Anaphora, Connectives and Resetting: Prosodic and Pragmatic Parameters Interactions in the Marking of Discourse Structure

Cyril Auran[1,2] *& Daniel Hirst*[2]

1 – IUFM d'Aix-Marseille / 2 - Laboratoire Parole et Langage, CNRS UMR 6057, Aix-en-Provence, France
cyril.auran@lpl.univ-aix.fr ; daniel.hirst@lpl.univ-aix.fr

Abstract

This paper tackles the issue of the interaction of three types of linguistic cohesion markers. Automatic analyses of the prosodically annotated (British English) Aix-MARSEC corpus show that anaphoric pronouns and connectives, though often grouped into a general category of cohesion devices, do behave differently in relation with the phonetic realization of resettings. Anaphoric pronouns, more particularly, are demonstrated to interact with resettings in a hypothesized complex interplay of production and pragmatic constraints, whereas connectives are shown to have no significant effect on resettings.

1. Introduction

The aim of this paper is to shed some light on the issue of the complex interactions between prosodic and textual markers of cohesion in spoken British English discourse. More particularly, this study analyses the influence of the textual marking of cohesion (through the use of anaphoric pronouns and connectives) on the amplitude of resettings, thus aiming at bringing elements of answer to the questions raised in [12]. Our first hypothesis, therefore, is that interactions between anaphoric and prosodic cohesion markers do exist and can receive systematic explanations. The analyses in this work are based on the automatic extraction of data and parameters from the Aix-MARSEC corpus.

Section 2 focuses on the question of the nature of discourse as involving both a product and a process before presenting two competing views of discourse unity: More particularly, Halliday & Hasan's now classical conception is contrasted with views (mainly Charolles') which differentiate the three concepts of cohesion, coherence and connectivity in order to account for the unity of discourse. Most interesting for us in these approaches is the subtle (for Halliday & Hasan) or clear-cut distinction (for Charolles) operated between anaphoric pronouns and connectives.

Section 3 gives a closer look at three types of markers traditionally related to spoken discourse unity: anaphoric pronouns, connectives and prosodic markers. In this perspective, prosodic markers related to discourse topic units boundaries and internal structures are more particularly focused on.

The analysis of the Aix-MARSEC corpus is then presented in section 4. The Aix-MARSEC corpus is introduced and the data extraction and analysis are detailed. Results, finally, are given in the last part of this section.

Section 5, eventually, proposes a tentative explanation for the effects observed in this corpus study. The specific effects of anaphora-based cohesion are compared to those observed with connectives and receive a tentative explanation within a competing constraints framework.

2. Discourse and discourse unity

2.1. Discourse as both a product and a process

As already noted by Brown and Yule in their 1983 work, linguistic studies taking discourse as their object tend to fall in two categories ([2]: 23-24). The first category, or "text-as-product view", relies on an extension of the concepts traditionally used for the semantic-syntactic description of (often written) sentences to discourse. Halliday & Hasan's analysis of cohesive devices in English is an example of such an approach. The second category, "discourse-as-process", consists of studies in which "words, phrases and sentences which appear in the textual record of a discourse [are considered] to be evidence of an attempt by a producer (speaker / writer) to communicate his message to a recipient (hearer / reader)" ([2]: 24).

Though the actual marking of the unity which characterizes discourse does rely on linguistic/textual elements, it seems important to place the "product" aspect of discourse against the "process" background which gives rise to it. This position, adopted here and exemplified in [9], is consistent with a cognitive approach to discourse, and aims not only at analysing the complex relations which unite discourse forms and functions but also at shedding light on the psychological processes involved in the production and perception of spoken discourse.

2.2. Two views of the unity of discourse

The detailed analysis of the linguistic marking of cohesion given in [11] constitutes a landmark in text-oriented discourse analysis. In this study, Halliday & Hasan propose a definition of cohesion as a "semantic [concept referring] to relations of meaning that exist within the text, and that define it as a text" ([11]: 4). The authors, while rejecting the concept of text as a linguistic structure, insist on the linguistic status of cohesion and further split the general category of cohesion into two sub-categories: lexical cohesion and grammatical cohesion.

