EUG: é o mesmo valor //=COM=
       todas duas //=COM=
       se cê quiser comprar as duas //=COM=
       eu fico mais feliz /=COM= viu //=PHA= Ex. bpubdl02 [219.224];
       Ex. bpubdl02 [223]
       '*EUG: it's the same price //=COM=
       both of them //=COM=
       if you want to buy the two //=COM=
       I'll be happier /=COM= you know //=PHA='

(42)   *MAX: secondo me ne dimostrava di più //=COM=
       *LIA: **perché** era grosso //=COM=
       *LIA: <e poi non aveva capelli> //=COM=
       Ex. ifamcv01_taglio [190.192]; Ex. ifamcv01_taglio [191]
       '*MAX: I think he looked older //=COM=
       *LIA: **because** he was large //=COM=
       *LIA: <and he also was bald> //=COM='

(43)   *PAO: allora /=INP= crei un rapporto /=SCA= con loro /=SCA= come SR //=COM=
       e **questi non capiscono più una mazza** //=COM=
       te lo dico io //=COM=
       anche se sei chiaro //=COM= Ex. ipubcv01_taglio [301.304];
       Ex. ipubcv01_taglio [304]

'*PAO: so /=INP= you create a relationship /=SCA= with them /=SCA=
as SR //=COM=
and **they really don't get it** //=COM=
I'm telling you //=COM=
even if you're clear //=COM='

In the examples above, Cause, Time, Conditional and Concessive adverbial clauses are performed in what we called "pragmatic" insubordination: their main clauses are traceable within the contiguous context, but they still perform autonomous illocutions, as it can be perceived listening to them separately (the separate audio file is provided as well). Their pragmatic autonomy is then clear, as (42) shows in a noticeable way: *LIA performs an insubordinated causal clause (indirect cause in epistemic domain),[17] whose traceable main clause is within the utterance pronounced by *MAX.

Protases, on the contrary, are more often used in the "proper" insubordinated construction, i.e. without the presence of the apodosis in the context. As Lombardi Vallauri (2000, 2004, 2010; see also the referred bibliography) pointed out, insubordinated protases ("free conditionals" in his terms) are cross-linguistically widespread in speech. This would follow from the metadiscoursive function they perform in projecting the conditional relation existing with the (omitted) main clause. In this way, protases would have an inherent pragmatic force that, in speech, makes conditionals versatile and useful in order to perform different illocutions (Lombardi Vallauri finds out Offer, Request, Reassurance of the addressee, Desire, among others). Examples (45)-(48) below show instances of "proper" insubordination in BP and IT:

(44)   *PAO: perché non è che al primo momento /=TOP= se anche ti pigliano
        il giornale //=COM=
        tu vuoi approfondire //=COM=
        giustamente //=COM= Ex. ipubcv01_taglio [298.300];
        Ex. ipubcv01_taglio [298]

---

[17] With respect to insubordinated Cause clauses, then, it has to be noted that the relation of indirect cause between two utterances is detectable independently from whether the adverbial subordinator *porque/perché* occurs alone in a DCT unit at the beginning of the second utterance (see examples (32) and (33) above) or not. This issue has certainly to be deepened. Vieira & Raso (in preparation) explore the prosodic and functional characteristics of DCT, trying to find a better description of this dialogic unit.

'*PAO: because it's not that in the first time /=TOP= if ever they buy the
journal from you //=COM=
you want to deepen //=COM=
of course //=COM='

(45)   *NIC: allora vediamo //=COM=
se te la memorizzi //=COM=
*CEC: sì //=COM= Ex. ifamdl17 [210.212]; Ex. ifamdl17 [211]
'*NIC: so let's see //=COM=
if you decorate it //=COM=
*CEC: yes //=COM='

(46)   *ANE: eh /=PHA= se cê não tiver um carrinho que [/1]=SCA= que sobe
aqui //=COM=
*CES: uhn //=COM=
é //=COM=
isso não é muito bom //=COM= Ex. bfamdl05 [38.41];
Ex. bfamdl05 [38]
'*ANE: uh /=PHA= **if you don't have a good car that** [/1] = **that
climbs here** //=COM=
*CES: uhm //=COM=
yeah //=COM=
this is not very good //=COM='