Lexical cohesion, on the one hand, is described as "the cohesive effect achieved by the selection of vocabulary" ([11]: 274).

Grammatical cohesion, on the other hand, is obtained through the use of four types of linguistic marking involving grammatical items: reference, substitution and ellipsis, and conjunction. Anaphoric pronouns and connectives, which constitute the focus of this paper, are respectively parts of the 'reference' and "conjunction" subcategories.

Other (less text-centred) conceptions can be found in [8] or [3] who propose to rely on several characteristics (and not solely on cohesion) in order to capture the wholeness of discourse. Among the seven characteristics proposed by De Beaugrande & Dressler, Charolles keeps the first two (cohesion and coherence) and introduces connectivity. Coherence, cohesion and connectivity are considered as independent properties centred either on discourse product (cohesion and connectivity) or discourse processes (coherence).

In this perspective, Halliday & Hasan's concept of cohesion, redefined as dealing with the "marking of relations between utterances or utterance constituents" ([3]: 53, our translation) is contrasted with the concept of coherence, related with the interpretability of discourse: "Coherence is not a characteristic of texts [...]. The need for coherence, on the contrary, is a sort of a-priori mode of discourse reception" ([3]: 55, our translation). Connectivity, finally, constitutes a separate category and has to do with logical-semantic

relations (marked by connectives) between propositions and speech acts.

These two views of the exact role of cohesion in the unity of discourse obviously rely on the two basic conceptions of discourse mentioned above: when discourse is equated with text ("text-as-product view"), cohesion is regarded as its defining characteristic; whereas a more global view (involving both the processes and the product of discourse) treats cohesion as only one of the factors determining its unity.

Although both these approaches lay the emphasis on the importance of the formal linguistic ties which contribute to the unity of discourse, it seems interesting to explore Charolles' distinction between the role played by connectives and that played by cohesion markers (more particularly anaphoric pronouns in our case).

Our second hypothesis is that, if indeed anaphoric pronouns and connectives belong to two independent categories (cohesion vs. connectivity, both being centred on discourse product), then prosodic markers of cohesion should interact with other phenomena within the same category (i.e. anaphoric pronouns within the cohesion category), but not with phenomena belonging to the other category (i.e. connectives).

3. Discourse unity markers

3.1. Anaphoric pronouns

Anaphoric pronouns undoubtedly are some of the most typical discourse cohesion marks. Indeed, according to the approach, these marks are considered as "endophoric personal referents" ([11]), as members of "anaphoric chains" (cf. [4]: 204-205) or as expressions pointing to "highly accessible referents" (cf. for instance Ariel's or Gundel's work and [5] for a detailed overview); all these conceptions eventually boil down to the same fundamental property: anaphoric pronouns are mainly used by speakers to refer to already salient referents and seldom to introduce new referents into the discourse. In this respect, anaphoric pronouns permit the thematic preservation ([7]) necessary for discourse to be cohesive.

3.2. Connectives

Connectives are other well known discourse cohesion devices. Halliday & Hasan ([11]: chapter 5), for instance, propose to further categorize this class into 3 groups: adverbs (including simple adverbs and compound adverbs in "-ly" and "there-"), other adverbs (such as "furthermore") and prepositional phrases (e.g. "on the contrary") and prepositional expressions containing the demonstrative "that" (e.g. "instead of that"). Regarded as "procedural" markers within a Relevance Theoretic framework (see [12] and [18] for example), or items expressing "certain meanings which presuppose the presence of other components in the discourse" [11]: 226), their main function unanimously seems to relate to the integration of elements in the discourse context.

3.3. Prosodic markers

Among the numerous works tackling the issue of prosody as a marker of cohesion in discourse, [2] and [19] are particularly interesting for us as they question the prosodic marking of discourse topic units both in their boundaries and their internal structuring. For obvious reasons of space, however, only topic unit boundaries will be dealt with in this paper.

According to [2], topic-shifts in spoken discourse are prosodically marked as the boundaries of "structural units of spoken discourse which take the form of 'speech paragraphs' and have been called paratones" ([2]: 100-101). Even if, as in [13], we do not fully agree with this strict hierarchy position, numerous studies have demonstrated that tone groups do fit into some kind of hierarchic structure. We will more particularly focus on major and minor tone groups here as representatives of two levels in such a hierarchic organisation. Phonetic features related to the beginnings and ends ([19]) of major units have been studied extensively; most noticeable are the extra high (F0) onset values recorded at their beginnings ([2], [19] and [6]): this phenomenon is known as "pitch reset" or "resetting"; such elements as very low pitch, loss of amplitude, lengthy pauses ([2]) and creaky voice ([19]: 57), on the other hand, have been shown to be frequent correlates of major unit endings.

4. Corpus study

4.1. The Aix-MARSEC Corpus

The Aix-MARSEC corpus, which was used in this study, constitutes a second evolution from the original SEC corpus, the MARSEC corpus constituting the first one.

The original SEC (*Spoken English Corpus*) is a collection of BBC recordings from the 1980s ([15], [16]). The data represent more than five and a half hours of natural-sounding British English from 53 different speakers. The corpus contains about 55.000 orthographically transcribed words, manual prosodic annotation of all the recordings (G. Knowles and B. Williams, [16]) and tagging and parsing of the data using the CLAWS I algorithm.

The SEC was subsequently modified to facilitate computer use and became the MARSEC (*MAchine Readable Spoken English Corpus*); the first change consisted in manually aligning the word and (minor-major) intonation unit boundaries with the sound. Second, some of the tonetic stress marks (TSM) were changed into ASCII symbols in order to have a computer compatible set of TSM ([17]).

The Aix-MARSEC corpus, finally, designed within the Aix-MARSEC Project ([1]), constitutes the latest development stage of the data. Automatic procedures, more particularly, were used to transcribe the 55.000 words of the corpus into phonemes (SAMPA and IPA alphabets), to optimize and align this transcription with the speech signal and to group and code phonemes in sub-syllabic constituents (onset, nucleus and coda), syllabic units, and rhythmic groups using Abercrombie and Jassem models. Phonemes were also grouped into words according to their initial phonemic composition. All constituents were eventually grouped into minor and major intonation units

The coding of intonation was carried out using the MOMEL-INTSINT methodology developed in Aix-en-Provence. The MOMEL algorithm, more particularly, aims at modelling the actual F0 curve so that any microsegmental characteristics (the "micro-prosodic component") should be factored out [10]. The resulting curve is thus similar to that found on a sequence of entirely sonorant segments and constitutes the "macro-prosodic component" [14]. The system involves quadratic spline functions which allow us to treat a sequence of target points as an appropriate phonetic representation of F0 curves.

4.2. Data extraction and analysis

In order to test the hypothesis proposed at the end of section 2.2, we extracted onset F0 values for all the tone groups which contained either a third person anaphoric pronoun or a connective. The whole of the Aix-MARSEC was used, except for the "E" type of recordings ("Daily Service"), the quality of which could not guaranty accurate F0 detection).

The experimental design thus displays one dependent variable (onset F0 value) and 2 independent variables:

- type of tone group ("major" vs. "minor", [16], [19]);
- anaphoric marker ("presence" vs. "absence" for he/him/his, she/her and they/them/their);

- connective marker ("presence" vs. "absence" for "and", "but" "however" and "then").

Onset F0 values were extracted from the modelled F0 curves obtained with the MOMEL algorithm. More particularly, mean modelled F0 was computed for the first stressed syllable of each tone group (a method comparable to that used in [19]).