(47)   *LUZ: dá muito trabalho agora desativar //=COM=
andando /=TOP= nũ dá não //=COM=
*LAU: **mas e se ela explodir** //=COM=
*LUZ: explode não//=COM= Ex. bfamdl03 [190.192];
Ex. bfamdl03 [192]
'*LUZ: it's really hard to turn it off now //=COM=
while driving /=TOP= it's impossible //=COM=
*LAU: **but and if it explodes** //=COM=
*LUZ: it won't explode //=COM=

Listening to (45) and (46) it becomes clear that the utterances *tu vuoi ap-
profondire* and *allora vediamo* definitely are not possible main clauses for the
insubordinated protases (and that, for example, in (46) *se* introduces an insubor-
dinated conditional clause, and not an indirect question depending from the pre-
vious clause).

The formally adverbial clauses are performed with a prosodic profile that
carries an illocution (actually, different types of illocutions, depending on the

case). In this way, their lack of syntactic autonomy, within the same utterance, in cases where tracing their respective main clauses is still possible ("pragmatic" insubordination) or within the global context ("proper" insubordination), is overridden by virtue of their pragmatic autonomy as independent utterances.

This is interesting for two reasons: first, because it represents a significant part of adverbial clauses use in spoken BP and IT; secondly, because the lack of illocutionary force is a property that has been commonly associated to (adverbial) dependent clauses (Foley & Van Valin 1984, Haiman & Thompson 1984, Mathiessen & Thompson 1988, Lehmann 1988).

4.   Conclusions

In this survey on finite adverbial clauses in spoken BP and IT we found that the two languages display very similar features. We observed a strong reduction of the traditional inventory of adverbial relations in both languages, since the relevant ones in the two *minicorpora* of Italian C-ORAL-ROM and C-ORAL-BRASIL are Cause, Conditional and Time.

Secondly, we observed that finite adverbial clauses are rarely used as proper subordinate clauses, since the linearized occurrences are very uncommon. Comparing this with the fact that the vast majority of complement clauses, on the contrary, are performed in linearized configuration in both BP and IT, we think that the different degree of syntactic and semantic integration adverbial and complement clauses display is iconically reflected in their different realizations in speech.

Adverbial clauses appear mostly in TOP/COM and CMM/CMM or COB/COB patterns, with different informational/pragmatic functions. In the former, adverbial clauses are used in TOP in order to provide background information for the illocution in the COM hosting the main clause. In this sense, the semantics of Time and Conditional clauses is consistent with the pragmatic function they assume in such pattern, and there is a sort of correspondence between their syntactic and "pragmatic" subordination.

This is not the case of the majority of Cause clauses, since the *porque*/*perché*-clauses are used to encode new information: therefore, they are not the best candidate to appear in TOP (where given information occurs more commonly) and are actually used consistently in final position, as a means to encode causality with a backward scope ($SoA_2 \Leftarrow SoA_1$).

In CMM and COB patterns, on the other hand, adverbial and main clauses are performed in units that stand at the same hierarchical pragmatic level, for

they all carry illocutionary force. In these patterns, adverbial clauses display a higher degree of freedom in their distribution (pre/postponed). In chains of COB the adverbial clauses display functions related to discourse construction, while in CMM serve as cues for characterizing the type of logic relation existing between the various patterned illocutions. As regards Cause clauses, it was also noted that in specific patterns the causal relation between two utterances can still be detectable, but in a domain (epistemic, pragmatic) different from the propositional one.

Finally, a wide set of insubordination uses of finite adverbial clauses was verified in the two *minicorpora*. This particular use represents another instance of the pragmatic orientation of spoken language, since formally dependent clauses acquire pragmatic autonomy by virtue of specific prosodic profiles, and therefore can perform speech acts. Such autonomous speech acts can be textually (and, apparently, also from a syntactic point of view) linked with other utterances ("pragmatic" insubordination) or not ("proper" insubordination).

It is clear that the analysis of speech has to take into account the prosodic/pragmatic dimension in order to adequately describe the peculiar properties of spoken syntax, and adequate speech corpora represent a fundamental linguistic resource in order to do it.

Acknowledgments

References

Acciardi, D. 2010. *Funzioni connettivali nella lingua italiana scritta e parlata e analisi corpus-based del connettivo* perché. PhD diss. Firenze, Università degli Studi di Firenze.

Austin, J. L.1962. How to do things with words. Oxford: Oxford University Press.

Bossaglia, G. 2014. Interface entre sintaxe e articulação informacional na fala espontânea: uma comparação baseada em corpus entre português e italiano. *Caligrama* 19 (2): 35-60.

Barbosa, P. A. 2000. Syllable-timing in Brazilian Portuguese: Uma crítica a Roy Major. *DELTA*, 16 (2): 369-402.

Barbosa, P. A. 2000. Incursões em torno do ritmo da fala. Campinas: Pontes Editora.

Bick, E. 2000. The parsing system 'Palavras': Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework. Aarhus: Aarhus University Press.

Blanche-Benveniste, C. 2002. Macro-syntaxe et micro-syntaxe: les dispositifs de la rection verbale. In H.L. Andersen & H. Nølke (eds), *Macro-syntaxe et macro-sémantique*. Berne: Peter Lang, 95-118.

Chafe, W. 1984. How people use adverbial clauses. *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*: 437-449.

Chafe, W. 1988. Linking intonation units in spoken English. In J. Haiman & S. A. Thompson (eds), *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins, 1-27.

Couper-Kuhlen, E. 1996. Intonation and clause combining in discourse: the case of Because. *Pragmatics* 6 (3): 389-426.

Cresti, E. 1995. Speech act units and informational units. In E. Fava (ed), *Speech Acts and Linguistic Research*. Nemo: Padova, 89-107.

Cresti, E. 2000. *Corpus di italiano parlato: Introduzione*. Vol. 1. Firenze: Accademia della Crusca.

Cresti, E. 2005. Notes on lexical strategy, structural strategies and surface clause indexes in the C-ORAL-ROM spoken corpora. In E. Cresti & M. Moneglia (eds), 2005, 209-256.

Cresti, E. 2014. Syntactic properties of spontaneous speech in the Language into Act Theory: Data on Italian complement and relative clauses. In T. Raso & H. Mello (eds), *Spoken Corpora and Linguistic Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 365-410.

Cresti, E. & Moneglia, M. (eds). 2005. *C-ORAL-ROM: integrated reference corpora for spoken romance languages*. Amsterdam/Philadelphia: John Benjamins.

Cresti, E. & Moneglia, M. 2010. Informational patterning theory and the corpus-based description of spoken language. The compositionality issue in the topic-comment pattern. In M. Moneglia & A. Panunzi (eds), *Bootstrapping information from corpora in a cross-linguistic perspective*. Firenze: FUP, 13-45.

Cristofaro, S. 2003. *Subordination*. Oxford: Oxford University Press.

Dancygier, B. & Sweetser, E. 2005. *Mental Spaces in Grammar: Conditional Constructions*. Cambridge: Cambridge University Press.

Diessel, H. 2005. Competing motivations for the ordering of main and adverbial clauses. *Linguistics* 43 (3): 449-470.

Foley, W. A. & Van Valin Jr., R. D. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Ford, C. E. 1993. *Grammar in interaction: Adverbial clauses in American English conversations*. Cambridge: Cambridge University Press.

Ford, C. E. & Thompson, S. A. 1996. nteractional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns". *Studies in interactional sociolinguistics* 13: 134-184.

Ford, C. E., Fox, B. A. & Thompson, S. A. 2002. Constituency and the grammar of turn increments. In C. E. Ford, B. A. Fox & S. A. Thompson (eds), *The Language of Turn and Sequence*. Oxford: Oxford University Press, 14-38.

Givón, T. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4: 333-378.

Givón, T. 1990. *Syntax: A functional-typological introduction*. Volume II. Amsterdam/Philadelphia: John Benjamins.