The type of tone group ("tone group factor": "major", vs. "minor") was taken into account in order to verify the traditional correspondence ([2], [6], [19]) established between topic units (realized as major prosodic units) and significantly higher onset F0 values. The reproduction of these observations for our data thus confirms the experimental accuracy and validity of our fully automatic methodology.

The presence or absence of anaphoric pronouns and connectives ("anaphoric marker factor" and "connective marker factor" respectively) constitutes the main factor studied here.

The data were automatically extracted using Perl scripts and statistical tests were computed using the R software. A total number of 12,272 tone groups were analysed, representing approximately 5 hours of speech.

Given that, even after logarithmic transform, the actual distribution of onset F0 values diverged from a normal distribution (kurtosis = 4.54 and 0.13 after log transform, and skewness = 1.73 and 0.5 after log transform), all ANOVA results were subsequently checked using two-sample Kolmogorov-Smirnov tests (KST) which do not presuppose normal distribution of the data. More particularly, transitive and intransitive binary comparisons were computed in this respect.

4.3. Results

The analysis of the effect of the tone group factor on the resetting phenomenon confirmed the relation traditionally noted between higher onset F0 values and major units beginnings. Onset values for major tone groups (figure 1) were significantly (ANOVA: F=513.7, p<2e-16) higher (approximately 4.5 semitones (ST) for both mean and median values) than onset values for minor tone groups.

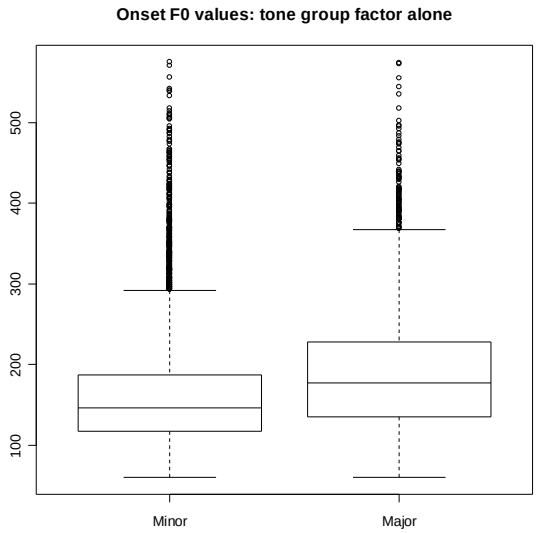

Figure 1: Onset F0 values (tone group factor)

The effect of the anaphoric marker factor (figure 2) was also found to be significant (ANOVA: F=54.94, p=1.321e-13). The observed 3.9 ST difference measured between "non-anaphoric onsets" and "anaphoric onsets" means ("anaphoric onsets" being higher) was confirmed to be significant using a KST (p<2.2e-16).

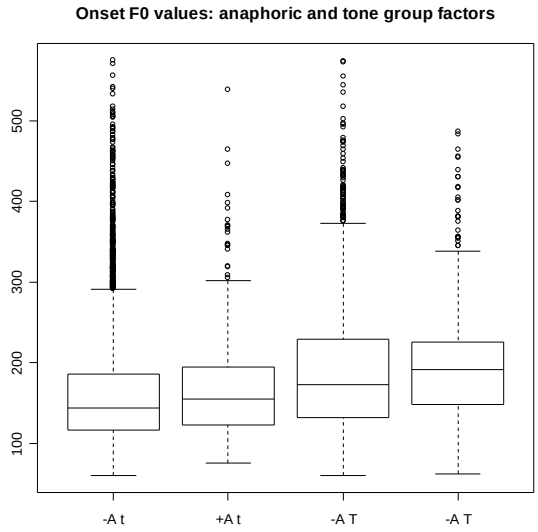

Figure 2: Onset F0 values for both anaphoric and tone group factors.

When all connectives were taken into account simultaneously, the effect of the connective marker factor was found to be significant (ANOVA: F=22.95, p=1.688e-6) for minor tone groups, but not significant (ANOVA: F=1.888, p=0.1696) at major tone group boundaries.