Haegeman, L. 2004. The syntax of adverbial clauses and its consequences for topicalisation. *Antwerp Papers in Linguistics* 107: 61-90.

Haegeman, L. 2012. *Adverbial clauses, Main Clause Phenomena, and Composition of the Left Periphery: The Cartography of Syntactic Structures*. Oxford: Oxford University Press.

Haiman, J. 1983. Iconic and Economic Motivation. *Language* 59 (4): 781-819.

Haiman, J. & Thompson, S. A. 1984. 'Subordination' in Universal Grammar. In C. Brugmann & M. Macauley (eds), *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley: University of California Press, 510-523.

Halliday, M. A. K. 1985. *An introduction to functional grammar*. London: Arnold.

t'Hart, J.T., Collier, R. & Cohen, A. 1990. *A perceptual study on intonation: an experimental approach to speech melody*. Cambridge: Cambridge University Press.

Heath, J. 1985. Discourse in the field: clause structure in Ngandi. In J. Nichols, J. & A. C. Woodbury (eds), *Grammar inside and outside the clause: Some approaches to theory from the field*. Cambridge: Cambridge University Press, 89-110.

Hengeveld, K. 1993. Semantic type, factivity and the expression of adverbial clauses. In K. Hengeveld (ed), *The internal structure of adverbial clauses* (EUROTYP Working Papers V.5). Berlin: Mouton de Gruyter, 119-132.

Hengeveld, K. 1998. Adverbial clauses in the languages of Europe. In J. van der Auwera (ed), *Adverbial constructions in the languages of Europe* (Empirical approaches to language typology/Eurotyp 20-3). Berlin: Mouton de Gruyter, 335-419.

Hengeveld, K. & Wanders, G. 2007. Adverbial conjunctions in Functional Discourse Grammar. In M. Hannay & G. Steen (eds), *Structural-functional studies in English grammar: In honor of Lachlan Mackenzie*. Amsterdam/Philadelphia: John Benjamins, 211-227.

Hopper, P. & Thompson, S. A. 2008. Projectability and Clause Combining in Interaction. In R. Laury (ed), *Crosslinguistic Studies of Clause Combining: The Multifunctionality of Conjunctions*. Amsterdam/Philadelphia: John Benjamins, 99-124.

Kortmann, B. 1997. *Adverbial subordination: A typology and history of adverbial subordinators based on European languages*. Berlin: Walter de Gruyter.

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

Lehmann, Ch. 1988. Towards a typology of clause linkage. In J. Haiman & S.A. Thompson, (eds), *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins, 181-225.

Lombardi Vallauri, E. 2000. *Grammatica funzionale delle avverbiali italiane*. Roma: Carocci.

Lombardi Vallauri, E. 2004. Grammaticalization of syntactic incompleteness: Free conditionals in Italian and other languages. *SKY Journal of Linguistics* 17: 189-215.

Lombardi Vallauri, E. 2010. Free conditionals in discourse: the forming of a construction. *Lingvisticae Investigationes* 33(1): 50-85.

Longacre, R. E. & Thompson, S. A. 1985. Adverbial Clauses. In T. Shopen (ed), *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press, 171-234.

Martin, Ph. 2004. WinPitch Corpus: A text to Speech Alignment Tool for Multimodal Corpora. *Proceedings of the 4th International Conference on Language Resources Evaluation*. Lisbon, 26-28 may 2004, 537-540.

Matthiessen, Ch. & Thompson, S. A. 1988. The Structure of Discourse and 'Subordination'. In J. Haiman & S.A. Thompson, (eds), *Clause combining in grammar and discourse*. Amsterdam/Philadelphia: John Benjamins, 275-329.

Mello, H. 2014. Methodological issues for spontaneous speech corpora compilation: the case of C-ORAL-BRASIL. In T. Raso & H. Mello (eds), *Spoken Corpora and Linguistic Studies*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 27-68.

Mello, H., Raso, T., Mittmann, M. M., Vale, H. P. & Côrtes, P. O. 2012. Transcrição e segmentação do corpus C-ORAL-BRASIL: critérios de implementação e validação. In T. Raso & H. Mello (eds), *C-ORAL-BRASIL I. Corpus de referência do português brasileiro falado informal*. Belo Horizonte: UFMG, 125-176.