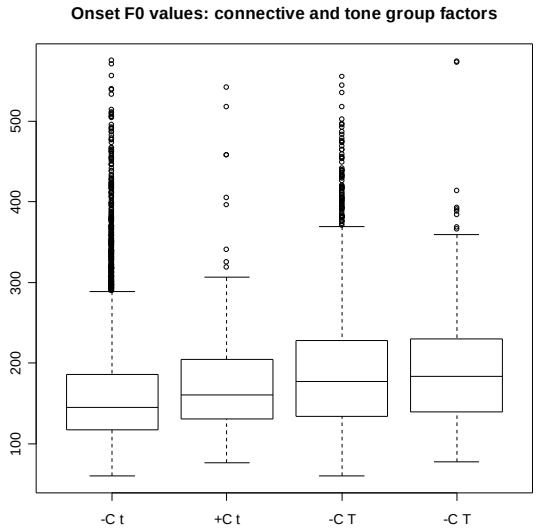

Figure 3: Onset F0 values for both connective and tone group factors.

However, of all the connectives taken in isolation, only "but" had a significant effect on onset F0 values. Table 1 presents the results for the relevant two-factor ANOVAs (connective and tone group factors). More particularly, p-values are detailed for each factor ("C" = connective, "TG" = tone group and "C : TG" = interaction). Only significant effects (with a 0.01 threshold) appear in bold face.

Table 1: Two-way ANOVA p-values for each connective

Connective	p-values		
	C	TG	C : TG
And	0.843	**< 2.2e-16**	0.676
But	**4.04e-9**	**< 2.2e-16**	0.0249
However	0.712	**< 2.2e-16**	0.631
Then	0.271	**< 2.2e-16**	0.854

5. Discussion

The preceding results call for two particular comments. First, the analysis of the anaphoric factor effects confirms our main hypothesis: the use of anaphoric cohesion marks influences resettings in a significant way. However, the influence observed may seem quite unexpected: the presence of anaphoric pronouns, which constitutes a reinforcement of local cohesion, is not accompanied by lower onset values (which would reinforce local cohesion even more), but, on the contrary, by significantly higher onset values.

This phenomenon may be related to the contrary constraints governing resetting phenomena themselves: indeed, one may consider that the actual value for a given onset results from a balance between pragmatic discourse constraints related to cohesion marking ("decrease onset to mark cohesion") and production and planning constraints ("increase onset to favour declination on the tone group span"; cf. [6]: 43). Both these constraints are exemplified by the higher onset values measured for major tone groups (compared to minor ones) and by the strong link noticed between these onset values and the actual duration of the tone groups.

The influence of anaphoric cohesion marks on resettings could thus be interpreted as resulting from the loosening of the pragmatic discourse constraints (which otherwise "increase" onset values) due to the cohesion guarantee provided by anaphoric marks; production and planning constraints thus gain relative influence and induce the observed increase in onset values.

Secondly, it is interesting to note that, apart from "but", connectives seem not to influence resetting phenomena. This constitutes an argument in favour of our second hypothesis, derived from Charolles' view attributing anaphoric markers and connectives to distinct and independent discourse unity categories. Indeed, both anaphoric pronouns and resetting phenomena are classified within a single "cohesion" category and do interfere with one another; connectives, on the contrary, belong to a category of their own and do not influence resetting phenomena. These two categories, therefore, seem to behave independently, which justifies the twofold division proposed by Charolles among text-centred discourse unity marks.

6. Conclusion and perspectives

Basing the analyses of this preliminary study on the automatic extraction of data from the Aix-MARSEC corpus ([1]), we have shown that interesting interactions between pragmatic and prosodic parameters do take place in spoken British English discourse. Cohesion prosodic marks such as resetting phenomena have more particularly been demonstrated to adapt to modifications in other linguistic phenomena (anaphoric cohesion) related to cohesion in discourse. This, together with the arguments presented here consistent with Charolles' view, seems to constitute interesting supplementary elements in favour of a more global approach to spoken discourse, taking into account both its aspects as product and psychological process in order to improve our understanding not only of prosody and discourse but also of the psychological constraints governing the use of language in spontaneous interaction.