Miller, J. E. & Weinert, R. 1998. *Spontaneous Spoken Language*. Oxford-New York: Clarendon Press.

Mittmann, M. M. 2012. *O C-ORAL-BRASIL e o estudo da fala informal. Um novo olhar sobre o Tópico no Português do Brasil*. PhD diss. Belo Horizonte, Universidade Federal de Minas Gerais.

Moneglia, M. 2005. I corpora dell'italiano parlato di LABLITA: criteri di costituzione, unità di analisi e comparabilità dei dati linguistici orali. In E. Burr (ed), *Il parlato: teoria - corpora - linguistica dei corpora*. Atti del VI convegno SILFI. Duisburg, 28.06-02.07.2000, 213-232.

Moneglia, M. 2011. Spoken corpora and pragmatics. *Revista Brasileira de Linguística Aplicada* 11(2): 479-519.

Moneglia, M. & Cresti, E. 2006. C-ORAL-ROM: Prosodic boundaries for spontaneous speech analysis. In Y. Kawaguchi, S. Zaima & T. Takagaki (eds), *Spoken Language Corpus and Linguistics Informatics*. Amsterdam/Philadelphia: John Benjamins, 89-114.

Moneglia, M. & Raso, T. 2014. Notes on Language into Act Theory (L-AcT). In T. Raso & H. Mello (eds), *Spoken Corpora and Linguistic Studies*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 468-495.

Panunzi, A. & Gregori, L. 2012. DB-IPIC: an XML database for the representation of information structure in spoken language. In H. Mello, A. Panunzi & T. Raso (eds), *Pragmatics and Prosody: Illocution, Modality, Attitude, Information Patterning and Speech Annotation*. Firenze: Firenze University Press,133-150.

Panunzi, A.& Mittmann, M. M. 2014. The IPIC resource and a cross-linguistic analysis of information structure. In T. Raso & H. Mello (eds), *Spoken Corpora and Linguistic Studies*, Amsterdam/Philadelphia: John Benjamins Publishing Company, 129-151.

Raso, T. & Mello, H. (eds). 2012a. *C-ORAL-BRASIL I. Corpus de referência do português brasileiro falado informal*. Belo Horizonte: UFMG.

Raso, T. & Mello, H. 2012b. The C-ORAL-BRASIL I: Reference Corpus for Informal Spoken Brazilian Portugues. *Lecture Notes on Artificial Intelligence* 7243: 362-368.

Raso, T. & Mittmann, M. M. 2012. As medidas da fala. In T. Raso & H. Mello (eds), *C-ORAL-BRASIL I. Corpus de referência do português brasileiro falado informal*. Belo Horizonte: UFMG, 177-220.

Schmidtke-Bode, K. 2009. *A Typology of Purpose Clauses*. Amsterdam/Philadelphia: John Benjamins.

Stassen, L. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.

Sweetser, E. 1990. *From etymology to pragmatics. Metaphorical and cultural aspects of semantic structure.* Cambridge: Cambridge University Press.

Thompson, S. A. 2002. 'Object complements' and Conversation: Towards a Realistic Account. *Studies in Language* 26 (1): 125-164.

Thompson, S. A. & Couper-Kuhlen, E. 2005. The Clause as a Locus of Grammar and Interaction. *Discourse Studies* 7 (4-5): 481-505.

Thompson, S. A., Longacre, R. E. & Hwang, S. J.. 2007. Adverbial Clauses. In T. Shopen, (ed), *Language Typology and Syntactic Description*, Second Edition. Cambridge: Cambridge University Press, 237-300.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. 1985. *A comprehensive grammar of the English language*. Harlow: Longman.

Vieira, M. & Raso, T. In preparation. *Descrição das unidades dialógicas por parâmetros fonéticos no português brasileiro*.

Voghera, M. 1992. *Sintassi e Intonazione nell'Italiano Parlato*. Bologna: Il Mulino.