Further research, obviously, remains to be made, notably concerning the specific behaviour and function of "but" (especially as opposed to both "and" and "however"), but also about a closer analysis of resetting as a relative value, calculated, for instance, as a speaker-normalised ratio or differential. Further analyses, finally, are under way, taking into account the exact reference of anaphoric pronouns and span of connectives in order to better qualify the influence of inter-tone-group discourse unity marks alone.

7. References

[1] Auran, C.; Bouzon, C. ; Hirst, D.J., The Aix-MARSEC project: an evolutive database of spoken British English. *Speech Prosody 2004*, March 23-26, 2004, submitted.

[2] Brown, G.; Yule, G., 1983. *Discourse Analysis*. Cambridge: C.U.P.

[3] Charolles, M., 1988. Les études sur la cohérence, la cohesion et la connexité depuis la fin des années 1960. *Modèles Linguistiques*, 10 (2). 45-66

[4] Chastain, C., 1975. Reference and context. In Gunderson (ed.): *Language, Mind and Knowledge*. Minneapolis: University of Minnesota Press, 194-269.

[5] Cornish, F., 1999. *Anaphora, Discourse and Understanding. Evidence from English and French*. Oxford: Clarendon Press.

[6] Couper-Kuhlen, E., 2001, Interactional prosody: High onsets in reason-for-the-call turns. *Language in Society* 30, 29-53.

[7] Danes, F. (ed.), 1974. *Papers on functional sentence perspective*. Prague: Publishing House of the Chechoslovak academy of sciences.

[8] De Beaugrande, R.A ; Dressler W.U., 1981. *Introduction to Text Linguistics*. London: Longman.

[9] Di Cristo A.; Auran C.; Bertrand R.; Chanet C.; Portes C., 2003. An integrative approach to the relations of prosody to discourse: towards a multilinear representation of an interface network. *Prosodic Interfaces*, March 27-29, Nantes.

[10] Di Cristo, A.; Hirst, D.J., 1986. Modelling French micromelody: analysis and synthesis. In Kholer (ed.): *Prosodic Cues for Segments* (*Phonetica* 43, 1-3), 11-30.

[11] Halliday, M.A.K.; Hasan, R., 1976. *Cohesion in English*. London: Longman.

[12] Higashimori, I.; Wilson, D., 1996. Questions on Relevance. *UCL Working Papers in Linguistics* 8, 1-14.

[13] Hirst, D., 1998. Intonation in British English. In Hirst, D.J.; Di Cristo, A. (eds.): *Intonation Systems: a Survey of Twenty Languages*. Cambridge University Press.

[14] Hirst, D.; Di Cristo, A.; Espesser, R., 2000. Levels of Representation and Levels of Analysis for the Description of Intonation Systems. In Horne, M. (ed.): *Prosody : Theory and Experiment. Text, Speech and Language Technbology*, *14*. Kluwer Academic Publishers, 51-87.

[15] Knowles, G.; Wichmann, A.; Alderson, P., 1996. *Working with Speech: perspectives on research into the Lancaster/IBM Spoken English Corpus*. England: Longman.

[16] Knowles, G.; Williams, B.; Taylor, L., 1996. *A Corpus of Formal British English Speech*. London: Longman.

[17] Roach, P.; Knowles, G.; Varadi, T.; Arnfield, S., 1993. MARSEC: A machine readable Spoken English corpus. *Journal of the International Phonetic Association* 23(2), 47-53.

[18] Wilson, D.; Sperber, D., 1993. Linguistic form and relevance. *Lingua* 90, 1-25.

[19] Wichman, A., 2000. *Intonation in Text and Discourse*. London : Longman.